Alma Mater Studiorum — Università di Bologna

DOTTORATO DI RICERCA IN ASTROFISICA

Ciclo 36

Settore Concorsuale: 02/C1 - ASTRONOMIA, ASTROFISICA, FISICA DELLA TERRA E DEI PIANETI
Settore Scientifico Disciplinare: FIS/05 - ASTRONOMIA E ASTROFISICA

Galaxy cluster cosmology in KiDS, Planck, and *Euclid* surveys

Presentata da: Giorgio Francesco Lesci

Coordinatore di Dottorato
Andrea Miglio

Supervisore
Federico Marulli

Co-supervisore
Lauro Moscardini

Esame finale anno 2024

Contents

Abstract

The large-scale distribution of matter in the Universe plays a crucial role in shedding light on the fundamental properties of dark matter, dark energy, and gravity on cosmic scales. Over recent decades, numerous multi-band surveys have been conducted to assemble representative samples of cosmic tracers. To achieve an accurate and precise understanding of the Universe on its largest scales, a combination of diverse methods and observations is employed. Among the primary cosmic tracers that can provide answers to the most profound questions in cosmology and fundamental physics, galaxy clusters assume a pivotal role. Reaching up to 10^{15} solar masses and radii up to 5 Mpc, galaxy clusters trace the deepest virialized potential wells of dark matter in the present Universe, lying in correspondence of the nodes of the cosmic web.

Since galaxy clusters mark the transition between linear and nonlinear gravitational perturbations, their formation and evolution can be theoretically described with high accuracy. Due to the dominant role of dark matter within galaxy clusters, it is sufficiently accurate, considering existing observational uncertainties, to model the formation and evolution of clusters through N-body dark-matter-only simulations. These simulations serve to calibrate both the mass function, namely the abundance as a function of mass and redshift, and mass profiles of galaxy cluster haloes. Having a theoretically calibrated halo mass function allows for the modelling of observed galaxy cluster counts, which, in turn, provide valuable constraints on fundamental cosmological parameters. Indeed, cluster counts are among the most effective cosmological probes, providing precise constraints on parameters like the matter density parameter, $\Omega_{\rm m}$, the square root of the mass variance at redshift $z = 0$ for scales of 8 h^{-1}Mpc, σ_8, and the parameters of the dark energy equation of state. Additionally, robust constraints on these cosmological parameters can be derived from measurements of cluster correlation functions and power spectra, commonly referred to as cluster clustering.

In this Thesis, we present a series of works that encompass the fundamental steps of cosmological analyses based on galaxy clusters, spanning from mass calibration to deriving cosmological constraints through counts and clustering. Throughout all these investigations, we assumed a standard Λ-cold dark matter (ΛCDM) cosmological model. Firstly, we focus on the work by Lesci et al. (2023) on the 3D two-point correlation function (2PCF) of the galaxy cluster sample by Planck Collaboration XXVII (2016), built up using observations of the Sunyaev-Zeldovich (SZ) effect. The masses of these clusters are expected to be underestimated, as they are derived from a scaling relation calibrated through X-ray observations. Recent studies, including Planck Collaboration XXIV (2016) and Planck Collaboration VI (2020), have demonstrated the need for very high values of the Planck mass bias, referred to as $b_{\rm SZ}$, to avoid discrepancies between the σ_8 constraints obtained from Planck cluster counts and CMB observations. Specifically, simulations and weak-lensing observations suggest $(1 - b_{\rm SZ}) \sim 0.8$, while reconciling cluster counts with CMB cosmological results requires $(1 - b_{\rm SZ}) = 0.62 \pm 0.03$. In our recent work, Lesci et al. (2023), we determined $(1 - b_{\rm SZ}) = 0.62 \pm 0.12$ using the Planck cluster 2PCF and assuming priors on cosmolo-

gical parameters consistent with Planck CMB results. This constraint agrees with the one obtained from the combination of CMB and cluster count observations, with this highlighting the need for very low $(1 - b_{\rm SZ})$ values to reconcile cluster statistics with CMB results. Lastly, by applying the same $b_{\rm SZ}$ priors as Planck Collaboration XXIV (2016), which are based on weak-lensing measurements, we obtained constraints on $\Omega_{\rm m}$ from the 2PCF. These constraints are not only consistent but also competitive, in terms of uncertainties, with those derived from cluster counts. Notably, our analysis revealed a reduction of approximately 30% in the uncertainty of $\Omega_{\rm m}$ when combining cluster counts and 2PCF. However, it is important to note that σ_8 remains unconstrained in our analysis.

Furthermore, in this Thesis we detail the works by Lesci et al. (2022a) and Lesci et al. (2022b), where we analysed the cluster counts and 2PCF, respectively, of the galaxy clusters detected by Maturi et al. (2019) in the third data release of the Kilo Degree Survey (KiDS-DR3, de Jong et al. 2017). Specifically, the cluster detection was carried out with the Adaptive Matched Identifier of Clustered Objects (AMICO, Bellagamba et al. 2018) algorithm. The sample covers an effective area of 377 deg^2, for a total of 7988 galaxy clusters. We assessed the incompleteness and the impurities of the cluster sample by exploiting a mock catalogue developed by Maturi et al. (2019), and we corrected our data accordingly. Both cluster counts and 2PCF analyses are based on the weak-lensing measurements by Bellagamba et al. (2019), which were fundamental for deriving the scaling relation between cluster mass and richness. In Lesci et al. (2022a), we considered a subsample containing 3652 galaxy clusters having an intrinsic richness of $\lambda^* \geq 20$ and redshift $z \in [0.1, 0.6]$. From the joint analysis of cluster counts and weak lensing, we found $\Omega_{\rm m} = 0.24^{+0.03}_{-0.04}$, $\sigma_8 = 0.86^{+0.07}_{-0.07}$, and $S_8 \equiv \sigma_8(\Omega_{\rm m}/0.3)^{0.5} = 0.78^{+0.04}_{-0.04}$, which are competitive constraints with the results from recent cosmic shear (Troxel et al. 2018; Hikage et al. 2019; Asgari et al. 2021), cluster counts (Bocquet et al. 2019; Costanzi et al. 2019), and CMB (Hinshaw et al. 2013; Planck Collaboration VI 2020) analyses. We also showed that the inclusion of cluster counts significantly improves the constraints on the mass-richness scaling relation derived from weak lensing only. In the work by Lesci et al. (2022b) on the 2PCF of the AMICO KiDS-DR3 clusters, we adopted the same redshift selection used for the counts' analysis, namely $z \in [0.1, 0.6]$. On the other hand, we chose a less restrictive threshold in richness, that is $\lambda^* > 15$, which guarantees a sample purity close to 100%, since our clustering model does not rely on sample completeness estimates. This results in 4934 galaxy clusters. Assuming the scaling relation constraints derived by Lesci et al. (2022a) as priors, we derived $\Omega_{\rm m} = 0.28^{+0.05}_{-0.04}$, $\sigma_8 = 0.82^{+0.14}_{-0.12}$, and $S_8 = 0.80^{+0.08}_{-0.08}$, which are consistent within 1σ with the results obtained from CMB experiments and from the most recent analyses of the late Universe. In addition, by fixing the cosmological parameters to the values derived by Planck Collaboration VI (2020, Table 2, TT, TE, and EE+lowE), we derived a robust constraint on the normalisation of the mass-richness relation. This confirms the key role of clustering in constraining cluster masses.

The fourth work presented in this Thesis is the one by Euclid Collaboration: Lesci et al. (2023). In this analysis, we introduce a novel approach to establish galaxy colour-redshift relations for the purpose of cluster weak-lensing analyses, regardless of the specific photometric bands in use. This method optimises selection completeness while maintaining a defined purity threshold. A notable aspect of our work is the presentation of colour-redshift relations as a continuous function of the lens redshift, denoted as $z_{\rm l}$. This is particularly valuable in cluster weak-lensing analyses, where accurate background selections are paramount, given that foreground and cluster member contamination can significantly affect the measured lensing signal. Based on the galaxy sample by Bisigello et al. (2020), we calibrated two colour selections, one relying on the ground-based *griz* bands (referred to as *griz* selection), and the other including the *griz* and *Euclid* $Y_{\rm E}J_{\rm E}H_{\rm E}$ bands (referred to as *griz*$Y_{\rm E}J_{\rm E}H_{\rm E}$ selection).

These selections show a purity exceeding 97%. The *griz* selection achieves a completeness of up to 84% within the lens redshift range $z_\mathrm{l} \in [0.2, 0.8]$. With the $griz Y_\mathrm{E} J_\mathrm{E} H_\mathrm{E}$ selection, the completeness improves by as much as 25 percentage points, enabling a z_l range that extends up to $z_\mathrm{l} = 1.5$. Notably, these calibrated colour selections prove their robustness against variations in sample limiting magnitudes and redshift ranges. The *griz* selection delivers compelling results when applied to real external datasets, even across different photometric aperture definitions sourced from various ground-based telescopes. The *griz* selection also exhibits larger purity at high redshifts and increased completeness at low redshifts compared to colour selections found in literature. Additionally, we demonstrated the consistency of our calibrated colour selections through an analysis of a simulated *Euclid* galaxy catalogue. Our work also shows that the calibrated colour selections produce reliable results even when data from a single band is missing from the ground-based dataset. Importantly, these selections do not introduce biases beyond the 1σ uncertainty in the multiplicative shear bias.

In addition, we present the preliminary work by Lesci et al. (in prep. b) on the weak-lensing mass calibration of the AMICO clusters detected by Maturi et al. (in prep.) in the fourth data release of KiDS (KiDS-1000, Kuijken et al. 2019). The cluster sample covers an effective area of 840 deg^2, for a total of 24 396 clusters with signal-to-noise S/N > 3.5 and within the redshift range $z \in [0.1, 0.8]$. We jointly modelled the stacked cluster weak-lensing signal obtained in bins of redshift and richness, deriving constraints on the mass-richness and concentration-mass scaling relations. We accounted for the main biases affecting weak-lensing analyses, such as halo miscentring, galaxy photo-z errors, impurities due to the background selection, halo orientation, and projection effects. This mass calibration will enable the cosmological analyses based on cluster counts and clustering, from which we expect remarkable improvements in the results compared to those derived by Lesci et al. (2022a) and Lesci et al. (2022b) in KiDS-DR3. Indeed, the AMICO KiDS-1000 cluster sample contains three times the objects in the DR3 sample, and the maximum sample redshift increased from $z = 0.6$ to $z = 0.8$.

The statistical analyses presented in this Thesis are based on CosmoBolognaLib (CBL, Marulli et al. 2016), a large set of *free software* C++/Python libraries, that provide an efficient numerical environment for cosmological investigations of the large-scale structure of the Universe. A relevant fraction of the numerical tools used in this work have been implemented in CBL by the author of this Thesis. Among the most relevant ones, we have the Poissonian likelihood including the contribution of the super-sample covariance, the modelling of cluster counts and clustering as a function of the mass proxy, and the measure, stacking, and modelling of galaxy cluster weak-lensing profiles.

Introduction

In accordance with the predictions of the standard cosmological model, known as Λ-cold dark matter (ΛCDM), we inhabit an expanding Universe characterized by two principal constituents: dark energy, described through the cosmological constant Λ, causing the accelerated expansion of the Universe and making up about 68% of the total energy and matter content of the cosmos, and cold dark matter, constituting approximately 27%. In contrast, only around 5% of the Universe consists of ordinary matter, which, despite its relatively small proportion, plays a significant role in astronomical and cosmological research as it forms the luminous side of the Universe.

Based on the ΛCDM model, we interpret the large-scale structure of the Universe as a consequence of the growth of primordial anisotropies in the matter density field, taking place within an expanding space. Such anisotropies originated from quantum fluctuations created at the end of the inflationary era, representing an exponential expansion of the Universe. The inflationary era is at the foundations of the so-called hot Big Bang model, depicting the evolution of the Universe as a thermal history. In this framework, the Universe cools down over cosmic time, and this temperature decrease allows the gravitational growth of the primordial fluctuations. Among the most compelling pieces of evidence supporting the Big Bang theory, we have the cosmic microwave background (CMB). Notably, the CMB is the black-body radiation that began to freely propagate following the first recombination of free ions into neutral hydrogen, and it provides a picture of the matter density field 379 000 years after the Big Bang. As such, CMB observations also provide robust information on the matter-energy content of the Universe.

Observations of the late Universe complement those based on CMB, also serving as a test of the evolutionary models based on the ΛCDM framework. To this aim, large samples of tracers of the large-scale structure of the Universe are needed. Among cosmic tracers, galaxy clusters play a critical role. Reaching masses up to 10^{15} $M_\odot$ and radii up to 5 Mpc, galaxy clusters are the largest gravitationally bound structures existing in the Universe, marking the most pronounced peaks in the matter density field. The steady advancement of cosmological experiments has paved the way for the detection of large galaxy cluster samples across various wavelengths. Specifically, observations in X-rays, optical, near-infrared, and microwave bands, along with gravitational lensing studies, have allowed the investigation of the statistical properties of galaxy clusters.

Since the dark matter component is dominant in galaxy clusters, modelling the cluster formation and evolution through N-body dark-matter-only simulations is accurate enough, given the current observational uncertainties. Such simulations allow us to calibrate the mass function and mass density profiles of galaxy cluster haloes. With a calibrated halo mass function at hand, it is possible to model the observed counts of galaxy clusters to derive constraints on fundamental cosmological parameters. Indeed, cluster counts are among the most powerful cosmological probes, providing excellent constraints on the matter density parameter, Ω_m, on the square root of the mass variance on scales of 8 h^{-1}Mpc at redshift $z = 0$,

σ_8, and on the parameters entering the dark energy equation of state. Robust constraints on such cosmological parameters can be derived also from cluster correlation function and power spectrum measurements, which are commonly referred to as cluster clustering. Furthermore, to constrain the cosmological parameters from observations of galaxy cluster statistics, the theoretical models must be convolved with instrumental effects, such as measurement errors and selection biases. In addition, due to the degeneracy between cluster masses and cosmological parameters, it is necessary to accurately model the cluster matter profiles in order to derive mass estimates. This is possible, for example, through X-ray or weak-lensing observations.

In this Thesis, after introducing the ΛCDM model and the main statistical properties of the Universe, we delve into the methodologies employed for the calibration of galaxy cluster masses and mass function, also comparing the clustering properties of galaxies and galaxy clusters. Then we present a series of works based on real observations of galaxy cluster statistics, developed over the past three years. More specifically, this Thesis is organised as follows:

- In Chapter 1, we provide the foundational knowledge required for the mathematical description of modern cosmological models. After presenting the fundamentals of the General Theory of Relativity and the derivation of the Friedmann Equations, we highlight the key components of the concordance ΛCDM cosmological model.

- In Chapter 2, we delve into the description of the matter power spectrum and correlation functions. We detail how these quantities depend on cosmological parameters, also introducing the basics of weak gravitational lensing.

- In Chapter 3, we provide an overview on the methodologies employed for galaxy cluster detection. We detail the alternative observational techniques used for measuring cluster masses, and how such mass estimates are exploited in cosmological analyses based on cluster counts and clustering.

- In Chapter 4, we detail the analysis by Lesci et al. (2023) on the 3D clustering of the galaxy cluster sample provided by Planck Collaboration XXVII (2016), based on observations of the Sunyaev-Zeldovich (SZ) effect. In this work, we derived constraints on the Planck mass bias, $b_{\rm SZ}$, and on $\Omega_{\rm m}$.

- In Chapter 5, we present the works by Lesci et al. (2022a) and Lesci et al. (2022b) on counts and clustering of the galaxy cluster sample developed by Maturi et al. (2019), based on the third data release of the Kilo Degree Survey (KiDS-DR3, de Jong et al. 2017). Through these cosmological analyses, we derived robust constraints on $\Omega_{\rm m}$, σ_8, and on the scaling relation linking cluster masses and mass proxies.

- In Chapter 6, the analysis carried out by Euclid Collaboration: Lesci et al. (2023) is presented. In this work, we developed a method for calibrating galaxy colour-redshift relations, which are fundamental for cluster weak-lensing analyses.

- In Chapter 7, we present the weak-lensing mass calibration, carried out by Lesci et al. (in prep. b), of the galaxy cluster sample provided by Maturi et al. (in prep.). This sample is based on the fourth data release of KiDS (KiDS-1000, Kuijken et al. 2019).

- Lastly, in Chapter 8 we summarise the results presented in this Thesis, preparing the way for future projects.

Chapter 1

Cosmological framework

In this Chapter, we present the theoretical framework at the foundation of the standard cosmological model. After the description of the spacetime metric, based on the cosmological principle and General Relativity, we introduce the main observational achievements of the last century and their theoretical interpretation. In addition, we detail the evolution of matter perturbations leading to the formation of cosmic structures, such as galaxies and galaxy clusters.

1.1 The cosmological principle

Cosmology investigates the formation and evolution of the Universe, describing its statistical properties through the assumption of the cosmological principle. Such principle states that the Universe is homogeneous and isotropic, and these are the two properties at the foundations of the Friedmann-Lemaître-Robertson-Walker metric, described in Sect. 1.3. The isotropy of the Universe is empirically supported within the observable Universe from the observed isotropy in the cosmic microwave background (CMB) and on scales greater than 100 Mpc, that is on scales larger than those of superclusters and voids. Notably, the cosmological principle entails the fair sample principle, stating that analysing sufficiently large and independent volumes of Universe is equivalent to considering many realisations of Universe. Thus, the fair sample principle tackles the non-reproducibility of the Universe, allowing for the statistical description of the Universe matter-energy content.

1.2 The central role of General Relativity

The theoretical description of the Universe in the standard cosmological model is based on General Relativity (GR; Einstein 1915). In GR, the content of energy and matter acts on the geometry of the spacetime, which is a four-dimensional differentiable manifold characterised by the metric tensor $g_{\alpha\beta}$. Specifically, spacetime points are called *events* and are described by one time-like and three space-like coordinates. The interval between two infinitesimally close events is expressed as

$$\mathrm{d}s^2 = g_{\alpha\beta}\,\mathrm{d}x^\alpha \mathrm{d}x^\beta \qquad (\alpha, \beta = 0,\, 1,\, 2,\, 3), \tag{1.1}$$

where repeated indexes imply summation. In particular, the three space coordinates x^1, x^2, x^3 are generally labeled with the indexes i, j, while the time coordinate is $x^0 = ct$, where c is the speed of light and t is the time. Due to the spacetime curvature, defined by $g_{\alpha\beta}$, free

particles move along non-straight lines, called geodesics, locally defining the shortest path between two events. In particular, the integral along the path of a particle is such that

$$\delta \int_{\text{path}} \mathrm{d}s = 0\,. \tag{1.2}$$

Particle paths can be derived from the geodesic equation, expressed as

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\lambda^2} + \Gamma^\alpha_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = 0\,, \tag{1.3}$$

where Γ are the Christoffel symbols, depending on the metric tensor, and λ is an affine parameter. In addition, the metric tensor is related to the content of energy and matter, defined by the energy-momentum tensor $T_{\alpha\beta}$, through the fundamental Einstein's field equations, expressed as

$$R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = \frac{8\pi G}{c^4} T_{\alpha\beta}\,, \tag{1.4}$$

where $R_{\alpha\beta}$ and $R \equiv g^{\alpha\beta} R_{\alpha\beta}$ are the Ricci tensor the Ricci scalar, respectively, both depending on Γ, while G is Newton's gravitational constant. By describing the content of the Universe as a perfect fluid with pressure p and energy density ρc^2, the energy-momentum tensor takes the following form

$$T_{\alpha\beta} = (p + \rho c^2)\, u_\alpha u_\beta - p g_{\alpha\beta}\,, \tag{1.5}$$

where $u_\alpha = g_{\alpha\beta} u^\beta$ is the four-velocity of the fluid.

1.3 Friedmann-Lemaître-Robertson-Walker metric

As discussed in the previous section, the theoretical description of the Universe in the standard model is based on GR. Thus, a cosmological theory must be rooted in a properly defined spacetime metric. This is the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, based on the cosmological principle. For a homogeneous and isotropic Universe, the mixed components of the metric tensor in comoving coordinates are zero. Thus, from Eq. (1.1), the following expression is obtained,

$$\mathrm{d}s^2 = g_{00} \mathrm{d}t^2 - g_{ij} \mathrm{d}x^i \mathrm{d}x^j = c^2 \mathrm{d}t^2 - \mathrm{d}l^2\,, \tag{1.6}$$

where the first and the second terms are time and space terms, respectively. At any instant defined by t, which in this case coincides with the proper time, or *cosmic time*, the spatial metric $\mathrm{d}l^2$ is identical in all the places and directions. The cosmic time monotonically varies, and it is zero where all the geodesics of the spacetime intersect.

To find a suitable functional form for the three-dimensional spatial metric, we consider the two-dimensional case first. The spaces satisfying the cosmological principle are the flat Euclidean space, the sphere and the hyperboloid. In the case of the flat Euclidean space we have

$$\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2, \tag{1.7}$$

which in polar coordinates $0 \leq \psi < \infty$, $0 \leq \phi < 2\pi$, is expressed as

$$\mathrm{d}l^2 = \mathrm{d}\psi^2 + \psi^2 \mathrm{d}\phi^2. \tag{1.8}$$

We define $\psi \equiv ar$, where a has the dimension of a length while r has no dimension, and $0 \leq r < \infty$. Then the final relation is

$$\mathrm{d}l^2 = a^2(\mathrm{d}r^2 + r^2 \mathrm{d}\phi^2). \tag{1.9}$$

In the case of a spherical surface, we have

$$\mathrm{d}l^2 = a^2(\sin^2\theta\,\mathrm{d}\phi^2 + \mathrm{d}\theta^2), \tag{1.10}$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$. Imposing $\sin\theta = r$, the previous equation has the form

$$\mathrm{d}l^2 = a^2\left[r^2\mathrm{d}\phi^2 + \frac{\mathrm{d}r^2}{1-r^2}\right]. \tag{1.11}$$

Similar to the spherical case, for a hyperbolic surface the following equation

$$\mathrm{d}l^2 = a^2(\sinh^2\theta\,\mathrm{d}\phi^2 + \mathrm{d}\theta^2) \tag{1.12}$$

becomes

$$\mathrm{d}l^2 = a^2\left[r^2\mathrm{d}\phi^2 + \frac{\mathrm{d}r^2}{1+r^2}\right]. \tag{1.13}$$

The general expression for $\mathrm{d}l^2$ in two dimensions is then

$$\mathrm{d}l^2 = a^2\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\phi^2\right], \tag{1.14}$$

where k is the so-called curvature parameter, being $k = 0$ for the flat space, $k = 1$ for the sphere and $k = -1$ for the hyperboloid. The curvature parameter is related to the amount of energy and matter densities and defines the sign of the *Gaussian curvature* C_G:

$$C_G = \frac{k}{a^2}\ . \tag{1.15}$$

In three dimensions, the inclusion of the solid angle, $\mathrm{d}\Omega$, expressed as

$$\mathrm{d}\Omega = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\,, \tag{1.16}$$

leads to the following expressions:

$$\mathrm{d}l^2 = a^2\left[\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\right] \qquad \text{(flat space)}, \tag{1.17}$$

$$\mathrm{d}l^2 = a^2\left[\frac{\mathrm{d}r^2}{1-r^2} + r^2\mathrm{d}\Omega^2\right] \qquad \text{(hypersphere)}, \tag{1.18}$$

$$\mathrm{d}l^2 = a^2\left[\frac{\mathrm{d}r^2}{1+r^2} + r^2\mathrm{d}\Omega^2\right] \qquad \text{(hyperboloid)}. \tag{1.19}$$

By generalizing these equations through the introduction of the k parameter, the FLRW metric can be derived from Eq. (1.6):

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right]. \tag{1.20}$$

In this framework $a(t)$ is called the cosmic scale factor, having the dimension of a length, t is the proper time, and r, θ, ϕ are the comoving coordinates, where r is dimensionless. As mentioned above, k can be either 0, 1 or -1, and in particular the space is flat ($k = 0$), closed ($k = 1$) or open ($k = -1$) if the density parameter $\Omega(t)$, defined as

$$\Omega(t) = \frac{\rho}{\rho_c}, \tag{1.21}$$

is, respectively, equal to, greater or less than unity. In Eq. (1.21), ρ is a density, and ρ_c is the critical density of the Universe, defined as the density for which the Universe is flat:

$$\rho_c = \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2 , \tag{1.22}$$

where $\dot{a}$ is the first time derivative of $a(t)$. At the current cosmic time, t_0, the value of the critical density, $\rho_{0,c}$, is

$$\rho_{0,c} = 1.9 \cdot 10^{-29}\, h^2 \ \ \mathrm{g\ cm^{-3}}, \tag{1.23}$$

where

$$h = \frac{\dot{a}(t_0)}{100\, a(t_0)} \ . \tag{1.24}$$

In addition, based on the FLRW metric, we can derive two fundamental quantities in cosmology, namely the proper distance and the comoving distance. The proper distance, d_p, of a point P from the origin of a set of polar coordinates (r, θ, ϕ), is defined as the geodesic passing through such points, obtained by imposing $\mathrm{d}t = \mathrm{d}\theta = \mathrm{d}\phi = 0$ in Eq. (1.20). Specifically, d_p has the following functional form

$$d_\mathrm{p} = a(t) \int_0^r \frac{\mathrm{d}r'}{(1 - kr'^2)^{1/2}} = a(t) f(r), \tag{1.25}$$

where $f(r)$ has a different form for each value of k, namely

$$f(r) = \sin^{-1} r \qquad (k = 1), \tag{1.26}$$

$$f(r) = r \qquad (k = 0), \tag{1.27}$$

$$f(r) = \sinh^{-1} r \qquad (k = -1). \tag{1.28}$$

Since d_p is defined for $\mathrm{d}t = 0$, it does not account for the limits on the propagation speed of information and then it can not be directly measured. However, d_p has a crucial role in the understanding of the expansion of the Universe, as we will see in the next section. At t_0, the proper distance coincides with the comoving distance, d_c, being defined as follows

$$d_\mathrm{c} = a_0 f(r) = f(r), \tag{1.29}$$

where $a_0 = a(t_0) = 1$. Therefore, the proper and comoving distances are related by the following equation

$$d_\mathrm{c} = \frac{1}{a} d_\mathrm{p}. \tag{1.30}$$

1.4 Hubble-Lemaître law

Based on the cosmological principle and thus on the FLRW metric, the Hubble-Lemaître law (Lemaître 1927; Hubble 1929) describes the expansion of our Universe. In fact, from the definition of the proper distance in Eq. (1.25), the radial velocity of a source at a point P with respect to the origin of a set of polar coordinates (r, θ, ϕ) can be expressed as follows

$$v_r = \dot{a} f(r) = \frac{\dot{a}}{a} d_\mathrm{p} \, . \tag{1.31}$$

This is the Hubble-Lemaître law, and $H(t) \equiv \dot{a}/a$ is often referred to as the Hubble parameter, which is constant at a given proper time. For the present cosmic time, we refer to $H(t_0) = H_0$ as the Hubble constant. The ESA Planck mission measured $H_0 = 67.4 \pm 0.5$ km s^{-1} Mpc^{-1}

(Planck Collaboration VI 2020) from the angular power spectrum of the cosmic microwave radiation (CMB), assuming a standard Λ-cold dark matter (ΛCDM) cosmological model (described in Sect. 1.7). However, this value of the Hubble constant is not in agreement with that obtained through other independent late-time cosmological probes, such as observations of type Ia supernovae (Dhawan et al. 2018) or Cepheids in nearby galaxies (Riess et al. 2019). These tensions in H_0 are subject of current debates and might imply the development of alternative cosmological models (for an extensive review, see Di Valentino et al. 2021), which will be tested through upcoming experiments such as *Euclid* (Laureijs et al. 2011) and the *Vera C. Rubin Observatory* (*Rubin*/LSST; Ivezic et al. 2008).

1.5 Cosmological redshift and distances

As the Universe expands following the Hubble-Lemaître Law, the light from extragalactic sources shifts towards longer wavelengths, namely redshifts. This cosmological redshift, z, is observable and it is defined as

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e} \ . \tag{1.32}$$

The quantity λ_e is the source rest-frame wavelength, emitted at time t_e and at a comoving coordinate r. On the other hand, λ_0 is the wavelength observed at time t_0 at the origin of the coordinate system. As light travels along null geodesics, we have $\mathrm{d}s^2 = 0$, and for $\mathrm{d}\theta = \mathrm{d}\phi = 0$ the Eq. (1.20) becomes

$$\int_{t_e}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = \int_0^r \frac{\mathrm{d}r}{\sqrt{1-kr^2}} = f(r). \tag{1.33}$$

We remark that $f(r)$, expressed in Eqs. (1.26) – (1.28), does not vary with time, because both the source and the observer follow the cosmological expansion and therefore r is constant. Consequently, by considering the light emitted from the source at $t' = t_e + \delta t_e$ and received by the observer at $t_0' = t_0 + \delta t_0$, we can write

$$\int_{t'}^{t_0'} \frac{c\,\mathrm{d}t}{a(t)} = f(r). \tag{1.34}$$

If the time intervals δt_e and δt_0 are small, from Eqs. (1.33) and (1.34) we derive that

$$\frac{\delta t_0}{a_0} = \frac{\delta t_e}{a(t_e)} \ . \tag{1.35}$$

Since the frequencies of emitted and observed light can be expressed as $\nu_e = 1/\delta t_e$ and $\nu_0 = 1/\delta t_0$, respectively, from Eq. (1.32) we obtain

$$1 + z = \frac{\lambda_0}{\lambda_e} = \frac{a_0}{a(t_e)} = \frac{1}{a(t_e)} \ . \tag{1.36}$$

This equation summarises the fact that in an expanding Universe, where $a(t)$ increases with time, the observed light of galaxies is redshifted. In addition, Eq. (1.36) allows to infer the distance of extragalactic sources from redshift measurements, achieved through photometry and spectroscopy. To this aim, proper motions due to gravitational effects must be taken into account, as they may significantly affect redshift estimates.

Specifically, Eq. (1.36) allows for *luminosity distance*, d_L, and *angular diameter distance*, d_A, measurements. These distance definitions are alternative to the proper distance in Eq.

(1.25), which is not directly related to the redshift. The luminosity distance is defined in order to preserve the the inverse-square law, and it is expressed as

$$d_{\rm L} = \left(\frac{L}{4\pi l}\right)^{1/2}, \tag{1.37}$$

where L is the luminosity of a source at comoving distance r, emitted at time t, while l is the flux measured at time t_0 by the observer, which is expressed as

$$l = \frac{L}{4\pi a_0^2 r^2}\left(\frac{a}{a_0}\right)^2, \tag{1.38}$$

where $4\pi a_0^2 r^2$ is the surface area of a sphere having its centre on the source and passing through the position of the observer. The factor $(a/a_0)^2$ accounts for cosmological redshift and time interval dilation. In fact, Eq. (1.36) implies that photons are redshifted by a factor a/a_0, and from Eq. (1.35) we know that the photons emitted in a time interval δt_e arrive to the observer in an interval $\delta t_0 = (a_0/a)\delta t_e$. Thus, by combining Eqs. (1.37) and (1.38), we obtain

$$d_{\rm L} = a_0^2 \frac{r}{a} = a_0 r(1+z). \tag{1.39}$$

In addition, the angular diameter distance is defined in order to preserve the variation of angular size with the distance, thus it is expressed as

$$d_{\rm A} = \frac{l}{\Delta\theta}, \tag{1.40}$$

where l is the physical size of a source with comoving coordinate r and emitting light at time t, while $\Delta\theta$ is the angle subtended by l. As $l = ar\Delta\theta$, we obtain

$$d_{\rm A} = ar = \frac{a_0 r}{1+z}. \tag{1.41}$$

By assuming a spacetime metric, luminosity and angular diameter distances can be derived for any extragalactic source having an associated redshift measure. Standard candles, that is objects with known luminosity, can be exploited to test the FLRW metric through Eq. (1.39), if the source redshift measurement is available. Similarly, objects with known physical size, namely standard rulers, provide tests for the assumed spacetime metric through Eq. (1.41). An additional test is provided by the following relation,

$$\frac{d_{\rm L}}{d_{\rm A}} = (1+z)^2, \tag{1.42}$$

derived from Eqs. (1.39) and (1.41). Eq. (1.42) is referred to as the *duality relation* and, by observing a source which is both a standard candle and a standard ruler, it can be exploited to quantify the deviations from the FLRW metric and in turn from homogeneity and isotropy. The duality relation provides a more robust test of the metric, compared to those relying either on luminosity or angular diameter distances. In fact, $d_{\rm L}$ and $d_{\rm A}$ depend on cosmological parameters (see Sect. 1.7), while Eq. (1.42) does not. However it is difficult to find objects that are both standard candles and standard rulers. To date, the most robust duality relation probes are strong gravitational lensing and compact radio sources (see, e.g., Qin et al. 2021; Tang et al. 2023; Li 2023).

1.6 Friedmann models

In this section we introduce the Friedmann models, which play a key role in cosmology. Indeed, based on the FLRW metric, the Friedmann models describe the matter-density content of the Universe in terms of the density, ρ, and the pressure, p, of a perfect fluid.

1.6.1 Friedmann equations

The relationship between the geometry of spacetime and the matter content in a Universe governed by General Relativity is established through the Einstein equations, Eq. (1.4). By employing these equations, we can derive the Friedmann equations, which rely on two fundamental assumptions: the FLRW metric and the representation of the Universe content as a perfect fluid. This implies that the energy-momentum tensor takes the form outlined in Eq. (1.5). Through these assumptions, the Einstein equations can be reduced to the following equation,

$$\ddot{a} = -\frac{4\pi}{3} G \left(\rho + 3\frac{p}{c^2}\right) a \,, \tag{1.43}$$

for the time-time component, and

$$a\ddot{a} + 2\dot{a}^2 + 2kc^2 = 4\pi G \left(\rho - \frac{p}{c^2}\right) a^2 \,, \tag{1.44}$$

for the space-space components, namely with $\alpha,\, \beta = 1$, $\alpha,\, \beta = 2$, $\alpha,\, \beta = 3$ in Eq. (1.4). Under the assumption of the cosmological principle, the spacetime components yield the identity $0 = 0$. From Eq. (1.43), the Eq. (1.44) takes the following form

$$\dot{a}^2 + kc^2 = \frac{8}{3}\pi G \rho a^2. \tag{1.45}$$

Equations (1.43) and (1.45) are the Friedmann equations (Friedmann 1922). Under the assumption of an adiabatic expansion of the Universe, such equations are related by the adiabaticity relation, expressed as follows:

$$\mathrm{d}(\rho\, c^2 a^3) = -p\, \mathrm{d}a^3. \tag{1.46}$$

1.6.2 The cosmological constant

Before the discovery of its expansion, the Universe was assumed to be a static and not evolving object. However, according to Eq. (1.43) the Universe can not be static, unless either the pressure, p, or the energy density, ρc^2, are negative. This contradicts the physical definitions of p and ρc^2, rendering such a scenario impossible. To try to address this issue, in 1917 Einstein introduced a cosmological constant, Λ, in his equations, thereby modifying the nature of gravity itself:

$$R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R - \Lambda g_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta}. \tag{1.47}$$

By selecting a sufficiently small value for Λ, one can achieve a static, though unstable, cosmological model without invalidating the description of planetary motions. After the discovery of the Hubble-Lemaître law in the 1920s, described in Sect. 1.4, the cosmological constant was not necessary anymore. Nevertheless, from the observed flux of distant Ia supernovae, Riess et al. (1998) showed that the expansion of the Universe is accelerated, contradicting the prediction of decelerated expansion derived from Eq. (1.43). As a result, the cosmological constant Λ has been reintroduced and it plays a role in the current cosmological model. An alternative, though fully equivalent approach, is to include Λ into the energy-momentum tensor, thus modifying the energy and matter content of the universe. Indeed, this can be interpreted as introducing a new energy density of an unknown *dark energy* component. Specifically, the modified energy-momentum tensor, $\tilde{T}_{\alpha\beta}$, takes the following form:

$$\tilde{T}_{\alpha\beta} = T_{\alpha\beta} + \frac{\Lambda c^4}{8\pi G} g_{\alpha\beta} = -\tilde{p} g_{\alpha\beta} + (\tilde{p} + \tilde{\rho} c^2) u_\alpha u_\beta, \tag{1.48}$$

where $\tilde{p}$ and $\tilde{\rho}$ are the effective pressure and the effective density, respectively. In the case of a perfect fluid, such quantities are expressed as

$$\tilde{p} = p - \frac{\Lambda c^4}{8\pi G}, \quad \tilde{\rho} = \rho + \frac{\Lambda c^2}{8\pi G} \, . \tag{1.49}$$

Therefore, the Einstein field equations can be written as

$$R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = \frac{8\pi G}{c^4} \tilde{T}_{\alpha\beta}, \tag{1.50}$$

and the Friedmann equations become

$$\ddot{a} = -\frac{4\pi}{3} G \left(\tilde{\rho} + 3\frac{\tilde{p}}{c^2} \right) a, \tag{1.51}$$

$$\dot{a}^2 + kc^2 = \frac{8}{3}\pi G \tilde{\rho} a^2. \tag{1.52}$$

We remark that the introduction of the cosmological constant is physically well motivated in GR, as its presence naturally descends from the general expression of the gravitational field second order action (Bianchi & Rovelli 2010).

1.6.3 Perfect fluid models

Given the equation of state of the perfect fluid describing the matter-energy content in the Universe, the Friedmann equations and the adiabaticity relation in Eq. (1.46) allow for the derivation of $a(t)$, $\rho(t)$, and $p(t)$, as a function of the cosmic time, t. As the Friedmann models are based on the cosmological principle, the pressure is isotropic. Furthermore, isotropic pressure is a necessary requirement for a perfect fluid. At rest, perfect fluids are fully characterised by their energy density ρc^2 and pressure p, allowing for the following definition of the fluid equation of state:

$$p = w \rho c^2, \tag{1.53}$$

where w is assumed to be constant with time, and it is defined in the range

$$0 \leq w \leq 1, \tag{1.54}$$

called the *Zel'dovich interval*. This range is derived from the definition of the adiabatic sound speed of the fluid, v_s, defined as

$$v_\mathrm{s}^2 = \frac{\partial p}{\partial \rho}\Big|_{S=\mathrm{const}} , \tag{1.55}$$

where S denotes the entropy. From Eq. (1.53), the sound speed would exceed the speed of light for $w > 1$, while for $w < 0$ the sound speed would be an imaginary variable. It is possible to show that in case of non-relativistic matter, also referred to as *dust*, $w = 0$ is a good approximation. This implies $p = 0$, which is expected for non-relativistic fluids. Indeed, the typical particle thermal energy, namely $k_\mathrm{B} T$, where k_B is the Boltzmann constant and T is the fluid temperature, is much lower than the particle rest mass, $m_\mathrm{p} c^2$, where m_p is the particle mass. Furthermore, for a non-degenerate radiative fluid composed of photons or ultrarelativistic particles in thermal equilibrium, $w = 1/3$ is obtained.

By combining the adiabaticity relation, Eq. (1.46), and the general equation of state, Eq. (1.53), it is possible to derive the following relation:

$$\rho_w = \rho_{0,w} \left(\frac{a}{a_0} \right)^{-3(1+w)} , \tag{1.56}$$

where $\rho_{0,w}$ is the density at the present time of a fluid with parameter of the equation of state w. In case of a dust-only (i.e. matter-only) universe, Eq. (1.56) becomes

$$\rho_{\rm m} = \rho_{0,\rm m} \left(\frac{a}{a_0}\right)^{-3} = \rho_{0,\rm m} \left(1+z\right)^3 , \tag{1.57}$$

while for a radiative universe we obtain

$$\rho_{\rm r} = \rho_{0,\rm r} \left(\frac{a}{a_0}\right)^{-4} = \rho_{0,\rm r} \left(1+z\right)^4 , \tag{1.58}$$

where $\rho_{\rm m}$ and $\rho_{\rm r}$ are the matter and radiation densities, respectively. Eq. (1.57) is expected from the expansion of the Universe, for which the dust particle density decreases as the cube of the scale factor a. The same is true for radiative fluids, which also undergo an increase in wavelength by a factor a. This explains the differences in the time evolution of $\rho_{\rm m}$ and $\rho_{\rm r}$.

As discussed in Sect. 1.6.2, the fluid energy-momentum tensor includes the contribution of the cosmological constant, Λ, implying the accelerated expansion of the Universe. In an empty universe, from Eq. (1.49) we can derive the pressure and density associated to the cosmological constant as follows:

$$p_\Lambda = -\frac{\Lambda c^4}{8\pi G}, \qquad \rho_\Lambda = \frac{\Lambda c^2}{8\pi G} . \tag{1.59}$$

Consequently, from Eq. (1.53), $w = -1$ holds for the perfect fluid equivalent of a cosmological constant. Consistent with this theoretical approximation, Planck Collaboration VI (2020) derived $w = -1.028 \pm 0.032$ from the combination of CMB, supernovae, and baryon acoustic oscillations observations, assuming that w does not change over cosmic time. In addition, by assuming $w = -1$, from Eq. (1.56) we can see that the energy density associated to the cosmological constant does not vary as the Universe expands.

1.6.4 Evolution of Hubble and density parameters

Based on the perfect fluid description of the matter-energy content of the Universe, we can theoretically derive the evolution of the scale parameter, $a(t)$, and, in turn, of the Hubble parameter, $H(t)$. Considering Eq. (1.45) at the present time, t_0, and dividing by a_0^2, we obtain

$$H_0(1-\Omega_{0,w}) = -\frac{kc^2}{a_0^2} , \tag{1.60}$$

where $\Omega_{0,w}$ is the density parameter at t_0 for a fluid component with equation of state parameter w, expressed as

$$\Omega_{0,w} = \frac{\rho_{0,w}}{\rho_{0,cr}} . \tag{1.61}$$

Thus, by considering Eq. (1.45) at a generic time t and dividing by a_0^2, we obtain

$$\frac{\dot{a}^2}{a_0^2} = H_0^2 \left[1 - \Omega_{0,w} + \Omega_{0,w} \left(\frac{a_0}{a}\right)^{1+3w}\right] . \tag{1.62}$$

We can express this relation in terms of the Hubble parameter as

$$H^2(t) = H_0^2 \left(\frac{a_0}{a}\right)^2 \left[1 - \Omega_{0,w} + \Omega_{0,w} \left(\frac{a_0}{a}\right)^{1+3w}\right] , \tag{1.63}$$

which can be generalised for universes with more than one fluid component as follows:

$$H^2(t) = H_0^2 \left(\frac{a_0}{a}\right)^2 \left[1 - \sum_i \Omega_{0,w_i} + \sum_i \Omega_{0,w_i} \left(\frac{a_0}{a}\right)^{1+3w_i}\right]. \tag{1.64}$$

Since observations show that the Universe expands, we have $\dot{a}(t) > 0$, and Eq. (1.43) implies that $\ddot{a}(t) < 0$ for any t if $(\rho + 3p/c^2) > 0$ or, equivalently, if $(1 + 3w) > 0$. Consequently, $a(t)$ is a concave function and it must be zero at a finite time in the past. Thus, at such time, the density and the Hubble parameter diverge. We conclude that models based on the cosmological principle and described by perfect fluids with $-1/3 < w < 1$ show a singularity at the time $t = 0$, which is commonly referred to as the *Big Bang singularity*. Such singularity can be avoided, for example, in universes with a non-zero cosmological constant or in models that do not follow the cosmological principle.

In general, the density parameter of a fluid component depends on time, or redshift, and it can be expressed as follows:

$$\Omega_w(z) = \frac{\rho_w(z)}{3H^2(z)/(8\pi G)} \,. \tag{1.65}$$

Considering Eq. (1.56) for $\rho_w(z)$ and Eq. (1.63) for $H(z)$, we obtain

$$\Omega_w(z) = \frac{\Omega_{0,w}(1+z)^{1+3w}}{(1-\Omega_{0,w}) + \Omega_{0,w}(1+z)^{1+3w}}, \tag{1.66}$$

so that:

$$\Omega_w^{-1}(z) - 1 = \frac{\Omega_{0,w}^{-1} - 1}{(1+z)^{1+3w}} \,. \tag{1.67}$$

This relation implies that $\Omega_w^{-1}(z) - 1$ can not change its sign over time, and that flat universes, for which the total density parameter is equal to 1, preserve their flatness. In other words, the expansion can not change the value of the curvature parameter.

1.6.5 Cosmological horizon

By knowing the evolution of the scale factor, $a(t)$, it is possible to derive the cosmological horizon, $R_{\rm H}(t)$. This quantity defines a sphere centred on a particle, containing the volume of the Universe which is in causal connection with that particle. In particular, $R_{\rm H}(t)$ is defined as follows:

$$R_{\rm H}(t) = a(t) \int_0^t \frac{c\,{\rm d}t'}{a(t')} \,. \tag{1.68}$$

We note that $R_{\rm H}(t)$ is a proper distance and in turn it accounts for the expansion of the Universe. In fact, Eq. (1.68) is derived from the definition of proper distance, Eq. (1.25), given the $f(r)$ expression for a photon in Eq. (1.34). If the integral does not diverge, $R_{\rm H}(t)$ defines the so-called *particle horizon*. From Eq. (1.62), it is possible to show that $a(t) \propto t^\beta$, with $\beta > 0$, for a generic Friedmann model. Therefore, from Eq. (1.68), we obtain

$$R_{\rm H}(t) \propto t^\beta \int_0^t t'^{-\beta} {\rm d}t' \,. \tag{1.69}$$

Consequently, $R_{\rm H}$ exists only if $\beta < 1$. Additionally, from Eq. (1.43) we have

$$\ddot{a} = -\frac{4\pi}{3} G a \left(\rho + \frac{3p}{c^2}\right) \propto \beta(\beta - 1) t^{\beta - 2}, \tag{1.70}$$

from which

$$\beta(\beta-1) \propto -\frac{4\pi}{3} G t^2 \left(\rho + \frac{3p}{c^2}\right). \tag{1.71}$$

From this equation, we note that the condition for a Big Bang singularity, namely $\ddot{a} < 0$, implies $\beta(\beta-1) < 0$, that is $0 < \beta < 1$. In turn, the presence of a Big Bang guarantees the finiteness of $R_{\rm H}(t)$. Lastly, for a flat Friedmann model it is possible to show that

$$R_{\rm H}(t) \simeq 3\frac{1+w}{1+3w} ct\,, \tag{1.72}$$

which, in the case of dust and radiation flat universes, becomes $R_{\rm H} \simeq 3ct$ and $R_{\rm H} \simeq 2ct$, respectively.

1.7 The standard cosmological model

The fundamental concept underlying the description of the observed Universe is the *hot Big Bang model*, which relies on the cosmological principle and the Friedmann equations. Within this framework, the evolution of the Universe is depicted as a thermal history, hence the term "hot". Indeed, as we trace back closer to the moment of the Big Bang, the Universe progressively becomes hotter. Notably, the present-day temperature of the Universe is measured to be $T = 2.7255 \pm 0.0009$ K (Fixsen et al. 1996). This is the temperature of the CMB, corresponding to the black-body radiation that began to freely propagate following the first recombination of free electrons, happened at redshift $z \simeq 1100$. As such, the CMB serves as one of the most compelling pieces of evidence supporting the Big Bang theory. In addition, CMB observations provide empirical support to the isotropy assumption enclosed in the cosmological principle. In fact, the CMB temperature shows very small fluctuations, of around 10^{-5} K, in any direction on the sky. On the other hand, we remark that these measurements do not prove the homogeneity of the spacetime, as this would require observations from different locations in the Universe.

As the same CMB temperature is measured in all directions, all regions of the Universe seem to be in causal connection at $z \simeq 1100$. However, given the expected cosmological horizon size at that time, causal connection should not hold at scales larger than about 2° on sky. This issue could be solved through cosmic inflation, first proposed by Guth (1981), implying an accelerated expansion with an exponential growth of the cosmic scale factor, $a(t)$, at early time. Most of the modern inflationary models, based on the work by Linde (1983), consider a scalar field with large initial potential energy, called inflaton, as the responsible of the inflationary era. In these models, the inflaton finally decays into particles through quantum fluctuations, at the end of the exponential expansion of the Universe. Consequently, inflation creates the initial conditions for the growth of the perturbations in the Universe, leading to the formation of the observed cosmic structures. In addition, cosmic inflation explains the flatness of the Universe spacetime geometry, observed via CMB experiments (de Bernardis et al. 2000).

The standard cosmological model, known as ΛCDM, is based on hot Big Bang and cosmic inflation models. Here, Λ represents the contribution of dark energy in the form of the cosmological constant, while CDM refers to cold dark matter, which is a dust-like typology of dark matter. In addition to CDM, a smaller contribution to the total matter content is given by baryonic matter. The estimation of this baryonic matter component can be obtained through various means, such as from the CMB, the mass-luminosity ratio of observable structures in the Universe, or studying primordial nucleosynthesis (Thuan & Izotov 2000). The mass-luminosity ratio can be determined through dynamical investigations of galaxies, galaxy

groups, and clusters, serving as a fundamental piece of evidence supporting the existence of dark matter.

Cold dark matter is included in the standard cosmological model as it attains a non-relativistic state earlier than hot dark matter. In fact, as the Universe is described in terms of its thermal history, cosmic time is intrinsically linked to temperature. As a result, we can estimate the time at which a fluid ceases to be relativistic using the relationship $k_{\rm B}T \simeq m_x c^2$, where $k_{\rm B}$ denotes the Boltzmann constant and m_x represents the characteristic mass of the fluid particles. Given that cold dark matter particles possess greater mass compared to hot dark matter ones, they undergo the transition to a non-relativistic state prior to their hot counterparts. Additionally, cold dark matter holds another notable advantage over hot dark matter, as it is already non-relativistic at the moment of decoupling from the radiation component (see Sect. 1.8). This facilitates the prompt emergence of CDM gravitational perturbations, which serves as a fundamental aspect of the widely accepted *hierarchical model* for the development of cosmic structure. This model entails the initial formation of small structures within the Universe, which subsequently experience mergers, resulting in the formation of larger gravitationally bound objects.

As we mentioned, another crucial element of the standard cosmological model is the inclusion of dark energy, represented by Λ. This energy component is essential to explain the accelerated expansion of the Universe, observed for the first time by Riess et al. (1998) and Perlmutter et al. (1999). In fact, by combining the first Friedmann equation, Eq. (1.43), and the general perfect fluid equation of state, Eq. (1.53), we obtain the following expression:

$$\ddot{a} = -\frac{4\pi}{3} G a \left(\rho + \frac{3p}{c^2}\right) = -\frac{4\pi}{3} G a \rho (1 + 3w)\,, \tag{1.73}$$

from which a decelerated expansion is expected, if the density ρ is positive and w ranges within the Zel'dovich interval, namely Eq. (1.54). As discussed in Sect. 1.6.3, a positive Λ is required to explain the accelerated expansion of the Universe, which implies a positive density and a negative pressure (see Eq. 1.59). Thus, $w_\Lambda = -1$ holds for the perfect fluid associated to Λ. As detailed in Sect. 1.6.4, models based on the cosmological principle and described by perfect fluids with $-1/3 < w < 1$ show a Big Bang singularity. As $w_\Lambda = -1$ for dark energy, the Big Bang may be eluded, in contrast with CMB observations. Nevertheless, the measured energy density associated to Λ is too low, as it corresponds to

$$\rho_{0,\Lambda} = \rho_\Lambda = \text{const} \simeq 10^{-29}\ \text{g cm}^{-3}, \tag{1.74}$$

and it is the dominant energy contribution in the Universe only since $z \simeq 0.67$. Therefore, the dark energy density is not sufficiently high to avoid the Big Bang.

Alternatives to the cosmological constant, such as quintessence models (Tsujikawa 2013), predict a variation of w with time. Stage IV cosmological surveys, like the *Euclid* mission (Laureijs et al. 2011), will allow investigating the evolution of w_Λ. The most widely used w_Λ parameterisation is that proposed by Chevallier & Polarski (2001) and Linder (2003) (CPL), for which w_Λ depends on redshift through the following expression:

$$w_\Lambda(z) = w_0 + w_{\rm a}\,\frac{z}{1+z}\,. \tag{1.75}$$

Models with varying w_Λ are commonly referred to as wCDM. On the other hand, the ΛCDM cosmological model is based on the assumption of $w_\Lambda = -1$. In addition, considering a negligible radiation density and a flat geometry, the ΛCDM model is fully characterised by six parameters, namely:

- $\Omega_{\rm m}$: total matter density parameter;
- $\Omega_{\rm b}$: baryonic matter density parameter;
- H_0: Hubble constant;
- $A_{\rm s}$: normalisation of the primordial power spectrum;
- $n_{\rm s}$: spectral index of the primordial power spectrum;
- τ: reionization optical depth.

The matter power spectrum, depending on $A_{\rm s}$ and $n_{\rm s}$, will be detailed in Chapt. 2. The parameters $\Omega_{\rm m}$ and $\Omega_{\rm b}$ refer to their present-day values. Indeed, in this thesis we will drop the subscript "0" for simplicity. In addition, in flat universes, the dark energy density parameter is expressed as $\Omega_\Lambda = 1 - \Omega_{\rm m}$.

The most robust constraints on all the ΛCDM parameters can be obtained through CMB observations. From the combination of CMB temperature fluctuations and lensing, Planck Collaboration VI (2020) derived the following constraints: $\Omega_{\rm m}h^2 = 0.1430 \pm 0.0011$, $\Omega_{\rm b}h^2 = 0.02237 \pm 0.00015$, $H_0 = 67.36 \pm 0.54$ km s^{-1} Mpc^{-1}, $10^9 A_{\rm s} = 2.100 \pm 0.030$, $n_{\rm s} = 0.9649 \pm 0.0042$, $\tau = 0.0544 \pm 0.0073$. Here, h is the reduced Hubble constant, defined as $h = H_0/100$ km s^{-1} Mpc^{-1}. From these results, we can see that $\Omega_{\rm m}$ is dominated by the contribution of the cold dark matter density parameter, $\Omega_{\rm c}$, since $\Omega_{\rm c} = \Omega_{\rm m} - \Omega_{\rm b}$.

1.8 Structure formation

In the standard model of large-scale structure formation, the presence of observed gravitationally bound systems is attributed to the growth of small initial perturbations. As discussed in Sect. 1.7, these perturbations originate from quantum fluctuations created at the end of the inflationary era. The effects of these resulting perturbations can be observed in the temperature fluctuations of the CMB. Through the investigation of the growth rate of perturbations, referred to as the *growth factor*, it is possible to estimate the expansion rate of the Universe. The more the intensity of expansion increases, the more gravitational collapse is hampered. Hence, by inference, from the evolution of the perturbations it is possible to deduce the cosmic expansion rate. Consequently, observations at various redshifts allow estimating the Hubble parameter, $H(z)$, which serves as a parameterization of the Universe expansion rate.

In this section, we will focus on the growth of the density contrast of the fluctuations, δ, expressed as follows:

$$\delta = \frac{\rho - \overline{\rho}}{\overline{\rho}} = \frac{\delta\rho}{\rho}, \tag{1.76}$$

where ρ is the fluctuation density, and $\overline{\rho}$ is the mean density of the Universe. We will provide an overview on the analytical description of the growth of perturbations in linear regime, which is a reliable approximation for describing a remarkable part of the evolution of cosmic structures. Subsequently, we will explore the nonlinear regime and the methodologies employed to describe it.

1.8.1 Jeans linear theory

The Jeans theory describes the gravitational collapse of material within a homogeneous and isotropic fluid. Essentially, it postulates that density perturbations surpassing a critical scale within this fluid could initiate a gravitational collapse, leading to the growth of these perturbations. As these fluctuations become denser, they will accrete more material, ultimately

culminating in an instability that can lead to the formation of a self-gravitating structure.

Nevertheless, the formation of gravitationally bound structures is a complex process involving hydrodynamical and dynamical effects, such as gas heating and angular momentum increment, which hinder the collapse. These effects cannot be adequately described in a linear mathematical regime and severely impact the formation of galaxies. Nevertheless, interactions on large scales are primarily driven by gravity alone. In addition, as we discussed in Sect. 1.7, the matter content in the Universe is dominated by dark matter, which is weakly interactive and, in turn, does not undergo the same physical processes as baryonic matter. Hence, large scales and dark matter constitute two crucial components for a linear regime, which is the main framework that we will primarily focus on in the subsequent discussion. In addition, we will adopt the Newtonian formalism to describe the effects of gravity, as this is a good approximation on large scales.

To elucidate the basic concept of Jeans theory, let us consider a spherical perturbation with radius R, mass M, and density ρ, in a background fluid having mean density $\overline{\rho}$. Specifically, we assume $\rho > \overline{\rho}$, namely a positive density fluctuation $\delta\rho$:

$$\delta\rho = \rho - \overline{\rho} > 0. \tag{1.77}$$

To achieve a collapse, the absolute value of the gravitational potential energy, $E_{\rm p}$, expressed as

$$E_{\rm p} = -\frac{GM^2}{R} = -\frac{GM}{R}\rho R^3\,, \tag{1.78}$$

must exceed, the thermal kinetic energy, $E_{\rm k}$, expressed as

$$E_{\rm k} = \frac{1}{2}Mv^2\,, \tag{1.79}$$

where v is the mean fluid velocity, derived from the Maxwell-Boltzmann distribution. Thus, the condition for the collapse, $|E_{\rm p}| > E_{\rm k}$, translates into the following expression:

$$G\rho R^2 > \frac{v^2}{2}. \tag{1.80}$$

By defining the Jeans scale, $R_{\rm J}$, as follows,

$$R_{\rm J} = v\frac{1}{\sqrt{2G\rho}}\,, \tag{1.81}$$

from Eq. (1.80) we have $R > R_{\rm J}$. Now that we have familiarized ourselves with the principles of gravitational collapse in a fluid, we can delve into a more in-depth discussion of the evolution of perturbations.

1.8.2 Jeans instability

Jeans (1902) originally studied the collapse of a cloud of interstellar molecular gas and dust, wherein the local self-gravitation makes the impact of the Universe expansion negligible. Nevertheless, to describe the growth of density perturbations on large scales, we must account for the expansion of the Universe and, in turn, for the evolution of the cosmological horizon, $R_{\rm H}(t)$. As discussed in Sect. 1.6.5, $R_{\rm H}(t)$ denotes the extension of the regions in the Universe which are in causal connection at the time t. Consequently, on scales smaller than $R_{\rm H}(t)$, hydrodynamical and dynamical processes become dominant. In this case, Jeans theory provides a reliable description of the collapse in linear regime. On the other hand, on scales larger than $R_{\rm H}(t)$ the dominant interaction is gravity, and the collapse is not hindered

by microphysics.

Before delving into the formalism of Jeans instability, let us focus on the evolution of perturbations on scales larger than $R_{\rm H}(t)$. Let us consider a density fluctuation, treated as a small closed universe, evolving in a background flat universe. We can consider the background flat universe as composed by only one component. In fact, gravity couples the evolution of all fluid components, with the minor ones following the time evolution of the density contrast of the dominant component. From Friedmann equations, we obtain

$$H_{\rm b}^2 = \frac{8\pi}{3} G\rho_{\rm b}\,, \tag{1.82}$$

$$H_{\rm p}^2 = \frac{8\pi}{3} G\rho_{\rm p} - \frac{c^2}{a^2}\,, \tag{1.83}$$

where the subscripts "b" and "p" refer to background and perturbed universes, respectively, and H is the Hubble parameter. Initially, we impose that the Hubble parameters of the two universes are the same, thus the perturbation density contrast is expressed as follows:

$$\delta = \frac{\rho_{\rm p} - \rho_{\rm b}}{\rho_{\rm b}} = \frac{3c^2}{8\pi G}\frac{1}{\rho_{\rm b} a^2} \propto \rho_{\rm b}^{-1} a^{-2}\,. \tag{1.84}$$

At the beginning of the thermal history of the Universe, a radiative gas composed of photons and relativistic particles dominated the matter-energy content. We refer to $z_{\rm eq}$ as the equivalence redshift, at which the density of matter and radiation fluids became equal. Thus, for $z > z_{\rm eq}$ radiation dominates, and $\rho_{\rm b} \propto a^{-4}$, as discussed in Sect. 1.6.3. Conversely, matter dominates for $z < z_{\rm eq}$, implying $\rho_{\rm b} \propto a^{-3}$. As the background universe is flat and contains only one dominant component, from Eq. (1.62) we obtain

$$\frac{\dot{a}^2}{a_0^2} = H_0^2 \left(\frac{a_0}{a}\right)^{1+3w}\,, \tag{1.85}$$

which can be integrated, yielding

$$a(t) = a_0 \left(\frac{t}{t_0}\right)^{\frac{2}{3\,(1+w)}}\,. \tag{1.86}$$

Recalling that $w = 1/3$ for radiation and $w = 0$ for dust-like matter, from Eq. (1.86) we obtain the following expression for the perturbation density contrast:

$$\delta \propto \begin{cases} a^2 \propto t & \text{if } z > z_{\rm eq} \\ a \propto t^{2/3} & \text{if } z < z_{\rm eq} \end{cases}\,. \tag{1.87}$$

Consequently, perturbations on scales larger than $R_{\rm H}$ always grow with time.

Now we focus on perturbations on scales smaller than $R_{\rm H}$, following the Jeans gravitational instability formalism. Let us consider a universe composed of a homogeneous and isotropic perfect fluid with density $\rho_{\rm b}(\boldsymbol{x}, t)$, where $\boldsymbol{x}$ is the fluid position vector. In Newtonian approximation, the generic fluid equations of motion are expressed as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \boldsymbol{v} = 0\,, \tag{1.88}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\frac{1}{\rho}\nabla p - \nabla \phi\,, \tag{1.89}$$

$$\nabla^2 \phi = 4\pi G \rho\,, \tag{1.90}$$

$$p = p(S, \rho) = p(\rho)\,, \tag{1.91}$$

$$\frac{{\rm d}S}{{\rm d}t} = 0\,, \tag{1.92}$$

where Eq. (1.88) is the continuity equation, Eq. (1.89) is the Euler equation, Eq. (1.90) is the Poisson equation, Eq. (1.91) is the fluid equation of state, and Eq. (1.92) is the adiabatic condition. The latter is observationally motivated, as CMB experiments showed that primordial fluctuations are consistent with being adiabatic. Specifically, adiabatic perturbations affect all the particle species such that the relative ratios in the number densities remain unperturbed. In addition, in Eqs. (1.88) – (1.92), $\boldsymbol{v}$ is the flow velocity, ϕ is the gravitational potential, p is the pressure, and S is the entropy. Consequently, p can be expressed in terms of ρ only, thus we can consider only three variables, namely ρ, $\boldsymbol{v}$, and ϕ. In turn, only Eqs. (1.88) – (1.90) are necessary to study the fluid motion.

Now we introduce small perturbations, namely $\delta\rho \ll 1$, $\delta\boldsymbol{v} \ll 1$, and $\delta\phi \ll 1$, so that the linear regime is valid. Consequently, we have the following perturbed quantities:

$$\rho = \rho_{\rm b} + \delta\rho\,, \tag{1.93}$$

$$\phi = \phi_{\rm b} + \delta\phi\,, \tag{1.94}$$

$$\boldsymbol{v} = \boldsymbol{v}_{\rm b} + \delta\boldsymbol{v} = H\boldsymbol{x} + \boldsymbol{v}_{\rm pec}\,, \tag{1.95}$$

where "b" refers to background quantities, and H is the Hubble parameter. Notably, the velocity perturbation in Eq. (1.95), $\delta\boldsymbol{v}$, corresponds to the peculiar velocity, $\boldsymbol{v}_{\rm pec}$. We remark that Eq. (1.95) reduces to the Hubble-Lemaître Law for an unperturbed fluid, for which $\boldsymbol{v}_{\rm pec} = 0$. Then we substitute Eqs. (1.93) – (1.95) into Eqs. (1.88) – (1.90), neglecting the terms including $(\delta\rho)^2$, $(\delta\phi)^2$, $(\delta\boldsymbol{v})^2$, and the other mixed terms, as we are considering small perturbations. To simplify the computation, solutions are found in terms of Fourier modes, for which the generic form is expressed as follows:

$$f(r,t) = f_k(t)\, e^{ikr}\,, \tag{1.96}$$

where the general expression $f(r,t)$ stands for $\delta\rho$, $\delta\phi$, or $\delta\boldsymbol{v}$, while r is the comoving spatial coordinate, the wavenumber $k = |\boldsymbol{k}|$ is the absolute value of the wavevector $\boldsymbol{k}$, and $f_k(t)$ is the amplitude of the Fourier mode accounting for the expansion of the universe. As we shall discuss, the solving relation for Eqs. (1.88) – (1.90) is a differential equation referred to as the *dispersion relation*, which can be expressed in terms of the density contrast δ, defined in Eq. (1.76).

Perturbations in matter dominated universes

For matter dominated universes, namely for $z < z_{\rm eq}$ and thus $\rho_{\rm b} \propto a^{-3}$, the dispersion relation has the following form:

$$\ddot{\delta}_k + 2\frac{\dot{a}}{a}\,\dot{\delta}_k + \delta_k\left(k^2 v_{\rm s}^2 - 4\pi G\rho_{\rm b}\right) = 0\,, \tag{1.97}$$

where δ_k is the amplitude of the Fourier transform of $\delta(x,t)$ and $v_{\rm s} = \sqrt{\partial p/\partial\rho}$ is the sound speed. Furthermore, in Eq. (1.97), the second term represents a tidal contribution resulting from the expansion of the universe, which counteracts gravity. The second term within the brackets corresponds to the average gravitational effect, while the term $\delta_k k^2 c_s^2$ takes into account the characteristic velocity field of the fluid. Notably, this term hampers the gravitational collapse in terms of pressure, which is contained within the definition of the sound speed. We remark that the wavenumber, k, corresponds to a wavelength, λ, through the relation $k = 2\pi/\lambda$.

In a flat universe, for which the $a(t)$ expression in Eq. (1.86) holds, the background density expressed in Eq. (1.56) takes the following functional form:

$$\rho_{\rm b} = \frac{1}{(1+w)^2\, 6\pi G t^2}\,. \tag{1.98}$$

Thus, from Eqs. (1.86) and (1.98), for a matter dominated universe having $w = 0$ we have

$$\rho_{\rm b} = \frac{1}{6\pi G t^2}\,, \tag{1.99}$$

$$a = a_0 \left(\frac{t}{t_0}\right)^{2/3}\,, \tag{1.100}$$

$$\frac{\dot{a}}{a} = \frac{2}{3t}\,. \tag{1.101}$$

Substituting these equations into Eq. (1.97), the following dispersion relation is obtained:

$$\ddot{\delta}_k + \frac{4}{3t}\,\dot{\delta}_k - \frac{2}{3t^2}\delta_k \left(1 - \frac{k^2 v_{\rm s}^2}{4\pi G \rho_{\rm b}}\right) = 0\,. \tag{1.102}$$

Solutions to Eq. (1.102) are expressed as $\delta_k \propto t^\alpha$, and thus $\delta \propto e^{ikr}\, t^\alpha$. Specifically, we have

$$\delta_k = t^{-\{1\pm 5[1-(6v_{\rm s}^2 k^2)/(25\pi G\rho_{\rm b})]^{1/2}\}/6} = t^{-[1\pm 5(1-\lambda_{\rm J}^2/\lambda^2)^{1/2}]/6}\,, \tag{1.103}$$

where we defined the Jeans length as follows:

$$\lambda_{\rm J} = \frac{\sqrt{24}}{5}\, v_{\rm s} \left(\frac{\pi}{G\rho_{\rm b}}\right)^{1/2}\,. \tag{1.104}$$

For perturbations with $\lambda < \lambda_{\rm J}$, solutions are propagating as waves. For $R_{\rm H} > \lambda > \lambda_{\rm J}$, instead, we have a growing and a decaying solution. We also note that for $R_{\rm H} > \lambda \gg \lambda_{\rm J}$, the growing solution, δ_+, is expressed as

$$\delta_+ \propto t^{2/3} \propto a\,, \tag{1.105}$$

while the decaying solution, δ_-, has the form

$$\delta_- \propto t^{-1} \propto a^{-3/2}\,. \tag{1.106}$$

For a generic universe with cosmological constant, it is possible to derive an approximate form for the variation of the growing solution, namely the *growth factor*, f, expressed as

$$f \equiv \frac{{\rm d}\ln\delta_+}{{\rm d}\ln a} \simeq \Omega_{\rm m}^{0.55}(z) + \frac{\Omega_\Lambda(z)}{70}\left(1 + \frac{\Omega_{\rm m}(z)}{2}\right). \tag{1.107}$$

This expression is extremely important because the exponent 0.55 descends from General Relativity, and therefore the measure of f provides a test for the gravity theory. In addition, f mainly depends on $\Omega_{\rm m}$, and this is reasonable as matter interacts gravitationally. The dependency on the cosmological constant, Λ, which participates to the expansion of the universe, is weaker. Moreover, it has to be noted that $f = 1$ in a matter only flat universe, while the value of f deviates from unity in a ΛCDM universe.

Perturbations in radiation dominated universes

For radiation dominated universes, that is $z > z_{\rm eq}$ and $\rho_{\rm b} \propto a^{-4}$, the dispersion relation is expressed as follows:

$$\ddot{\delta}_k + 2\frac{\dot{a}}{a}\,\dot{\delta}_k + \delta_k \left(k^2 v_{\rm s}^2 - \frac{32}{3}\pi G\rho_{\rm b}\right) = 0\,. \tag{1.108}$$

Thus, from Eqs. (1.86) and (1.98), for a radiation dominated universe having $w = 1/3$ we obtain

$$\rho_{\rm b} = \frac{3}{32\pi G t^2}\,, \tag{1.109}$$

$$a = a_0 \left(\frac{t}{t_0}\right)^{1/2}\,, \tag{1.110}$$

$$\frac{\dot{a}}{a} = \frac{1}{2t}\,. \tag{1.111}$$

Consequently, Eq. (1.108) becomes

$$\ddot{\delta}_k + \frac{\dot{\delta}_k}{t} - \frac{\delta_k}{t^2}\left(1 - \frac{3k^2 v_{\rm s}^2}{32\pi G\rho_{\rm b}}\right) = 0\,. \tag{1.112}$$

Looking for power-law solutions, the Jeans scale is derived as

$$\lambda_{\rm J} = v_{\rm s}\left(\frac{3\pi}{8G\rho_{\rm b}}\right)^{1/2}\,. \tag{1.113}$$

From Eq. (1.109), and recalling that $v_{\rm s} = c/\sqrt{3}$ for a radiative fluid, the Jeans scale takes the following form:

$$\lambda_{\rm J} = \frac{2\pi c}{\sqrt{3}}\,t\,. \tag{1.114}$$

As we discussed in Sect. 1.6.5, by considering $R_{\rm H}$ for a flat universe dominated by radiation (see Eq. 1.72), we have $R_{\rm H} \simeq 2ct$. Therefore, the Jeans scale is larger than $R_{\rm H}$ at $t < t_{\rm eq}$, and in turn radiative gravitational instabilities can not arise inside the horizon.

Dark and baryonic matter perturbations

So far, we focused on universe models including only one component. However, studying the perturbation growth for different components is key to understand the formation of galaxies. At $t < t_{\rm eq}$, dark matter instabilities can take place after the decoupling, despite this is not possible for radiation perturbations. For a dark matter perturbation with $R_{\rm H} > \lambda \gg \lambda_{\rm J}$, it is possible to show that the perturbation amplitude, $\delta_k^{\rm DM}$, at $t < t_{\rm eq}$ is expressed as

$$\delta_k^{\rm DM} = 1 + \frac{3}{2}\frac{a}{a_{\rm eq}}\,, \tag{1.115}$$

where $a_{\rm eq}$ is the cosmic scale factor at $t_{\rm eq}$. Thus, $\delta_k^{\rm DM}(t_{\rm eq})/\delta_k^{\rm DM}(t_{\rm H}) < 5/2$, where $t_{\rm H}$ is the time at which a perturbation enters $R_{\rm H}$. Consequently, for $t < t_{\rm eq}$, dark matter perturbations within $R_{\rm H}$ can grow at most by a factor 5/2. Therefore, dark matter gravitational instabilities take place before equivalence and are not erased, but their growth is severely hindered inside the horizon. This stagnation of the dark matter perturbations is commonly referred to as the *Meszaros effect* (Meszaros 1974).

For $t > t_{\rm eq}$, the evolution of dark matter gravitational instabilities is described by Eq. (1.105). On the other hand, baryonic matter perturbations arise after the decoupling between baryons and radiation, namely at $t > t_{\rm dec}$, where $t_{\rm dec} > t_{\rm eq}$. In this case, the baryonic perturbation amplitude, $\delta_k^{\rm B}$, at $t > t_{\rm dec}$ is expressed as

$$\delta_k^{\rm B} = \delta_k^{\rm DM}\left(1 - \frac{a_{\rm dec}}{a}\right)\,, \tag{1.116}$$

where $a_{\rm dec}$ is the cosmic scale factor at $t_{\rm dec}$. From Eq. (1.116), we can see that baryonic perturbations rapidly grow after decoupling, with $\delta_k^{\rm B} \to \delta_k^{\rm DM}$ for $t \to t_0$. This effect is called *baryon catch up*, and it takes place thanks to the presence of dark matter perturbations at $t_{\rm dec}$. Indeed, baryonic matter falls into such dark matter haloes, and baryonic and dark matter density contrasts tend to become equal.

In addition, the decoupling between dark matter and radiation happens before the equivalence, as the dark matter-radiation collisional time scale becomes larger than the universe expansion time scale before $t_{\rm eq}$ (Coles & Lucchin 2002). In addition, cold dark matter becomes not relativistic before its decoupling from radiation, allowing for the prompt formation of virialised haloes. This does not hold for hot dark matter, which is still relativistic after its decoupling. Therefore, cold dark matter is the suitable candidate for the hierarchical formation of cosmic structures.

1.8.3 Spherical collapse

The formation and evolution of cosmic structures, such as galaxies and galaxy clusters, can not be accurately described through linear theory. A spherical perturbation at constant density represents the only case in which the evolution of a virialised structure can be analytically computed. This is the basis of the *spherical collapse* model, which is a good approximation for galaxy clusters and leads to the Press-Schechter formalism for the mass function (see Sect. 3.3.1).

In particular, this model is based on the assumption of a small perturbation, so that it can be considered as spherical. The perturbation is characterised in terms of a closed universe, and at the initial time, t_i, the density contrast, δ_i, is small, namely $0 < \delta_i \ll 1$. As the perturbation expands with the background universe, the initial peculiar velocity at its edge is null, with a Fourier mode $v_{k,i} = 0$. In addition, the background universe is assumed to be flat and composed of matter only.

As discussed in Sect. 1.8.2, the density contrast evolution is characterised by the combination of growing, δ_+, and decaying, δ_-, modes. In a flat universe containing only matter, the initial density contrast, δ_i, is expressed as follows (see Eqs. 1.105 and 1.106):

$$\delta_i = \delta_+(t_i) \left(\frac{t}{t_i}\right)^{2/3} + \delta_-(t_i) \left(\frac{t}{t_i}\right)^{-1} . \tag{1.117}$$

Assuming null velocities at t_i provides the following relation (Coles & Lucchin 2002)

$$\delta_+(t_i) = \frac{3}{5}\delta_i \,, \tag{1.118}$$

thus the remaining 2/5 of δ_i decays with time, and tends to become negligible. As the collapse is described in terms of a closed universe, the perturbation density parameter, $\Omega_{\rm p}$, is larger than 1. In particular, we have

$$\Omega_{\rm p}(t_i) = \frac{\rho_{\rm p}(t_i)}{\rho_{\rm c}(t_i)} = \frac{\rho_{\rm b}(t_i)(1+\delta_i)}{\rho_{\rm c}(t_i)} = \Omega_{\rm b}(t_i)(1+\delta_i) > 1 \,, \tag{1.119}$$

where $\rho_{\rm p}$ is the perturbation density, $\rho_{\rm c}$ is the critical density, and $\rho_{\rm b}$ is the background density. Consequently, we have $(1+\delta_i) > 1/\Omega_{\rm b}(t_i)$, and from the relation describing the evolution of the density parameter, Eq. (1.66), we obtain

$$\delta_+(t_i) = \frac{3}{5}\delta_i > \frac{3}{5}\frac{1-\Omega_{0,\rm b}}{\Omega_{0,\rm b}(1+z)} \,, \tag{1.120}$$

where $\Omega_{0,\mathrm{b}}$ is the background density parameter at the current cosmic time, t_0. From Eq. (1.120), we note that in closed background universes, having $\Omega_{0,\mathrm{b}} > 1$, any positive density contrast leads to gravitational collapse. Conversely, in open background universes, with $\Omega_{0,\mathrm{b}} < 1$, the expansion inhibits the collapse if δ_i is not sufficiently greater than zero.

From Friedmann equations, one can demonstrate that the perturbation expands following the expansion of the background universe until it reaches a maximum physical scale, R_{max}, at the time t_{max}. Subsequently, the perturbation experiences a *turn-around*, and the gravitational collapse begins. The growing mode at t_{max} is expressed as follows:

$$\delta_+(t_{\mathrm{max}}) \simeq \frac{\rho_{\mathrm{p}}(t_{\mathrm{max}})}{\rho_{\mathrm{b}}(t_{\mathrm{max}})} - 1 \simeq 4.6\,. \tag{1.121}$$

Thus, as the linear regime holds for $\delta \ll 1$, we can see that a nonlinear perturbation has formed at t_{max}. On the other hand, by extrapolating the linear evolution of δ_+, one would obtain $\delta_+(t_{\mathrm{max}}) \simeq 1.07$. For $t > t_{\mathrm{max}}$, the scale of the perturbation decreases until $t_{\mathrm{coll}} \simeq 2t_{\mathrm{max}}$, when the collapse is arrested by internal thermal motions. Numerical simulations show that the virialisation is reached at the time $t_{\mathrm{vir}} \simeq 3t_{\mathrm{max}}$. From the scalar form of the virial theorem, the total energy of the perturbation, E_{tot}, is expressed as follows

$$E_{\mathrm{tot}}(t_{\mathrm{vir}}) = -\frac{1}{2}\frac{3}{5}\frac{GM^2}{R_{\mathrm{vir}}}, \tag{1.122}$$

where R_{vir} is the virial radius. Assuming that there are no mass and energy losses since the time t_{max}, namely $E_{\mathrm{tot}}(t_{\mathrm{max}}) = E_{\mathrm{tot}}(t_{\mathrm{vir}})$, we obtain the following relation:

$$-\frac{1}{2}\frac{3}{5}\frac{GM^2}{R_{\mathrm{vir}}} = -\frac{3}{5}\frac{GM^2}{R_{\mathrm{max}}}, \tag{1.123}$$

because at $t = t_{\mathrm{max}}$ only the potential energy contributes to E_{tot}. Thus, $R_{\mathrm{vir}} = R_{\mathrm{max}}/2$, and the perturbation density at $t = t_{\mathrm{vir}}$ is $\rho_{\mathrm{p}}(t_{\mathrm{vir}}) = 8\rho_{\mathrm{p}}(t_{\mathrm{max}})$. Then we obtain the following relations:

$$\delta_+(t_{\mathrm{coll}}) \simeq \frac{8\rho_{\mathrm{p}}(t_{\mathrm{max}})}{\rho_{\mathrm{b}}(t_{\mathrm{max}})}\left(\frac{t_{\mathrm{coll}}}{t_{\mathrm{max}}}\right)^2 \simeq 180, \tag{1.124}$$

$$\delta_+(t_{\mathrm{vir}}) \simeq \frac{8\rho_{\mathrm{p}}(t_{\mathrm{max}})}{\rho_{\mathrm{b}}(t_{\mathrm{max}})}\left(\frac{t_{\mathrm{vir}}}{t_{\mathrm{max}}}\right)^2 \simeq 400. \tag{1.125}$$

We stress that these results depend on the background cosmology, which in this case is a matter-only flat universe. By contrast, the extrapolation of the linear growth, for which $\delta_+ \propto t^{2/3}$, provides the following results:

$$\delta_+(t_{\mathrm{coll}}) \simeq \frac{3}{5}\left(\frac{3\pi}{4}\right)^{2/3}\left(\frac{t_{\mathrm{coll}}}{t_{\mathrm{max}}}\right)^{2/3} \simeq 1.68, \tag{1.126}$$

$$\delta_+(t_{\mathrm{vir}}) \simeq \frac{3}{5}\left(\frac{3\pi}{4}\right)^{2/3}\left(\frac{t_{\mathrm{vir}}}{t_{\mathrm{max}}}\right)^{2/3} \simeq 2.2\,. \tag{1.127}$$

1.8.4 N-body simulations

Due to the intricate and nonlinear nature of the processes involved in the formation of virialised structures, the analytical study of the generic evolution of perturbations in the nonlinear regime is not possible. As we discussed in Sect. 1.8.3, the spherical collapse model provides a fair approximation for the formation of dark matter haloes hosting galaxy clusters. Nevertheless, the spherical collapse formalism relies on simplistic assumptions, such as a background

flat universe containing only matter and a null initial peculiar velocity at the edge of the perturbation. Thus, numerical simulations are needed to make detailed predictions of the large-scale structure of the Universe, to be tested against observations.

N-body simulations containing only dark matter are excellent tools to accurately reproduce galaxy cluster observations (see, e.g., Angulo & Hahn 2022, and references therein). Indeed, as we discussed in Sect. 1.7, dark matter is the dominant matter component in the Universe. In addition, due to the collisionless nature of dark matter, N-body simulations exclusively consider the gravitational interactions between particles. This greatly streamlines the creation of large simulated volumes of universe, as hydrodynamical effects, which demand extensive computational resources, are neglected.

The initial conditions for the matter density field are established based on inflation theories. Subsequently, within the framework of the Newtonian approximation, the following set of equations needs to be solved:

$$\begin{cases} \boldsymbol{F}_i = GM_i \sum\limits_{i \neq j} \dfrac{M_j}{r_{ij}^2} \hat{\boldsymbol{r}}_{ij} \\ \ddot{\boldsymbol{x}}_i = \dfrac{\mathrm{d}^2 \boldsymbol{x}_i}{\mathrm{d}t^2} = \dfrac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} = \dfrac{\boldsymbol{F}_i}{M_i} \\ \dot{\boldsymbol{x}}_i = \dfrac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}t} = \boldsymbol{v}_i \end{cases} \tag{1.128}$$

where M is the particle mass, $\boldsymbol{F}_i$ is the gravitational force acting on the ith particle, $\boldsymbol{x}_i$ and $\boldsymbol{v}_i$ are the comoving coordinates and the velocity of the ith particle, respectively, r_{ij} is the comoving distance between ith and jth particles, and $\hat{\boldsymbol{r}}_{ij}$ is the related versor. In addition, base particles have the same mass in simulations. At a given time t, the force $\boldsymbol{F}_i$ is computed for each particle. Subsequently, after a time step δt, the new velocities and positions, $\boldsymbol{v}_i(t+\delta t)$ and $\boldsymbol{x}_i(t+\delta t)$, are calculated by solving the system (1.128). The force is recalculated for these updated positions, and this iterative procedure is repeated at every time step. This methodology constitutes the foundation of the particle-particle (PP) method, which stands as the most basic technique utilised in N-body simulations. Although PP is very accurate, its application implies several issues. In a first instance, as the gravitational force diverges at null distances, the definition of a *softening length*, ϵ, is required. In particular, ϵ defines the minimum particle size, and it is a free parameter. In addition, the PP method incurs high computational costs, as it necessitates calculating the distances between all particle pairs, which amounts to $N(N-1)/2$ computations for each time step. This results in a computational complexity scaling as $\mathcal{O}(N^2)$.

An alternative to PP is the particle-mesh (PM) method. In this case, the gravitational potential is calculated on a regular grid under periodic boundary conditions, enabling the use of Fast Fourier Transform (FFT) techniques. The process involves interpolating particle positions to compute a grid-based density field, from which the density FFT is determined. By leveraging the Poisson equation, the gravitational potential and force in Fourier space are derived across the grid. Performing an inverse FFT provides the force in real space on the grid, followed by particle-specific force calculation through interpolation. Since a significant portion of the computations occurs in Fourier space, the operational complexity scales as $\mathcal{O}(N \log N)$, outperforming the PP approach. However, the spatial resolution is degraded.

Another technique adopted in N-body simulations is the so-called hierarchical tree (HT). It consists in partitioning the simulated universe volume into cells, with their sizes organized hierarchically according to particle spatial distribution. Initially, the volume is divided into large cells, and then each cell can be further subdivided if it contains more than one particle. This subdivision process continues until cells host one particle or are empty. For a given

particle, named P, the gravitational interaction is computed with individual particles that are in close proximity to P. On the other hand, when dealing with distant clusters of particles, it is sufficient to compute the force generated by a cell containing the cluster, whose barycentre is considered as the distance between the cluster and particle P. With this method, the number of involved operations scales as $\mathcal{O}(N \log N)$.

As we shall detail in Chapter 3, N-body simulations provide accurate predictions on the formation and evolution of galaxy clusters. Indeed, the structure of these objects is dominated by dark matter, and hydrodynamical effects caused by baryonic matter imply second-order effects on their counts and clustering. Nevertheless, a realistic description of galaxy evolution can be attained only by accounting for hydrodynamical effects, such as gas cooling, shocks, turbulence, and magnetic fields. To this aim, several hydrodynamical simulations have been developed over the past years (see, e.g., Le Brun et al. 2014; Crain et al. 2015; McCarthy et al. 2017; Springel et al. 2018; Davé et al. 2019). However, such simulations are computationally expensive, thus they can be performed for relatively small volumes of universe. In addition, due to the uncertainties on galaxy formation, many specific choices on galaxy initial properties have to be made, and the modelling of physical processes is fine-tuned according to the simulation resolution. Consequently, the theoretical uncertainties on galaxy formation and evolution are still large. In the context of cosmological analyses based on galaxy statistics, such uncertainties can be encapsulated in a free bias parameter entering the matter power spectrum, as we shall discuss in Sect. 2.1.

Chapter 2

Statistical properties of the Universe

In this Chapter, we delve into the statistical properties of the large-scale structure of the Universe, which play a crucial role in observational cosmology. Specifically, we discuss the evolution of the matter power spectrum and correlation function, providing a description of the distortions showing up in these statistical functions due to galaxy proper motions and cosmological assumptions. In addition, we present the basic theory of gravitational lensing, as the correlation functions involving galaxy image distortions are among the most important probes in contemporary cosmology. Furthermore, gravitational lensing plays a key role in estimating the mass of galaxy clusters, which is a fundamental step for the cosmological investigations presented in this thesis.

2.1 Correlation function and power spectrum

The distribution of matter in the large-scale structure of the Universe is not random. Instead, galaxy positions are correlated, and this behaviour is named *clustering*. To describe this spatial correlation, let us consider the probability to find an object in an infinitesimal comoving volume $\mathrm{d}V_1$, namely

$$\mathrm{d}P_1 = \bar{n}\,\mathrm{d}V_1\,, \tag{2.1}$$

where $\bar{n}$ is the mean number density of objects, which is independent of position. The joint probability to find an object in two small volumes, namely $\mathrm{d}V_1$ and $\mathrm{d}V_2$, separated by a comoving distance r, is expressed as follows

$$\mathrm{d}P_{12} = \bar{n}^2\,\mathrm{d}V_1\mathrm{d}V_2\,[1+\xi(r)]\,, \tag{2.2}$$

where $r \in \mathbb{R}$, since we assumed homogeneity and isotropy. In Eq. (2.2), $\xi(r)$ is the two-point correlation function (2PCF), and we have $\xi(r) = 0$ if the distribution is random, $\xi(r) > 0$ if positions are correlated, and $-1 \leq \xi(r) < 0$ if positions are anticorrelated. In addition, we remark that $\xi(r)$ is dimensionless. N-point correlation functions with $N > 2$ can be defined in a similar way (Peebles 1980).

Now we study the case of a continuous density of objects with mass m, namely $\rho(x) = n(x)m$, having mean equal to $\bar{\rho} = \bar{n}m$. In this case, we have:

$$\mathrm{d}P_{12}(r) = \langle n(x+r)n(x)\rangle\,\mathrm{d}V_1\mathrm{d}V_2 = \bar{n}^2\,\frac{\langle\rho(x+r)\rho(x)\rangle}{\bar{\rho}^2}\,\mathrm{d}V_1\mathrm{d}V_2\,, \tag{2.3}$$

where 1 and 2 refer now to two different positions. Consequently, from Eq. (2.2), we have

$$1 + \xi(r) = \frac{\langle \rho(x+r)\rho(x) \rangle}{\bar{\rho}^2} \,. \tag{2.4}$$

To estimate ξ from observations, let us consider an observed catalogue containing N_D objects and a random sample containing N_R objects having uncorrelated positions. The so-called natural estimator of ξ, namely ξ_N, is expressed as follows (Peebles & Hauser 1974):

$$\xi_N(r) = \frac{N_{RR}}{N_{DD}} \frac{DD(r)}{RR(r)} - 1 \,, \tag{2.5}$$

where $DD(r)$ and $RR(r)$ are the number of data-data and random-random pairs with separation $r \pm \Delta r$, respectively, while N_{DD} and N_{RR} are the total number of data-data and random-random pairs, respectively. Due to the discreteness of the sample, the estimator in Eq. (2.5) is affected by low accuracy at large scales. In general, more precise estimators that take into account the cross terms between the data and random catalogues are used. One of such estimators is that proposed by Landy & Szalay (1993, LS), namely ξ_{LS}, having the following expression:

$$\xi_{LS}(r) = \frac{N_{RR}}{N_{DD}} \frac{DD(r)}{RR(r)} - 2\frac{N_{RR}}{N_{DR}} \frac{DR(r)}{RR(r)} + 1 \,, \tag{2.6}$$

where $DR(r)$ is the number of data-random pairs with separation $r \pm \Delta r$, and N_{DR} is the total number of data-random pairs. In particular, ξ_{LS} has a nearly Poissonian variance and provides an unbiased estimate of the 2PCF for $N_R \to \infty$, with minimum variance.

From the definition of density contrast, δ, in Eq. (1.76), the 2PCF in Eq. (2.4) can be expressed as

$$\xi(r) = \langle \delta(x+r)\delta(x) \rangle \,. \tag{2.7}$$

We remark that $\xi(r)$ is an auto-correlation, as it describes the correlation of the same class of cosmic structure tracers, and it is defined in three space dimensions. Cross-correlation functions between different tracers can be also defined (Peebles 1980), and correlation functions in two space dimensions, obtained by integrating over redshift, are widely used in cosmological surveys when accurate measurements of redshift are not available, like in photometric surveys (Laureijs et al. 2011).

The general expression of the configuration-space density contrast, $\delta(\boldsymbol{x}) \equiv \delta\rho(\boldsymbol{x})/\rho$, in terms of its Fourier counterpart is

$$\delta(\boldsymbol{x}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \delta(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}^3\boldsymbol{k} \,. \tag{2.8}$$

Specifically, as discussed in Sect. 1.8.2, $\delta(\boldsymbol{x})$ is the Fourier antitransform of $\delta(\boldsymbol{k})$. The quantity $\delta(\boldsymbol{x})$ has no dimension, while $\delta(\boldsymbol{k})$ has the dimension of a volume. In addition, $\delta(\boldsymbol{x})$ is assumed to be a continuum and stochastic field, originated from the primordial fluctuations produced at the end of the inflationary era (see Sect. 1.7). Our focus is not on the precise value of $\delta(\boldsymbol{x})$ at point $\boldsymbol{x}$, but rather on examining the average properties of this field. As we mentioned in Sect. 1.1, the cosmological principle entails the assumption of the ergodic principle, or *fair sample* principle, stating that the mean of a quantity from many realisations of the Universe is equivalent to the mean obtained by considering separate, and sufficiently large, volumes in the Universe. Their size is determined by the cosmic time and the degree of nonlinearity of the perturbations, which increases over time. We note that inflation produces random fluctuations in $\delta(\boldsymbol{k})$, which is an imaginary quantity. In turn, $\delta(\boldsymbol{k})$ phases are stochastically

distributed, and imaginary numbers with random phases follow a Gaussian distribution. Therefore, $\delta(\boldsymbol{x})$ is Gaussian.

From Eqs. (2.7) and (2.8), ξ can be expressed as follows:

$$\xi(r) = \frac{1}{(2\pi)^6} \int \mathrm{d}^3\boldsymbol{k} \int \mathrm{d}^3\boldsymbol{k}' \langle \delta(\boldsymbol{k})\, \delta(\boldsymbol{k}') \rangle e^{i\boldsymbol{k}\cdot(\boldsymbol{x}+\boldsymbol{r})} e^{i\boldsymbol{k}'\cdot\boldsymbol{r}}\,, \tag{2.9}$$

and from the Wiener-Khinchin theorem, which states that the 2PCF is the inverse Fourier transform of the power spectrum, we define

$$\langle \delta(\boldsymbol{k})\, \delta(\boldsymbol{k}') \rangle \coloneqq (2\pi)^3 P(\boldsymbol{k})\, \delta_D^{(3)}(\boldsymbol{k}+\boldsymbol{k}')\,, \tag{2.10}$$

where $P(\boldsymbol{k})$ is the *power spectrum* and $\delta_D^{(3)}(\boldsymbol{k}+\boldsymbol{k}')$ is the three-dimensional Dirac delta. Specifically, we have

$$\xi(r) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\boldsymbol{k}\, P(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\,. \tag{2.11}$$

In general the power spectrum is $P(k) \propto \langle |\delta(k)|^2 \rangle$, because

$$\langle \delta(\boldsymbol{k})\, \delta^*(\boldsymbol{k}) \rangle \equiv \langle |\delta(\boldsymbol{k})|^2 \rangle \coloneqq (2\pi)^3 P(\boldsymbol{k})\, \delta_D^{(3)}(\boldsymbol{k}-\boldsymbol{k})\,, \tag{2.12}$$

where $\delta^*(\boldsymbol{k})$ in the complex conjugate of $\delta(\boldsymbol{k})$, and $\delta^*(\boldsymbol{k}) = \delta(-\boldsymbol{k})$. Similarly to $\xi(r)$, $P(k)$ depends only on the modulus of $\boldsymbol{k}$, namely k, for the isotropy assumption.

As $\delta(\boldsymbol{x})$ is normally distributed, it is fully described by its mean and variance. It is possible to show that the variance, σ^2, has the following functional form:

$$\sigma^2 \coloneqq \langle \delta^2(\boldsymbol{x}) \rangle = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\boldsymbol{k}\, P(\boldsymbol{k}) = \frac{1}{2\pi^2} \int \mathrm{d}k\, k^2 P(k)\,. \tag{2.13}$$

From the operational point of view, to derive σ^2 we must account for the fact that the density field is not continuous, and that we observe distributions of objects, such as galaxies or galaxy clusters, that trace the underlying total matter distribution. Thus, we define the mass density contrast, δ_M, as follows:

$$\delta_M(\boldsymbol{x}) \coloneqq \delta(\boldsymbol{x}) * W(\boldsymbol{x}, R)\,, \tag{2.14}$$

which is a convolution of $\delta(\boldsymbol{x})$ and the window function $W(\boldsymbol{x}, R)$, depending on the physical radius $R = (3M/4\pi\rho_\mathrm{m})^{1/3}$, where M is the mass. $W(\boldsymbol{x}, R)$ is usually assumed to be a top-hat or Gaussian function. Therefore, the mass variance, σ_M^2, has the following functional form:

$$\sigma_M^2 = \frac{1}{2\pi^2} \int \mathrm{d}k\, k^2 P(k) \hat{W}^2(k, R)\,, \tag{2.15}$$

where $\hat{W}(\boldsymbol{k}, R)$ is the Fourier transform of the window function. For high values of R, or analogously for high values of M, large scales are filtered and the mass variance σ_M^2 tends to zero. Conversely, for small values of R we have $\sigma_M^2 \to \sigma^2$. In addition, galaxies and galaxy clusters are biased tracers of the underlying matter distribution, thus the density contrast of the tracer, δ_trc, can be approximated as follows

$$\delta_\mathrm{trc} = b(M, z)\delta_M\,, \tag{2.16}$$

where $b(M, z)$ is the tracer bias, depending on the mass and redshift of the tracers. In addition, $b(M, z)$ depends on cosmological parameters, and it grows with increasing mass and redshift. Functional forms for $b(M, z)$ can be obtained analytically and then calibrated through N-body simulations (Mo & White 1996; Sheth et al. 2001; Tinker et al. 2010), holding for dark matter haloes. Such analytic bias functions are good approximations for galaxy clusters, as the structure of these objects is dominated by dark matter and their evolution can be accurately described through dark matter-only N-body simulations. In cosmological analyses based on galaxy statistics, instead, $b(M, z)$ is a free parameter due to the large theoretical uncertainties on galaxy formation and evolution (see Sect. 1.8.4).

2.2 Linear evolution of the power spectrum

In Sect. 1.8.2 we showed that, in linear regime, matter perturbations grow with time without any dependence on the scale k. In addition, inflation does not create perturbations on privileged scales, thus we can assume that the initial power spectrum, $P_{\rm in}$, is expressed as a power-law:

$$P_{\rm in}(k) = A_{\rm s} k^{n_{\rm s}} \,, \tag{2.17}$$

where $A_{\rm s}$ is the normalisation of the primordial power spectrum, and $n_{\rm s}$ is the spectral index. As we discussed in Sect. 1.7, both $A_{\rm s}$ and $n_{\rm s}$ are fundamental parameters of the ΛCDM cosmological model. Inflationary models predict $n_{\rm s} \simeq 1$, which in terms of gravitational potential translates into perturbations with the same quadratic amplitude on all scales. In agreement with this theoretical prediction, Planck Collaboration VI (2020) derived $n_{\rm s} \simeq 0.96$. CMB observations also allow for constraining the slow-roll parameters ϵ and η defined in inflationary models (Linde 2007), where the spectral index is expressed as $n_{\rm s} = 1 + 2\epsilon - 6\eta$.

As $P(k) \propto \langle|\delta(k)|^2\rangle$, in linear regime the power spectrum is expressed as a constant shift of the initial power spectrum:

$$P(k,t) = \delta_+^2(t) P_{\rm in}(k) \,, \tag{2.18}$$

where $\delta_+(t)$ is the growing mode, and its time evolution can be derived through the approximated expression of the growth factor, f, in Eq. (1.107). As discussed in Sect. 1.8.2, matter perturbations that enter the horizon before the matter-radiation density equivalence experience the Meszaros effect, which nearly inhibits their growth. Thus, before equivalence, perturbations that enter the horizon can no longer undergo significant growth. On the other hand, perturbations on scales larger than the horizon continue to grow at the same rate, resulting in a constant shift of the primordial power spectrum. Consequently, the power spectrum at the equivalence epoch exhibits a peak in correspondence of $k_{\rm eq}$, namely the scale of the cosmological horizon at the equivalence, and this effect is characterised by the transfer function, $T(k)$. Specifically, the power spectrum is expressed as follows:

$$P(k,t) = \delta_+^2(t) P_{\rm in} T^2(k) = A_{\rm s} k^{n_{\rm s}} \delta_+^2(t) T^2(k), \tag{2.19}$$

where, for cold dark matter, $T^2(k) \propto k^0$ for $k < k_{\rm eq}$, and $T^2(k) \propto k^{-4}$ for $k > k_{\rm eq}$. In Fig. 2.1, left panel, we show the shape of $P(k)$ and its evolution with redshift.

Additionally, nonlinearities become relevant on small physical scales first, or equivalently at large wavenumbers k. This descends from the fact that σ_M^2 increases with decreasing physical scale, and we can assume that nonlinearities hold for $\sigma_M^2 \gtrsim 1$. Indeed, as discussed in Sect. 1.8.2, the linear regime holds for $\delta \ll 1$, and $\sigma_M \propto \delta$ as outlined in Eq. (2.13). As σ_M^2 increases, the probability distribution of the density contrast tends to deviate from a Gaussian shape. The original Gaussian distribution, centred around zero, can not extend to density contrasts lower than -1, resulting in a positive skewness in the distribution. This skewness tends to favour the formation of underdensities, and this explains the existence of the *cosmic web*. Furthermore, we remark that nonlinearities affect the power spectrum at high values of k. As a result, $P(k)$ in the linear regime is reflected in observable structure statistics mainly at high redshifts, or in the late Universe on large scales. Notably, in the local Universe, nonlinearities are attained on scales of approximately 8 h^{-1}Mpc. Consequently, cosmological analyses based on galaxy and galaxy cluster statistics constrain the parameter σ_8^2, which is the mass variance computed on scales of 8 h^{-1}Mpc at redshift $z = 0$. In particular, σ_8^2 is often assumed as the normalisation of the linear power spectrum, replacing $A_{\rm s}$ in Eq. (2.19). Operationally, the normalised linear power spectrum, $P_{\rm norm}$, is expressed as follows:

$$P_{\rm norm}(k,z) = \frac{\sigma_8^2}{\sigma_8^2(z)} P(k,z|A_{\rm s}=1) \,, \tag{2.20}$$

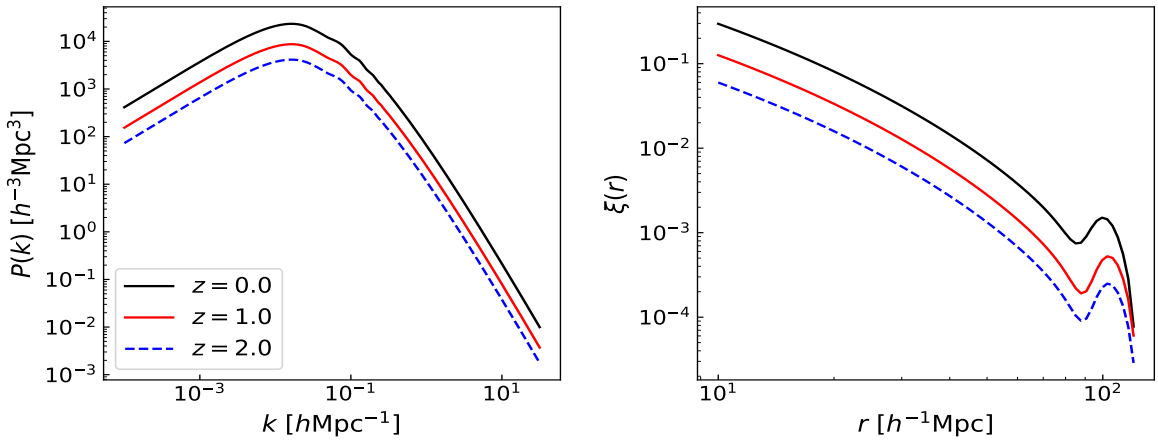

Figure 2.1: Linear matter power spectrum (left panel) and 2PCF (right panel), computed with CAMB (Lewis & Challinor 2011), at $z = 0$ (solid black lines), $z = 1$ (solid red lines), and $z = 2$ (dashed blue lines). The peak of $P(k)$ is in correspondence of $k_{\rm eq}$. As discussed in Sect. 2.3, the $P(k)$ wiggles at $k > k_{\rm eq}$ and the $\xi(r)$ bump at $r \simeq 100\ h^{-1}$Mpc are caused by the BAO.

where $\sigma_8^2(z)$ is computed at redshift z.

2.3 Baryon acoustic oscillations

In the '90s, observations of galaxy and quasar distributions showed an excess of clustering at scales around 100 h^{-1}Mpc (Broadhurst et al. 1990; Deng et al. 1994). This feature is described by the baryon acoustic oscillations (BAO) theory (Eisenstein & Hu 1998). Specifically, prior to recombination and decoupling, the Universe was composed of a high-temperature plasma of photons and baryons that were strongly coupled through Thomson scattering. The competition between radiation pressure and gravity induced oscillations within this fluid. In addition, in the baryon-photon plasma, density fluctuations propagated outwards as acoustic waves, with velocity equal to the sound speed, $c_{\rm s}$. At recombination the Universe became neutral, leading to the removal of radiative pressure on the baryons. As a result, photons started to freely propagate, giving rise to the CMB that we observe today, while the baryon waves stalled. The characteristic radius of the spherical shell formed during recombination, namely $s \sim 100\ h^{-1}$Mpc, is imprinted on the baryon distribution as a density excess. In particular, s is similar to the sound horizon at the time of recombination. As baryons and dark matter interact gravitationally, dark matter also tends to clump preferentially on this scale. Consequently, galaxies tend to form in correspondence of such density excess. Notably, since BAO take place within the cosmological horizon, their footprint in the matter $P(k)$ is observed at $k > k_{\rm eq}$. The BAO feature in both $P(k)$ and $\xi(r)$ is shown in Fig. 2.1.

Similar to the highest peak of the CMB angular power spectrum, whose observations showed that the Universe geometry is consistent with being flat (de Bernardis et al. 2000; Roos & Harun-or-Rashid 2000), the BAO scale present in galaxy clustering is a statistical standard ruler (Bassett & Hlozek 2010; Cuceu et al. 2019). In turn, BAO allow us to constrain the Hubble constant, as well as the matter and dark energy density parameters. Indeed, we recall that the angular diameter distance is expressed as follows

$$d_{\rm A}(z) := \frac{s}{\Delta\theta} \propto \int_0^z \frac{{\rm d}z'}{H(z')}\,, \tag{2.21}$$

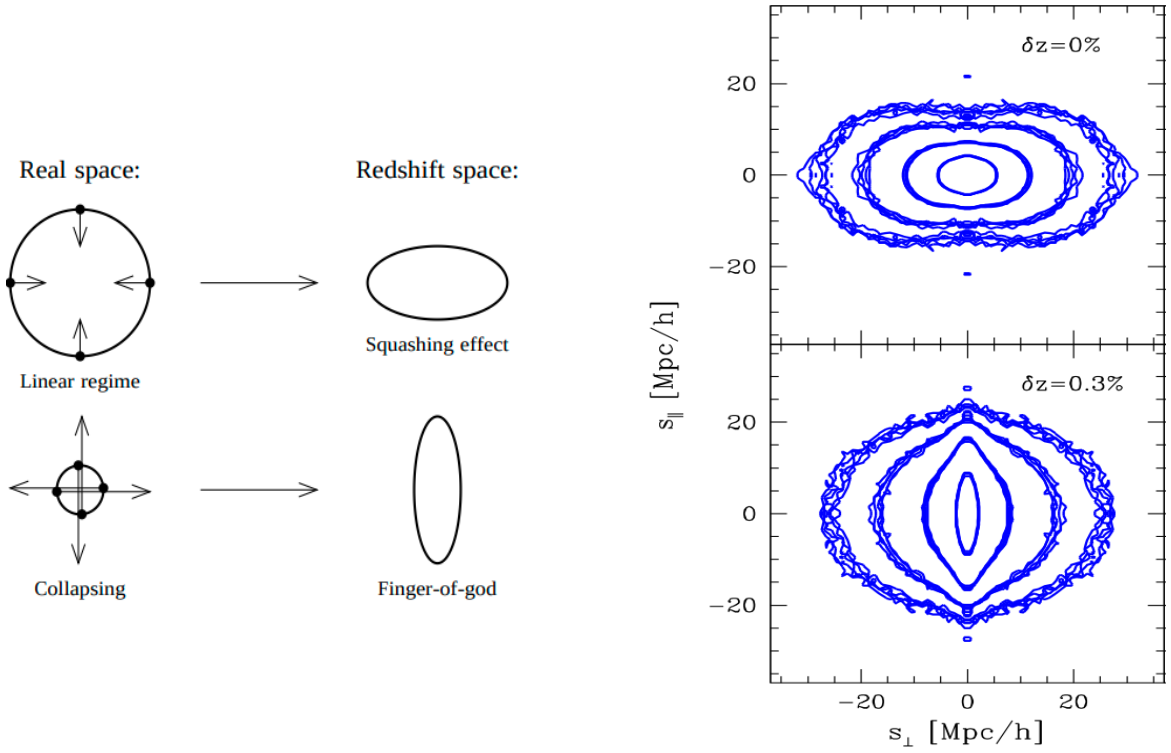

Figure 2.2: *Left panel*: Matter perturbations in real-space are shown on the left, while on the right the effects of redshift measurements are shown. On top, we represent the apparent squashing of a perturbation due to the Kaiser effect on large scales. On the bottom, we show the fingers-of-God effect caused by incoherent motions on small scales and redshift statistical errors. Image from Hamilton (1998). *Right panel*: Iso-correlation contours of $\xi(s_\perp, s_\parallel)$ derived from mock dark matter halo catalogues. On top, redshift errors, quantified by the parameter δz, are null and only the Kaiser effect squashing is present. In the bottom panel, $\delta z = 0.3\%$ and the fingers-of-God feature is evident. Image from Marulli et al. (2012).

where $\Delta\theta$ is the angle subtended by s. From Eq. (1.64), we know that the Hubble parameter, $H(z)$, depends on H_0 and on the density parameters.

2.4 Redshift-space distortions

When studying the large-scale structure of the Universe, it is necessary to account for the peculiar velocities of cosmic tracers, causing distortions in the observed clustering (Kaiser 1987; Hamilton 1998). Indeed, measured redshifts include the contributions of both cosmic expansion and peculiar velocities. The latter induce distortions in the measured comoving distances, $\boldsymbol{s}$, also referred to as redshift-space coordinates. This effect is commonly referred to as redshift-space distortions (RSD), and can be marginalised over by measuring the clustering as a function of angular separation only. However, as we shall discuss in the following, RSD provide robust cosmological information.

To characterise the impact of RSD, $\boldsymbol{s}$ can be decomposed into two components perpendicular and parallel to the line-of-sight, namely $\boldsymbol{s} = (s_\perp, s_\parallel)$. Without the presence of peculiar velocities, the mapping between real-space coordinates, $\boldsymbol{r}$, and $\boldsymbol{s}$ is an identity, and the iso-correlation contours of $\xi(s_\perp, s_\parallel)$ are circles. On the other hand, RSD squash such iso-correlation contours along $s_\parallel$, as we show in Fig. 2.2, due to coherent bulk flows from underdense regions to high density regions. This is the so-called Kaiser effect (Kaiser 1987),

which can be approximated as a linear effect enhancing the clustering on all scales. Specifically, the power spectrum accounting for RSD caused by the Kaiser effect has the following expression (Kaiser 1987; Hamilton 1992; Fisher et al. 1994):

$$P_{\rm RSD}(k,\mu) = P(k)\left(b + f\mu^2\right)^2 , \tag{2.22}$$

where $k = \sqrt{k_\perp^2 + k_\parallel^2}$, with $k_\parallel$ and $k_\perp$ being the wave-vector components parallel and perpendicular to the line of sight, respectively, $\mu := k_\parallel/k$, f is the growth rate of matter perturbations, and b is the tracer bias. Consequently, RSD provide robust cosmological information, as they trace the growth of matter perturbations (see, e.g., Peacock et al. 2001; da Ângela et al. 2005; Marulli et al. 2012; Tonegawa et al. 2020). Improvements of Eq. (2.22) can be derived from simulations (see, e.g., Scoccimarro 2004; Taruya et al. 2010), accounting for the nonlinear coupling between the density and velocity fields.

Additional effects distorting $\xi(s_\perp, s_\parallel)$ take place on small scales, where incoherent motions within virialized structures, such as galaxy clusters, have a remarkable impact on redshift measurements. Such motions cause nonlinear effects in the observed clustering, elongating $\xi(s_\perp, s_\parallel)$ along $s_\parallel$ and generating the so-called *fingers-of-God* effect (Jackson 1972), as we show in Fig. 2.2. Besides peculiar velocities, also random errors affecting redshift measurements, namely σ_z, contribute to the fingers-of-God effect. Specifically, the observed redshift, $z_{\rm ob}$, can be expressed as follows:

$$z_{\rm ob} = z + \frac{v_\parallel}{c}(1+z) \pm \sigma_z , \tag{2.23}$$

where z is the cosmological redshift, c is the speed of light, and $v_\parallel$ is the component of the peculiar velocity along the line of sight. To account for both peculiar velocities and redshift statistical uncertainties, Eq. (2.22) can be rewritten as follows:

$$P_{\rm RSD}(k,\mu) = P(k)\left(b + f\mu^2\right)^2 D_{\rm FoG}(k,\mu,f,\sigma_{\rm v,tot}) . \tag{2.24}$$

Eq. (2.24) is the dispersion model (Peebles 1980; Davis & Peebles 1983; Peacock & Dodds 1996), accounting for both linear and nonlinear dynamics. $D_{\rm FoG}$ is a damping function describing the fingers-of-God effect, which is usually expressed as a Gaussian damping,

$$D_{\rm FoG}(k,\mu,f,\sigma_{\rm v,tot}) = e^{-k^2\mu^2 f^2 \sigma^2_{\rm v,tot}} , \tag{2.25}$$

or as a Lorentzian damping,

$$D_{\rm FoG}(k,\mu,f,\sigma_{\rm v,tot}) = \frac{1}{1 + k^2\mu^2 f^2 \sigma^2_{\rm v,tot}} . \tag{2.26}$$

Notably, $D_{\rm FoG}$ implies a damping of $\xi(s)$ at low s (see, e.g., Marulli et al. 2012). We remark that Eq. (2.24) holds in case of Gaussian redshift statistical errors. Moreover, in Eq. (2.24), $\sigma_{\rm v,tot}$ is expressed as follows:

$$\sigma_{\rm v,tot} = \sqrt{\sigma_{\rm v} + \sigma_{{\rm v},z}} , \tag{2.27}$$

where $\sigma_{\rm v}$ is the pairwise velocity dispersion induced by the matter density field, which can be either assumed as a free parameter in cosmological analyses, or analytically derived from linear theory (see, e.g., Taruya et al. 2010). In Eq. (2.27), $\sigma_{{\rm v},z}$ is the apparent velocity dispersion caused by redshift errors, having the following functional form:

$$\sigma_{{\rm v},z} \equiv \frac{c\,\sigma_z(1+\bar{z})}{H(\bar{z})} , \tag{2.28}$$

where $\bar{z}$ is the mean redshift of the sample, c is the speed of light, $H(z)$ is the Hubble function, and σ_z is the typical redshift statistical uncertainty of the sample. As we shall discuss in Chapt. 4, $\sigma_{{\rm v},z}$ can be derived directly from the clustering shape when redshift error estimates are not available.

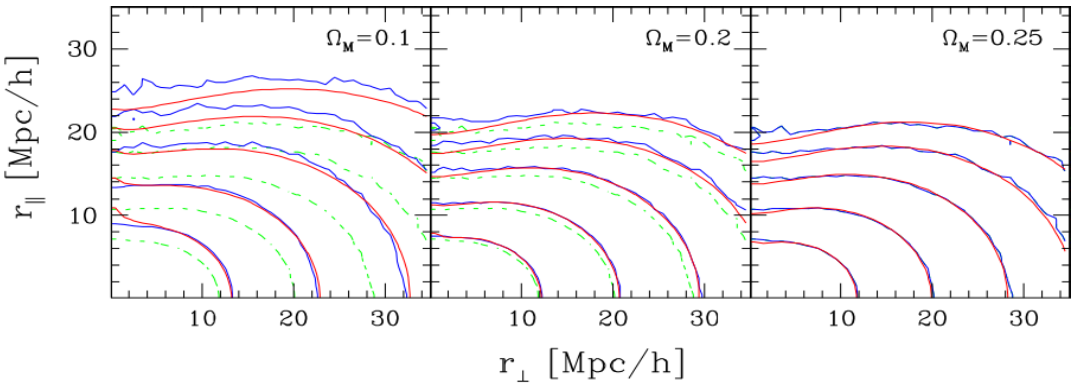

Figure 2.3: Impact of geometric distortions on simulated data. The blue solid curves represent the measured redshift-space iso-correlation contours, derived by assuming different values of the matter density parameter. The red solid lines show the best-fit dispersion model of the 2PCF, while the green dotted lines show the geometrically undistorted 2PCF, measured at the true value of the matter density parameter, namely $\Omega_{\rm m}^{\rm true} = 0.25$. Image by Marulli et al. (2012).

2.5 Geometric distortions

To measure the 3D clustering of a sample as a function of physical separation, a fiducial cosmological model must be assumed. In general, such fiducial model does not coincide with the true one, and this mismatch causes the so-called *geometric distortions*. In turn, by knowing the impact of RSD, one can infer the true cosmological model by leveraging on geometric distortions. This procedure is referred to as the Alcock & Paczynski (1979, AP) test. One possible method to apply the AP test is through the 2PCF in Cartesian coordinates, $\xi(s_\perp, s_\parallel)$. Specifically, the components of the physical separation derived by assuming a test cosmology, namely $s_\parallel^{\rm test}$ and $s_\perp^{\rm test}$, can be expressed as a function of those obtained by assuming the fiducial cosmology, namely $s_\parallel^{\rm fid}$ and $s_\perp^{\rm fid}$, as follows:

$$
\begin{aligned}
s_\perp^{\rm test} &= \frac{D_{\rm A}^{\rm test}(z)}{D_{\rm A}^{\rm fid}(z)}\, s_\perp^{\rm fid}\,, \\
s_\parallel^{\rm test} &= \frac{H^{\rm fid}(z)}{H^{\rm test}(z)}\, s_\parallel^{\rm fid}\,,
\end{aligned}
\tag{2.29}
$$

where $D_{\rm A}(z)$ and $H(z)$ are the angular diameter distance and the Hubble parameter, respectively. The correct values of the cosmological parameters provide $\xi(s_\perp^{\rm fid}, s_\parallel^{\rm fid}) = \xi(s_\perp^{\rm test}, s_\parallel^{\rm test})$, as we show in Fig. 2.3 (right panel). However, we remark that the effect of geometric distortions can be subdominant compared to the impact of redshift errors (Marulli et al. 2012).

2.6 Weak gravitational lensing

From General Relativity, we know that the trajectory of light rays traveling from a source to an observer is influenced by both the gravitational field of local structures and the geometry of the Universe (Schneider et al. 1992). Consequently, images of background sources bear the signature of gravitational lensing caused by the presence of cosmic structures. This allows for investigating the mass distribution of cosmic objects (Blandford & Narayan 1992). In turn, lensing is key for estimating the mass of galaxy clusters (Umetsu 2020), as we shall see in Chapt. 3. In addition, tomographic studies of observed lensing signal, performed on large survey areas and in bins of redshift, probe the redshift evolution of the large scale

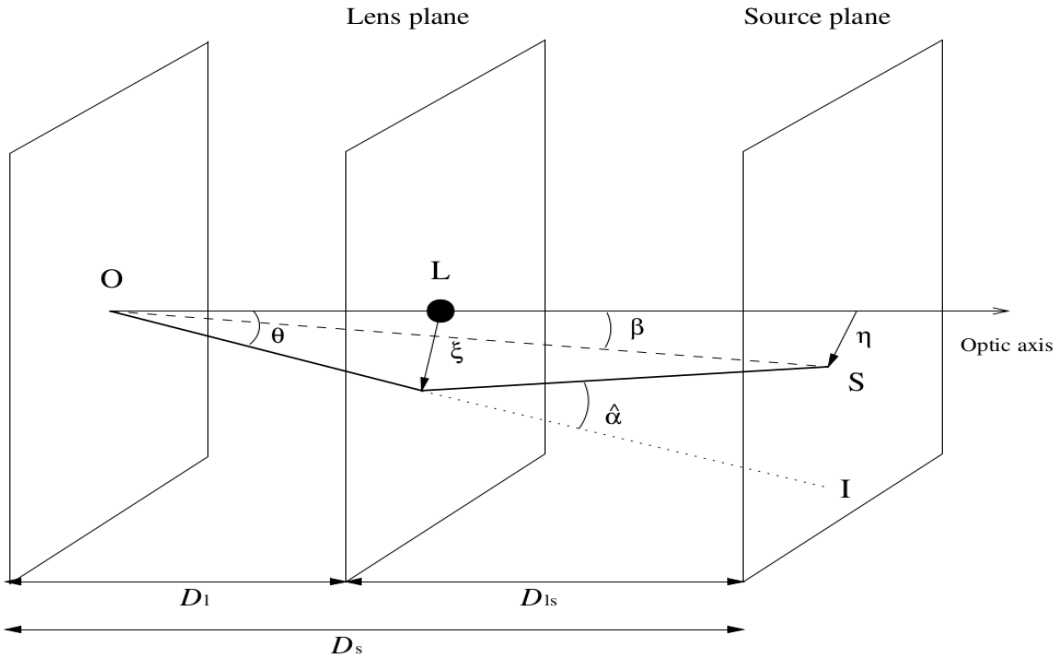

Figure 2.4: Figure by Umetsu (2020), representing the lens system described in Sect. 2.6.1.

structure of the Universe and yield strong constraints on the cosmological parameters. In fact, such tomographic studies, also referred to as *cosmic shear* analyses, are among the primary cosmological probes in current (see, e.g., Giblin et al. 2021; Amon et al. 2022) and future (Ivezic et al. 2008; Laureijs et al. 2011) surveys.

In this section, we present the basic concepts of gravitational lensing, focusing on the so-called weak lensing. Compared to its strong counterpart, which is observable in case of an alignment of lens and lensed source along the line of sight, weak lensing is much more widespread and thus it is best suited for statistical studies.

2.6.1 Lens equation

To introduce the lens equation, describing the bending of light ray paths due to intervening matter along the line of sight, let us consider the system shown in Fig. 2.4. We adopt the *thin lens* approximation, for which the lens matter distribution is fully described by its surface density. This is a good approximation, as the physical size of the lens is generally much smaller than the distances between observer, lens and source.

We define the *optical axis* as the line passing through the observer and the lens locations. The optical axis is perpendicular to the observer, the lens, and the source planes, referred to as O, L, and S, respectively. Specifically, we consider a distant source behind the lens at a position $\eta = \beta D_\mathrm{s}$ on S, where β is the angle subtended by η and D_s is the observer–source angular diameter distance. The lens deflects the path of a light ray emitted by the source by a bending angle $\hat{\alpha}$, so that the observer measures an angle θ between the optical axis and the source position on L, namely $\xi = \theta D_\mathrm{l}$, where D_l is the observer-lens angular diameter distance.

If β, θ, and $\hat{\alpha}$ are small, we obtain the following expression relating the source observed position, I, and its true position:

$$\theta D_\mathrm{s} = \beta D_\mathrm{s} + \hat{\alpha} D_\mathrm{ls} \,, \tag{2.30}$$

where $D_{\rm ls}$ is the lens-source angular diameter distance. Defining the reduced deflection angle, $\alpha = \hat{\alpha} D_{\rm ls}/D_{\rm s}$, we obtain the lens equation, expressed as follows:

$$\beta = \theta + \alpha(\theta)\,. \tag{2.31}$$

We remark that, for a source position β, the lens equation may have multiple solutions $\boldsymbol{\theta}$ (Hattori et al. 1999; Kneib & Natarajan 2011).

2.6.2 Convergence and shear

To relate the light ray path deflection described by the lens equation, Eq. (2.31), to gravity, let us consider a lens in terms of an extended distribution of matter. We express the lens effective gravitational potential, $\hat{\Psi}$, by projecting the three-dimensional Newtonian potential, Φ, on the lens plane, and by conveniently rescaling it:

$$\hat{\Psi}(\boldsymbol{\theta}) := \frac{D_{\rm ls}}{D_{\rm l} D_{\rm s}} \frac{2}{c^2} \int \Phi(\boldsymbol{\theta} D_{\rm l}, z)\, {\rm d}z\,, \tag{2.32}$$

where c is the speed of light and z is the redshift. Eq. (2.32) implies that the gradient of $\hat{\Psi}$ equals the reduced deflection angle, namely

$$\nabla \hat{\Psi}(\boldsymbol{\theta}) = \boldsymbol{\alpha}(\boldsymbol{\theta})\,, \tag{2.33}$$

and that the Laplacian of $\hat{\Psi}$ is equal to twice the *convergence*, κ,

$$\nabla^2 \hat{\Psi}(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta})\,, \tag{2.34}$$

where κ is expressed as follows:

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\rm cr}}\,. \tag{2.35}$$

In Eq. (2.35), Σ is the surface mass density, while $\Sigma_{\rm cr}$ is the critical surface mass density, having the following definition:

$$\Sigma_{\rm cr} := \frac{c^2}{4\pi G} \frac{D_{\rm s}}{D_{\rm l} D_{\rm ls}}\,. \tag{2.36}$$

If the source redshift is much larger than the lens redshift, $\Sigma_{\rm cr}$ weakly depends on the source redshift (Umetsu 2020), making the precise definition of the source redshift distribution less crucial (Okabe et al. 2010a).

From Eq. (2.31), we obtain the following Jacobian matrix, A, mapping the $\boldsymbol{\beta}$ points on the source plane onto $\boldsymbol{\theta}$ points on the lens plane:

$$A = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial \alpha_i(\boldsymbol{\theta})}{\partial \theta_j} \right) = \left(\delta_{ij} - \Psi_{ij} \right)\,, \tag{2.37}$$

where δ_{ij} is the Kronecker delta, while θ_i and θ_j are the ith and jth components of $\boldsymbol{\theta}$, respectively, and Ψ is defined as follows:

$$\Psi_{ij} := \frac{\partial^2 \hat{\Psi}(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\,. \tag{2.38}$$

Now we focus on the antisymmetric part of A, derived as follows:

$$\left(A - \frac{1}{2} {\rm tr} A \cdot I \right)_{ij} = \begin{pmatrix} -\frac{1}{2}(\Psi_{11} - \Psi_{22}) & -\Psi_{12} \\ -\Psi_{12} & \frac{1}{2}(\Psi_{11} - \Psi_{22}) \end{pmatrix}\,, \tag{2.39}$$

where tr$A$ is the trace of A and I is the identity matrix. Eq. (2.39) is an antisymmetric and trace-free matrix, referred to as the *shear matrix*. It provides a measurement of the gravitational tidal field projection, that is the gradient of the gravitational force, which accounts for the distortions observed in background sources. Specifically, the shape distortion of extended sources occurs due to the differential deflection of light bundles. We can define the *shear* pseudovector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$ on the lens plane, where the γ_1 and γ_2 components are expressed as follows:

$$\gamma_1 = \frac{1}{2}(\Psi_{11} - \Psi_{22})\,, \tag{2.40}$$

$$\gamma_2 = \Psi_{12} = \Psi_{21}\,. \tag{2.41}$$

Since the eigenvalues of the shear matrix are $\pm\sqrt{\gamma_1^2 + \gamma_2^2} = \pm\gamma$, it is possible to find a coordinate rotation by an angle ϕ such that the following expression holds:

$$\begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix} = \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}\,. \tag{2.42}$$

The remaining part of A has the following form:

$$\frac{1}{2}\mathrm{tr}A \cdot I = (1-\kappa)\delta_{ij}\,. \tag{2.43}$$

Consequently, contrary to the shear, the convergence induces isotropic transformations of the source shape, rescaling images by a constant factor in all directions. From Eqs. (2.39) and (2.43), the Jacobian matrix is expressed as follows:

$$A = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}\,. \tag{2.44}$$

The eigenvalues are $\lambda_\pm = 1 - \kappa \pm \gamma$, with $\lambda_+ \geq \lambda_-$. From Eq. (2.44), we note that in the regime in which $|\kappa| \ll 1$ and $|\gamma| \ll 1$ the image distortions are small. We refer to this limiting case as the *weak-lensing* regime.

In addition, as a consequence of Liouville's theorem, lensing conserves the surface brightness of background sources. Thus, lensed images undergo magnification, whose effect is quantified by the magnification factor, μ, expressed as follows:

$$\mu = \frac{\delta\Omega^I}{\delta\Omega^S} = \frac{1}{\det A} = \frac{1}{(1-\kappa)^2 - \gamma^2}\,, \tag{2.45}$$

where $\delta\Omega^I$ and $\delta\Omega^S$ are the lensed and unlensed image solid angles, respectively, and $\det A$ is the determinant of A.

2.6.3 Weak-lensing signal from observed ellipticity

As discussed above, weak lensing is more common than its strong counterpart. This is due to a mere statistical reason, as strong lensing requires a good alignment of source and lens along the line of sight (Hattori et al. 1999). In turn, weak lensing is widely exploited in cosmology as it maps the large scale structure of the Universe (Kilbinger 2015). In addition, weak lensing is key for estimating galaxy cluster masses (Umetsu 2020), which is a crucial step for performing cosmological analyses based on the statistical properties of these objects (Allen et al. 2011).

Nevertheless, measuring the small image distortions produced by weak lensing is not

trivial, due to the shape measurement noise, the isotropic smearing component of the point spread function (PSF), and the impact of instrumental PSF anisotropy. Kaiser et al. (1995) first proposed a method to precisely extract the lensing signal from noisy images of source galaxies, which was extended and improved during the last decades (see, e.g., Kuijken 1999; Bridle et al. 2002; Heymans et al. 2006; Refregier et al. 2012; Mandelbaum et al. 2018a). This method is based on the quadrupole moments, Q_{ij} with $i, j = 1, 2$, of the source brightness distribution, $I(\boldsymbol{\theta})$, which are expressed as follows (Kaiser et al. 1995):

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})] \Delta\theta_i \Delta\theta_j}{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})]} \,, \tag{2.46}$$

where $q_I[I(\boldsymbol{\theta})]$ is a weight depending on the brightness (Bartelmann & Schneider 2001), and $\Delta\theta_i = \theta_i - \overline{\theta}_i$ is the offset from the image centroid. We can express the complex ellipticity, $e = e_1 + ie_2$, as follows:

$$e = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22}} \,, \tag{2.47}$$

from which we can derive the intrinsic source ellipticity, e_s, as

$$e_\mathrm{s} = \frac{e - 2g + e^* g^2}{1 + |g|^2 - 2\Re(e^* g)} \,, \tag{2.48}$$

where $g = \gamma/(1-\kappa)$ is the reduced shear, and e^* is the complex conjugate of e. Since e is invariant under π rotations, the e_s expectation value, $\mathbb{E}(e_\mathrm{s})$, is null. Under this condition, Eq. (2.48) becomes (Schneider & Seitz 1995)

$$0 = \sum_n w_n \frac{e_n - \delta_g}{1 - \Re(e_n^* \delta_g)} \,, \tag{2.49}$$

where e_n and w_n are the image ellipticity and the statistical weight of the nth object, respectively, and δ_g is the complex distortion, which is expressed as follows:

$$\delta_g = \frac{2g}{1 + |g|^2} \,. \tag{2.50}$$

In the weak-lensing regime, that is $|\kappa|, |\gamma| \ll 1$, Eq. (2.48) reduces to $e_\mathrm{s} \simeq e - 2g$. Therefore, assuming randomly oriented sources, the observed reduced shear in the weak-lensing regime can be expressed as

$$\langle g \rangle \simeq \frac{\langle e \rangle}{2} \,, \tag{2.51}$$

where $\langle e \rangle$ is the average of observed ellipticities of a source ensemble. The statistical uncertainty on $\langle g \rangle$ decreases as $\propto \sigma/\sqrt{N}$, where N is the number of sources and σ is the source ellipticity dispersion, dominated by the intrinsic shape noise. To account for biases caused by observational and instrumental effects, the observed reduced shear, g^ob, is usually expressed as a function of the true reduced shear, g^tr (Heymans et al. 2006; Massey et al. 2007; Mandelbaum et al. 2014):

$$g_i^\mathrm{ob} = (1 + m_i)\, g_i^\mathrm{tr} + c_i \quad (i = 1, 2) \,, \tag{2.52}$$

where m_i and c_i estimates are based on observed and simulated galaxy samples.

2.6.4 Tangential and cross shear

From Eq. (2.51), we know that in the weak-lensing regime the shear can be estimated from observed ellipticities. The shear components γ_1 and γ_2, introduced in Sect. 2.6.2, depend on the chosen coordinate frame. To define shear components that are coordinate independent with respect to a given reference point, we define the polar coordinate frame (ϑ, φ) centred on $\boldsymbol{\theta}_\mathrm{c}$, such that $\boldsymbol{\theta} = (\vartheta\cos\varphi, \vartheta\sin\varphi) + \boldsymbol{\theta}_\mathrm{c}$. The convergence averaged within a circle centred on $\boldsymbol{\theta}_\mathrm{c}$ and with radius ϑ, namely $\overline{\kappa}(\vartheta)$, is expressed as:

$$\overline{\kappa}(\vartheta) = \frac{2}{\vartheta^2}\int_0^\vartheta \mathrm{d}\vartheta'\,\vartheta'\kappa(\vartheta') = \frac{\overline{\Sigma}(\vartheta)}{\Sigma_\mathrm{cr}}\,, \tag{2.53}$$

$$\kappa(\vartheta) = \oint \frac{\mathrm{d}\varphi}{2\pi}\,\kappa(\vartheta,\varphi) = \frac{\Sigma(\vartheta)}{\Sigma_\mathrm{cr}}\,, \tag{2.54}$$

where $\Sigma(\vartheta)$ is the surface mass density averaged along a circle with radius ϑ, $\overline{\Sigma}(\vartheta)$ is the average surface mass density within a circle with radius ϑ, and Σ_cr is the surface critical density, defined in Eq. (2.36).

Then, we introduce the tangential shear, γ_+, and the cross shear component given by a 45° rotation, $\gamma_\times$, having the following expressions:

$$\gamma_+(\boldsymbol{\theta}) = -\gamma_1(\boldsymbol{\theta})\cos(2\varphi) - \gamma_2(\boldsymbol{\theta})\sin(2\varphi)\,, \tag{2.55}$$

$$\gamma_\times(\boldsymbol{\theta}) = +\gamma_1(\boldsymbol{\theta})\sin(2\varphi) - \gamma_2(\boldsymbol{\theta})\cos(2\varphi)\,. \tag{2.56}$$

From these equations, we find (Kaiser 1995):

$$\gamma_+(\vartheta) = \oint \frac{\mathrm{d}\varphi}{2\pi}\,\gamma_+(\vartheta,\varphi) = \overline{\kappa}(\vartheta) - \kappa(\vartheta) = \frac{\Delta\Sigma(\vartheta)}{\Sigma_\mathrm{cr}}\,, \tag{2.57}$$

$$\gamma_\times(\vartheta) = \oint \frac{\mathrm{d}\varphi}{2\pi}\,\gamma_\times(\vartheta,\varphi) = 0\,, \tag{2.58}$$

where the excess surface mass density, $\Delta\Sigma(\vartheta)$, is expressed as follows (Miralda-Escude 1991):

$$\Delta\Sigma(\vartheta) = \overline{\Sigma}(\vartheta) - \Sigma(\vartheta)\,. \tag{2.59}$$

From Eq. (2.58), we note that $\gamma_\times$ provides a useful null test for weak-lensing measurements, as it is expected to be consistent with zero. We can derive the reduced tangential shear averaged over a circle with radius ϑ and centre $\boldsymbol{\theta}_\mathrm{c}$, $g_+(\vartheta)$, as follows:

$$g_+(\vartheta) = \oint \frac{\mathrm{d}\varphi}{2\pi}\,g_+(\vartheta,\varphi) = \oint \frac{\mathrm{d}\varphi}{2\pi}\,\frac{\gamma_+(\vartheta,\varphi)}{1-\kappa(\vartheta,\varphi)}\,. \tag{2.60}$$

If the projected mass distribution has an approximately circular symmetry, then the reduced tangential shear can be expressed as follows:

$$g_+(\vartheta) \simeq \frac{\gamma_+(\vartheta)}{1-\kappa(\vartheta)}\,. \tag{2.61}$$

As we shall discuss in Chapt. 3, this formalism is key to describe the weak-lensing signal produced by galaxy clusters, for which $\boldsymbol{\theta}_\mathrm{c}$ is assumed to be the cluster centre.

Chapter 3

Galaxy cluster cosmology

According to the standard ΛCDM cosmological model, described in Sect. 1.7, galaxy clusters are the largest gravitationally bound systems in the present Universe, formed through a hierarchical assembly of smaller structures. Lying in correspondence of the nodes of the cosmic web (Kaiser 1984), galaxy clusters trace the deepest virialised potential wells of the matter distribution in the Universe, reaching masses of 10^{15} $M_{\odot}$ and radii of 5 Mpc. The critical role of galaxy clusters in both astrophysics and cosmology dates back to the work by Zwicky (1933), based on observations of the Coma Cluster. Indeed, the Coma Cluster mass derived from the velocity dispersion of member galaxies was observed to exceed the combined mass of all member galaxies by factors of approximately 200 to 400. This discrepancy necessitated the proposition of the existence of a substantial amount of dark matter. Additional support for the existence of dark matter came from the observation of an extended hot intracluster medium (ICM) emitting X-rays through thermal bremsstrahlung (Giacconi et al. 1974). Measurements of the gas temperature further indicated that the depth of potential wells in clusters necessitates the presence of a significant dark matter component. As such, galaxy clusters allow for the investigation of dark matter interaction models (Peter et al. 2013; Robertson et al. 2017; Eckert et al. 2022) and gas astrophysics (Vazza et al. 2017; CHEX-MATE Collaboration 2021; Zhu et al. 2021).

In addition, galaxy clusters play a crucial role in cosmology because they mark the transition between the linear and the nonlinear regime of the gravitational perturbations, so that their formation and evolution can be theoretically described with excellent accuracy (Borgani 2008). Since the dark matter component is dominant in galaxy clusters, it is accurate enough, given the current observational uncertainties, to model the cluster formation and evolution through N-body dark matter simulations (Borgani & Kravtsov 2011; Angulo et al. 2012; Giocoli et al. 2012). Such simulations allow us to calibrate the mass function of galaxy cluster haloes (e.g. Sheth & Tormen 1999; Tinker et al. 2008; Despali et al. 2016), as well as finding a functional form describing the halo bias (e.g. Sheth et al. 2001; Tinker et al. 2010) and the cluster dark matter profiles (Navarro et al. 1997; Baltz et al. 2009). Many attempts have also been made to investigate the impact of the baryonic physical processes on cluster statistics (e.g. Cui et al. 2012; Velliscig et al. 2014; Bocquet et al. 2016; Castro et al. 2021). Through calibrated cluster halo mass function and bias, it is possible to model the observed cluster counts and clustering in order to derive groundbreaking constraints on fundamental cosmological parameters (Sereno et al. 2015; Veropalumbo et al. 2016; Costanzi et al. 2019; Marulli et al. 2021; Moresco et al. 2021; To et al. 2021a; Euclid Collaboration: Fumagalli et al. 2022; Lesci et al. 2022a,b, 2023; Fumagalli et al. 2023). In turn, it is necessary to derive the cluster masses in order to perform such cosmological analyses (Allen et al. 2011), as we shall discuss in Sect. 3.2. Furthermore, promising techniques to constrain the cosmological

parameters concern the study of the weak-lensing peak counts in cosmic shear maps (e.g. Maturi et al. 2011; Reischke et al. 2016; Shan et al. 2018; Martinet et al. 2018; Giocoli et al. 2018) and galaxy cluster sparsity (e.g. Balmès et al. 2014; Corasaniti et al. 2021).

In this Chapter, we detail the detection techniques leading to the collection of large samples of galaxy clusters, which allow for carrying out cosmological analyses. In addition, we describe the mass calibration processes exploited in current cosmological surveys, which yield scaling relations between cluster masses and observable mass proxies. As we shall detail, based on such mass scaling relations, it is possible to extract cosmological information from summary statistics of cluster distributions, such as cluster counts and clustering.

3.1 Cluster detection and mass proxies

Large samples of galaxy clusters can be built by exploiting different techniques thanks to their multi-wavelength emission. In particular, galaxy clusters can be detected through the bremsstrahlung emission of the intracluster medium in the X-ray band (e.g. Böhringer et al. 2004; Clerc et al. 2014; Pierre et al. 2016), through the detection of the Sunyaev-Zeldovich effect in the CMB (e.g. Hilton et al. 2018), or through their gravitational lensing effect on the background galaxies (e.g. Maturi et al. 2005; Miyazaki et al. 2018b). Furthermore, galaxy clusters can be detected in the optical band by looking for overdensities and peculiar features characterising cluster members in galaxy surveys (e.g. Rykoff et al. 2014; Bellagamba et al. 2018).

Ongoing and forthcoming wide extragalactic surveys, from the lowest to the highest frequencies, will provide complete and pure galaxy cluster samples up to high redshifts and down to low masses. These surveys include the Kilo Degree Survey[1] (KiDS, de Jong et al. 2017; Kuijken et al. 2019), the Dark Energy Survey[2] (The Dark Energy Survey Collaboration 2005; Abbott et al. 2021), the Vera C. Rubin Observatory LSST[3] (LSST Dark Energy Science Collaboration 2012, 2021), and *Euclid*[4] (Amendola et al. 2018; Euclid Collaboration et al. 2019, 2020; Euclid Collaboration: Scaramella et al. 2022) in the optical and near-infrared, the South Pole Telescope[5] (Bayliss et al. 2016; Chown et al. 2018), the Atacama Cosmology Telescope[6] (Naess et al. 2020; Orlowski-Scherer et al. 2021), and the Simons Observatory[7] (Ade et al. 2019; Xu et al. 2021) surveys at high-radio frequencies, and eROSITA[8] (Brunner et al. 2022; Liu et al. 2022) in X-rays.

In this section, we detail the most commonly employed cluster detection techniques, providing the description of the mass proxies that are usually related to cluster masses. In fact, deriving the relation between observable mass proxies and cluster masses is key for cosmology, as masses, along with redshifts, are fundamental cosmological variables on which theoretical models are based (Pratt et al. 2019; Clerc & Finoguenov 2023). In the following, we also provide a comparison between cluster detection methods, outlining the differences in the cluster sample selections.

[1] http://kids.strw.leidenuniv.nl/
[2] https://www.darkenergysurvey.org
[3] https://www.lsst.org/
[4] http://sci.esa.int/euclid/
[5] https://pole.uchicago.edu/
[6] https://act.princeton.edu/
[7] https://simonsobservcatory.org/
[8] http://www.mpe.mpg.de/eROSITA

3.1.1 Optical and near-infrared

Historically, galaxy clusters were identified as galaxy overdensities in optical images, and Abell (1958) provided the first catalogue based on this method. This technique, however, can be affected by projection effects, for which field galaxies along the line of sight could be mistakenly identified as cluster members. One way to address this issue is to leverage the photometric characteristics exhibited by galaxies within a cluster. Galaxy clusters contain a well-defined, highly regular population of elliptical and lenticular galaxies, which are observed in optical and near-infrared (NIR) bands. In a colour-magnitude diagram, these member galaxies form a distinct sequence known as the *red-sequence* (Bower et al. 1992). Notably, the red-sequence shows a tight linear relationship between colour and magnitude, indicating that redder galaxies are brighter. Consequently, the red-sequence plays a significant role in identifying galaxy clusters. An example of an algorithm utilizing this feature is the red-sequence Matched-filter Probabilistic Percolation (redMaPPer; Rykoff et al. 2014, 2016) algorithm, which has been widely employed for constructing cluster catalogues from the Sloan Digital Sky Survey (SDSS; Gunn et al. 1998) and DES.

As forthcoming optical and NIR surveys will provide deeper observations of the sky, cluster detection based on the red-sequence may become not optimal. Indeed, this feature could not be prominent enough for high redshift clusters. Among the detection algorithms which do not rely on the red-sequence, the Adaptive Matched Identifier of Clustered Objects (AMICO; Bellagamba et al. 2018) algorithm is one of the most employed in current surveys, such as KiDS and *Euclid*. AMICO is based on the optimal matched-filtering technique (Maturi et al. 2005), based on the assumption of a model for the density and luminosity profiles of galaxy clusters. Then, AMICO identifies clusters as galaxy overdensities by matching them with the defined model, estimating membership probabilities which strongly depend on galaxy photometric redshifts. The AMICO algorithm also provides reliable estimates of the cluster sample selection function, based on mock galaxy catalogues (Maturi et al. 2019). As we shall discuss in Chapts. 5 and 7, the application of AMICO to KiDS data provided highly pure and complete cluster samples down to low masses and high redshift.

As cluster detection in optical and NIR surveys relies on the observed galaxy distribution, one of the main detection observables is the richness. It quantifies the number of cluster member galaxies which satisfy conditions on, for example, magnitude or distance from the detection centre (Rykoff et al. 2016; Maturi et al. 2019). In fact, cluster richness is key for cosmology, as it has observationally proven to correlate with mass (Bellagamba et al. 2019; Costanzi et al. 2019; Abbott et al. 2020). As we shall discuss in Sect. 3.2, this allows for building up scaling relations between richness and mass. Furthermore, matched-filter detection algorithms such as AMICO provide an additional mass proxy, commonly referred to as signal amplitude (Bellagamba et al. 2018).

3.1.2 X-rays

X-ray observations have a critical role in the construction of large galaxy cluster samples (Arnaud 2005; Ettori & Molendi 2011) and, in turn, in cosmological studies (Marulli et al. 2018; Pierre 2022; Clerc & Finoguenov 2023). Indeed, galaxy clusters appear as extended X-ray sources in the sky, as the deep potential well of the dark matter halo leads to virial temperatures of approximately $10^7 - 10^8$ K for the diffuse ICM. As hydrogen is the most abundant element in the Universe, the ICM is mostly composed of free electrons and protons, emitting X-rays through free-free interactions, referred to as bremsstrahlung. The other primary X-ray emission processes are electron recombination and line emissions. In each of these cases, the emissivity is directly proportional to the square of the electron density, which

ranges from approximately 10^{-1} cm^{-3} in the inner regions of bright cool core clusters to approximately 10^{-5} cm^{-3} in the outskirts. This implies an optically thin regime, and thus a non-auto-absorbed emission.

As the ICM properties depend on the halo potential well depth, the cluster total mass is expected to strongly correlate with X-ray observable mass proxies, such as gas temperature, gas mass, and X-ray luminosity. The temperature can be estimated from spectroscopic data, which provide information on light emission of heavy elements, or from the cut-off of the continuum ICM emission. The luminosity is proportional to the normalisation of the observed profile, while the gas mass can be derived through the emission measure, corresponding to the integral of the electron density along the line of sight (Arnaud 2005). As these mass proxies trace the ICM virialisation, they characterise the inner regions of galaxy clusters. Specifically, the X-ray mass proxies are usually measured within R_{500c}, namely the radius within which the mean mass density is 500 times the critical density of the Universe at the cluster redshift. As shown by Eckert et al. (2020), X-ray emission is a low-scatter mass proxy at $\sim 0.5R_{500c}$.

Similarly to some optical and NIR detection algorithms, X-ray detection can rely on a model describing the expected cluster profiles (Käfer et al. 2019; Xu et al. 2022). Among the simplest models adopted in literature there is the isothermal β model (Cavaliere & Fusco-Femiano 1976), based on the assumption of hydrostatic equilibrium and spherical symmetry (Voit 2005), having the following functional form:

$$\rho_{\rm gas}(r) \propto [1 + (r/r_{\rm c})^2]^{-3\beta/2}, \tag{3.1}$$

where $\rho_{\rm gas}(r)$ is the gas density at a distance r from the centre, $r_{\rm c}$ is a core radius, and $\beta \simeq 2/3$ (Jones & Forman 1984). From Eq. (3.1), it is possible to derive an analytic surface brightness profile (Gorenstein et al. 1978), and X-ray cluster candidates are expected to follow such theoretical profile. In addition, Eq. (3.1) can be exploited to estimate the cluster gas mass from observations, as an alternative to the emission measure (Arnaud 2005).

Despite massive clusters can be easily identified as X-ray sources, optical and NIR observations allow for building up samples down to lower masses, including groups of galaxies (see, e.g., Maturi et al. 2019; Bellagamba et al. 2019). In addition, X-ray cluster surveys need complementary observations in optical and NIR to measure the redshifts (Klein et al. 2018), if resolved emission lines are not available in X-ray spectra (Yu et al. 2011; Borm et al. 2014). On the other hand, X-ray observations are less affected by projection effects, which may significantly bias the mass proxy and, in turn, the cluster mass estimates (Myles et al. 2021; Wu et al. 2022; Shi et al. 2023b). Furthermore, the cluster sample completeness can be analytically derived for flux-limited X-ray surveys (Tozzi 2007; Borgani 2008), as the ICM emission is continuous and stable. Conversely, estimating the completeness of optical and NIR selected cluster samples requires simulations (Maturi et al. 2019), as the cluster flux depends on the adopted photometric filters, the number and the concentration of member galaxies.

3.1.3 Thermal Sunyaev-Zeldovich effect

When CMB photons pass through the hot ICM of a galaxy cluster, they experience a frequency shift due to inverse Compton scattering with free electrons, commonly known as the Sunyaev-Zeldovich effect (SZ; Sunyaev & Zeldovich 1972). When the interaction is due to the non-relativistic thermal motion of the electrons, this effect is called thermal SZ (tSZ, referred to as SZ hereafter for simplicity). As a result of the SZ effect, low-frequency photons gain energy, leading to a distortion in the black body spectrum of the CMB. This distortion

causes a decrease in the number of photons with frequencies below $\nu_0 \simeq 217$ GHz and a corresponding increase in the number with frequencies above $\nu_0 \simeq 217$ GHz. Notably, ν_0 serves as a characteristic frequency at which the spectrum remains undistorted.

The SZ effect is purely a spectral distortion which does not depend on redshift. The amplitude of this distortion is quantified by the Compton parameter, $Y_{\rm SZ}$, expressed as follows:

$$Y_{\rm SZ} = \frac{\sigma_T}{m_e c^2} \int P_e \, {\rm d}l \, , \tag{3.2}$$

where σ_T represents the Thomson cross-section, m_e denotes the electron mass, c stands for the speed of light, and P_e represents the thermal pressure of free electrons integrated over the distance traveled by photons in the ICM. In the low-frequency approximation, the SZ effect causes variations in the intensity of the signal, I, corresponding to

$$\frac{\delta I}{I} = \frac{\delta T}{T} = -2 Y_{\rm SZ} \, . \tag{3.3}$$

The typical variation in the signal is $\delta I/I \sim 10^{-4}$, indicating that this effect is not negligible since the CMB itself exhibits variations on the order of $\sim 10^{-5}$. As shown in Eq. 3.3, unlike X-ray and optical measurements, the SZ signal of a cluster does not experience surface brightness dimming. Consequently, SZ surveys are well suited for detecting massive clusters at high redshifts.

Matched filter detection algorithms are commonly adopted for cluster identification through SZ, relying on pressure profile models, such as the isothermal β-model, to describe the SZ signal (Melin et al. 2006). Such algorithms aim at maximising the SZ signal-to-noise (S/N) by filtering the CMB map, and provide an estimate of $Y_{\rm SZ}$. We remark that $Y_{\rm SZ}$ quantifies the amplitude of the signal, and in turn it is a robust cluster mass proxy. As we shall discuss in Sect. 3.2, SZ and X-ray observations can be combined to derive robust constraints on cluster masses, based on mass proxies.

Similar to X-ray observations, the cluster SZ emission depends on ICM properties, and the mass range of detected clusters is compatible with that derived from X-ray surveys. Among the similarities with X-rays, we also mention that the SZ signal is poorly affected by projection effects, and that the sample completeness can be analytically derived (Planck Collaboration XX 2014). In addition, SZ selected clusters need complementary optical and NIR observations for measuring redshifts.

Throughout this section, we referred to the thermal SZ effect simply as SZ. This must be distinguished from the kinetic SZ effect (kSZ; Carlstrom et al. 2002), caused by the Doppler effect of clusters having a non-zero peculiar velocity along the line of sight, implying a proper motion with respect to the Hubble flow. Unlike its thermal counterpart, kSZ does not produce distortions in the shape of the spectrum, but the bulk temperature of the CMB spectrum does change. As a result, kSZ serves as a reliable tracer of the peculiar velocity field, despite inducing temperature variations of $|\partial T/T| \sim 10^{-5}$, which is an order of magnitude lower than those produced by SZ.

3.2 Mass measurements in cosmological surveys

Galaxy cluster formation and evolution can be theoretically modelled with high accuracy, this allowing the derivation of reliable cluster mass estimates from observations. This is not feasible for galaxies, due to the large uncertainties in galaxy formation and evolution models. Consequently, compared to cosmological analyses based on galaxies, cluster mass estimates compensate the lower abundance of large scale structure tracers.

In this section, our focus is on two highly precise methods for measuring cluster masses in cosmology, specifically those based on weak-lensing and X-ray hydrostatic equilibrium measurements. Other methods, based on strong lensing (Monna et al. 2017; Remolina González et al. 2021; Fox et al. 2022) and on the dynamics of member galaxies (Kodi Ramanah et al. 2020; Ho et al. 2022; Biviano & Mamon 2023), require long exposure time and spectroscopy, which are hard to achieve in large galaxy cluster surveys.

3.2.1 Weak lensing

As detailed in Sect. 2.6, observed galaxy ellipticities can be related to the reduced shear, $g = \gamma/(1-\kappa)$, in the weak-lensing regime. In turn, it is possible to reconstruct the gravitational potential of the lens from observations. Optical and NIR observations are suitable for achieving this task, as photometry yields accurate estimates of galaxy shapes (Ivezic et al. 2008; Laureijs et al. 2011; Giblin et al. 2021; Amon et al. 2022).

In Sect. 2.6.4, we showed that the excess surface mass density averaged over a circle, $\Delta\Sigma_+$, is related to the tangential component of the shear, γ_+, as

$$\Delta\Sigma_+(R) = \overline{\Sigma}(< R) - \Sigma(R) = \Sigma_{\rm cr}\gamma_+(R)\,, \tag{3.4}$$

where $\Sigma_{\rm cr}$ is the surface critical density, expressed in Eq. (2.36), $\Sigma(R)$ is the mass surface density, and $\overline{\Sigma}(< R)$ is its mean within the radius R, expressed as follows

$$\overline{\Sigma}(< R) = \frac{2}{R^2}\int_0^R {\rm d}r\, r\Sigma(r)\,. \tag{3.5}$$

In turn, galaxy ellipticities are proxies of $\Delta\Sigma_+$, as $g_+ \simeq \gamma_+$ in the weak-lensing limit (see Eq. 2.61). Following the formalism adopted in Sect. 2.6.4, we can derive the tangential and cross ellipticity components, namely e_+ and $e_\times$, respectively, as follows (Viola et al. 2015):

$$\begin{pmatrix} e_+ \\ e_\times \end{pmatrix} = \begin{pmatrix} -\cos(2\varphi) & -\sin(2\varphi) \\ \sin(2\varphi) & -\cos(2\varphi) \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}\,, \tag{3.6}$$

where φ is the position angle of the source with respect to the lens centre, while e_1 and e_2 are the ellipticity components introduced in Sect. 2.6.3. Therefore, the $\Delta\Sigma_+$ estimator has the following expression (Sheldon et al. 2004; Viola et al. 2015; Bellagamba et al. 2019):

$$\Delta\Sigma_+(R_j) = \left(\frac{\sum_{i\in j}\left(w_i\,\Sigma_{{\rm cr},i}^{-2}\right)e_{+,i}\,\Sigma_{{\rm cr},i}}{\sum_{i\in j}\left(w_i\,\Sigma_{{\rm cr},i}^{-2}\right)}\right)\frac{1}{1+K(R_j)}, \tag{3.7}$$

where j is the radial annulus index, with mean radius R_j, and w_i is the statistical weight assigned to the source ellipticity measure of the ith galaxy (Sheldon et al. 2004). An optimal choice for the statistical weight is $w = 1/\sigma_g^2$, where σ_g is the statistical uncertainty on the reduced shear. In addition, $K(R_j)$ is the average correction due to the multiplicative noise bias in the shear estimate and is computed as follows

$$K(R_j) = \frac{\sum_{i\in j}\left(w_i\,\Sigma_{{\rm cr},i}^{-2}\right)m_i}{\sum_{i\in j}\left(w_i\,\Sigma_{{\rm cr},i}^{-2}\right)}, \tag{3.8}$$

where m_i is the multiplicative shear bias for the ith galaxy (see Eq. 2.52).

In large photometric surveys, the weak-lensing signal of individual clusters is too low to precisely measure the density profiles. For this reason, $\Delta\Sigma_+$ is usually derived for ensembles of clusters, selected according to their mass proxies and their redshifts. This is possible by

stacking the lensing signal of galaxy clusters through a weighted sum. Specifically, the excess surface density profile for the Kth cluster bin is expressed as follows

$$\Delta\Sigma_{+,K}(R_j) = \frac{\sum_{k\in K} W_{k,j}\,\Delta\Sigma_{+,k}(R_j)}{\sum_{k\in K} W_{k,j}}, \tag{3.9}$$

where k runs over the clusters in the bin, and $W_{k,j}$ is the total weight for the jth radial bin of the kth cluster, estimated as

$$W_{k,j} = \sum_i w_i\,\Sigma_{\mathrm{cr},i}^{-2}, \tag{3.10}$$

where i runs over the background galaxies in the jth radial bin. We remark that Eq. (3.7) is valid for approximately spherical lenses. In case of merging and interacting systems, the lens mass distribution is complex and measurements of the two-dimensional shear field are needed (Umetsu 2020). Nevertheless, for large ensembles of galaxy clusters, the average mass distribution around their centres tends to be spherically symmetric. This is strictly true if selection biases towards the matter distribution are negligible. In addition, we stress that $\Delta\Sigma_+(R)$ measurements require the assumption of a cosmological model for the computation of Σ_{cr} and radial distances from cluster centres. For cosmological analyses, the cluster masses derived by modelling $\Delta\Sigma_+(R)$ can be appropriately rescaled as a function of cosmological parameters (Sereno 2015). Alternatively, reduced shear estimates as a function of the angular distance from the cluster centre do not depend on cosmology (Umetsu 2020).

Having an estimator for $\Delta\Sigma_+$, cluster masses can be derived by assuming a theoretical density profile model. As the dark matter component is dominant in galaxy clusters, N-body simulations provide an excellent theoretical description of the matter distribution of such objects. The Navarro-Frenk-White (NFW; Navarro et al. 1997) density profile is extensively adopted in literature, and it is expressed as follows:

$$\rho_{\mathrm{NFW}}(r) = \frac{\rho_{\mathrm{s}}}{(r/r_{\mathrm{s}})\,(1+r/r_{\mathrm{s}})^2}\,, \tag{3.11}$$

where $\rho_{\mathrm{NFW}}(r)$ is a 3D density, r is the radius, r_{s} is the scale radius, and ρ_{s} is the characteristic density. The scale radius is expressed as $r_{\mathrm{s}} = r_\Delta/c_\Delta$, where r_Δ is the radius enclosing a mass such that the corresponding mean density is Δ times the critical density of the Universe, and c_Δ is the concentration. Specifically, assuming spherical symmetry, r_Δ is expressed as follows:

$$r_\Delta = \left(\frac{3M_\Delta}{4\pi\Delta\rho_{\mathrm{c}}}\right)^{1/3}, \tag{3.12}$$

where ρ_{c} is the critical density of the Universe, and M_Δ is the cluster mass. Similarly to r_{s}, also ρ_{s} depends on c_Δ and M_Δ (see, e.g., Aguena et al. 2021).

Oguri & Hamana (2011) showed that the Baltz-Marshall-Oguri model (BMO; Baltz et al. 2009) better describes cluster profiles outside the viral radius in simulations, compared to the original NFW profile. In particular, the BMO is a truncated version of the NFW profile, and it is expressed as follows:

$$\rho_{\mathrm{BMO}}(r) = \frac{\rho_{\mathrm{s}}}{(r/r_{\mathrm{s}})\,(1+r/r_{\mathrm{s}})^2}\left(\frac{r_{\mathrm{t}}^2}{r^2+r_{\mathrm{t}}^2}\right)^2, \tag{3.13}$$

where $r_{\mathrm{t}} = F_{\mathrm{t}} r_\Delta$ is the truncation radius, with $F_{\mathrm{t}} \simeq 3$ (Oguri & Hamana 2011; Sereno et al. 2017). Another popular truncated density profile is the one proposed by Diemer & Kravtsov (DK14; 2014). This is a composite model, based on the Einasto (1965) profile describing the inner region of the dark matter halo, and a parametric infall term representing

the external accreting matter. In turn, the DK14 profile is extensively used to estimate the splashback radius from data (see Rana et al. 2023, and references therein), that is the radius where accreted particles reach the apocenter of their first orbit. Nevertheless, NFW and BMO profiles are commonly used for mass calibration, mainly for the lower number of model parameters compared to DK14. In addition, to constrain the DK14 infall term parameters, weak-lensing measuments up to large radii are required. However, cluster mass calibration is usually performed within a few Mpc from cluster centres, as the weak-lensing signal at large radii does not improve mass constraints. In addition, such signal may be significantly biased for photometrically selected clusters (Sunayama et al. 2020; Park et al. 2023).

To model $\Delta\Sigma_+$ measurements, the 3D dark matter density, ρ, must be appropriately projected into two dimensions. The mass surface density in Eq. (3.4) can be written as follows

$$\Sigma(R) = \int_{-\infty}^{\infty} \mathrm{d}\chi\, \Delta\rho(\sqrt{R^2+\chi^2})\,, \tag{3.14}$$

where $\Delta\rho$ is the excess matter density, expressed as

$$\Delta\rho(r) = \rho(r) - \rho_\mathrm{m}, \tag{3.15}$$

where $\rho_\mathrm{m} = \Omega_\mathrm{m}\rho_\mathrm{cr}(1+z)^3$. An additional contribution to the surface density, due to the correlated matter nearby the galaxy cluster halo, can be described by the 2-halo term, having the following expression

$$\Sigma_\mathrm{2h}(\theta, M, z) = \frac{\rho_\mathrm{m}(z)\, b_\mathrm{h}(M,z)}{(1+z)^3\, D_\mathrm{l}^2(z)} \int \frac{\mathrm{d}l\, l}{2\pi} J_0(l\theta) P(k_\mathrm{l}, z)\ , \tag{3.16}$$

where θ is the angular radius, J_0 is the 0th order Bessel function, $k_\mathrm{l} = l/(1+z)/D_\mathrm{l}(z)$, D_l is the lens angular diameter distance, b_h is the halo bias, and $P(k_\mathrm{l}, z)$ is the matter linear power spectrum. Eq. (3.16) is an alternative to the DK14 infall term, through which cosmological parameters and halo bias can be constrained (Giocoli et al. 2021; Ingoglia et al. 2022).

Since the weak-lensing measurements are based on the choice of cluster centres, estimating the offcentring scale is crucial. In fact, cluster finders may provide biased centre estimates due to partially masked detections or large photometric noise. In addition, the member galaxy distribution centre may not coincide with the dark matter halo centre, especially in not relaxed clusters. A possible way to address this problem is to associate the cluster centre to the position of the brightest central galaxy (BCG, Dalal et al. 2021; Herbonnet et al. 2022; Shi et al. 2023a). Moreover, the offset between baryonic distribution and dark matter halo centres can be estimated by studying the cluster SZ and X-ray signal in simulations (Saro et al. 2015; Seppi et al. 2023; Sommer et al. 2023). To this aim, the offcentring is statistically estimated for ensembles of clusters, and the excess surface density is expressed as follows

$$\Delta\Sigma_+(R) = (1 - f_\mathrm{off})\, \Delta\Sigma_{+,\mathrm{cen}}(R) + f_\mathrm{off}\, \Delta\Sigma_{+,\mathrm{off}}(R), \tag{3.17}$$

where $\Delta\Sigma_{+,\mathrm{cen}}$ is the centred excess surface density, $\Delta\Sigma_{+,\mathrm{off}}$ is its offcentred counterpart, and f_off is the fraction of offcentred clusters. In particular, the probability of a lens being at distance R_s from the assumed centre, namely $P(R_\mathrm{s})$, is usually described by a Rayleigh distribution (Johnston et al. 2007; Viola et al. 2015; Bellagamba et al. 2019)

$$P(R_\mathrm{s}) = \frac{R_\mathrm{s}}{\sigma_\mathrm{off}^2} \exp\left[-\frac{1}{2}\left(\frac{R_\mathrm{s}}{\sigma_\mathrm{off}}\right)^2 \right], \tag{3.18}$$

where σ_off is the root mean square deviation (rms) of the halo misplacement distribution on the plane of the sky. The azimuthally averaged profile of a population of haloes misplaced

by a distance R_s is given by (Yang et al. 2006)

$$\Sigma(R|R_\mathrm{s}) = \frac{1}{2\pi}\int_0^{2\pi} \Sigma_\mathrm{cen}\left(\sqrt{R^2+R_\mathrm{s}^2+2RR_\mathrm{s}\cos\theta}\right)\mathrm{d}\theta, \tag{3.19}$$

thus the surface mass density distribution derived from an off-centred population of haloes has the following expression

$$\Sigma_\mathrm{off}(R) = \int P(R_\mathrm{s})\,\Sigma(R|R_\mathrm{s})\,\mathrm{d}R_\mathrm{s}. \tag{3.20}$$

From Eqs. (3.4), (3.5), and (3.20), $\Delta\Sigma_{+,\mathrm{off}}$ in Eq. (3.17) can be derived.

We remark that weak lensing traces both dark and baryonic matter components. The baryonic impact on galaxy cluster profiles can be encapsulated within the halo concentration, c_Δ, (Shirasaki et al. 2018). Constraining c_Δ from data is ideal, as the assumption of theoretical c_Δ-M_Δ relations may lead to biases in M_Δ (Henson et al. 2017; Lee et al. 2018). In Chapt. 7, we will show that c_Δ-M_Δ and richness-M_Δ relations can be simultaneously constrained from KiDS weak-lensing data.

Another cluster mass measurement technique relying on weak lensing is the so-called CMB lensing (Zaldarriaga & Seljak 1999; Lewis & Challinor 2006), caused by the intervening matter distribution of clusters. Despite its lower S/N, CMB lensing has become a robust alternative to galaxy weak lensing in SZ cluster statistics analyses (Planck Collaboration XXIV 2016).

3.2.2 Hydrostatic equilibrium

As we mentioned in Sect. 3.1.2, mass measurements can be directly carried out from X-ray observations, through the assumption of the ICM hydrostatic equilibrium (for an extensive review, see Ettori et al. 2013). Specifically, a spherical ICM is assumed, satisfying the following condition:

$$\frac{\mathrm{d}P_\mathrm{gas}}{\mathrm{d}r} = -\rho_\mathrm{gas}\frac{\mathrm{d}\phi}{\mathrm{d}r} = -\rho_\mathrm{gas}\frac{GM(r)}{r^2}\ , \tag{3.21}$$

where P_gas and ρ_gas are the ICM pressure and density, respectively, while ϕ and M are the total gravitational potential and the total mass of the cluster, respectively. Therefore, by describing the ICM as an ideal gas, the total mass contained within the radius r, $M(r)$, can be expressed as

$$M(r) = -\frac{r}{G}\frac{kT}{\mu m_p}\left(\frac{\mathrm{d}\ln\rho_\mathrm{gas}}{\mathrm{d}\ln r}+\frac{\mathrm{d}\ln T}{\mathrm{d}\ln r}\right), \tag{3.22}$$

where μ is the mean molecular weight ($\mu \simeq 0.59$ for primordial composition[9]), and m_p is the mass of the proton. To describe the ICM density profile, a β-model is usually assumed, as detailed in Section 3.1.2. In addition, as we mentioned above, temperature measurements can be achieved through X-ray emission lines and continuum cut-off. When detailed temperature profiles from X-ray observations are not available, the ICM can be described as isothermal or polytropic (Borgani 2008). In addition, X-ray temperature and luminosity measurements can be related to masses through the assumption of a self-similar growth of galaxy clusters. This assumption is motivated by the fact that gravity does not act on preferred scales, and thus clusters of different sizes are expected to be the scaled version of each other. Specifically, the self-similarity assumption leads to the following scaling relations:

$$T_X \propto M^{2/3}(1+z), \tag{3.23}$$

$$L_X \propto T_X^2(1+z)^{3/2}, \tag{3.24}$$

[9]Note that the ICM has not a primordial composition, due to supernova and AGN feedback.

where T_X and L_X are the X-ray temperature and luminosity, respectively, and M is the cluster mass. Nevertheless, observations showed that the ICM self-similarity does not hold (see, e.g., Xue & Wu 2000), leading to $L_X \propto T^{\alpha}$, with $\alpha \simeq 2.5 - 3$ (for a review, see Borgani 2008).

Compared to richness estimates provided by optical and NIR detections, the link between cluster masses and the X-ray mass proxies is theoretically better understood. Indeed, as we discussed in this section, X-ray observables can be related to masses through the assumption of hydrostatic equilibrium. The link between cluster richness and mass, instead, is harder to model, due to strong uncertainties on galaxy formation and halo occupation distribution (Qin et al. 2022; Korytov et al. 2023; Contreras & Zehavi 2023). Nevertheless, simulations show that X-ray masses are typically 10–15% underestimated due to the assumption of hydrostatic equilibrium, for which bulk motions and turbulence in the ICM are neglected (Nagai et al. 2007; Meneghetti et al. 2010; Rasia et al. 2012; Le Brun et al. 2014). Therefore, weak-lensing observations are needed to properly assess this bias. In Chapt. 4 we will extensively discuss this effect, commonly referred to as the hydrostatic bias, by analysing the clustering signal of the SZ galaxy cluster catalogue derived by Planck Collaboration XXVII (2016). In fact, as we discussed in Sect. 3.1, SZ provides a reliable estimate of the thermal pressure by free electrons and, in combination with temperature measurements from X-ray observations, it allows for estimating cluster masses (Grego et al. 2001). Specifically, $Y_{\rm SZ}$ is a robust cluster mass proxy as it quantifies the amplitude of the signal, and Kravtsov et al. (2006) proposed the X-ray equivalent of $Y_{\rm SZ}$, namely the Y_X parameter, defined as the product of the gas mass and the cluster temperature. Consequently, based on $Y_{\rm SZ}$ measurements, scaling relations between Y_X and X-ray masses can be adopted to estimate the mass of SZ selected clusters (Planck Collaboration XX 2014).

3.2.3 Cluster statistics

A natural degeneracy between cluster masses and cosmological parameters exists, as both affect the power spectrum of matter perturbations. This makes mass calibration a key step of cosmological pipelines based on cluster statistics. Similarly, cluster abundance and clustering are suitable probes for mass calibration if a cosmological model is assumed (Murata et al. 2019; Chiu et al. 2020; Lesci et al. 2022b, 2023). In Chapts. 4 and 5 we will present mass calibration analyses based on the clustering of Planck and KiDS galaxy clusters, showing that this summary statistics provides robust constraints on the normalisation of the scaling relation between cluster masses and mass proxies. In addition, in Chapt. 5 we will discuss the constraints on such scaling relation from the combination of cluster weak lensing and counts. Notably, cluster counts improve the constraints on the mass-richness scaling relation slope, and on its dependence on redshift.

3.3 Cluster counts

The abundance of galaxy clusters, commonly referred to as cluster counts, is the primary cosmological probe based on this class of objects. Indeed, dark matter halo abundance can be accurately calibrated through N-body simulations (see Sect. 1.8.4). As we discussed in this Chapter, the galaxy cluster composition is dominated by dark matter, thus N-body simulations provide a good approximation also for the abundance of these objects. In fact, cluster counts provide groundbreaking constraints on fundamental cosmological parameters, such as σ_8 and $\Omega_{\rm m}$ (Costanzi et al. 2019; Bocquet et al. 2019; Lesci et al. 2022a), which are competitive with the results from cosmic shear analyses (Troxel et al. 2018; Hikage et al.

2019; Asgari et al. 2021).

In this section, we provide an overview on literature theoretical mass functions. After the description of the analytical mass function provided by Press & Schechter (1974), we discuss the theoretical models adopted in current cosmological analyses and their dependence on cosmological parameters. Lastly, we detail the models adopted to link the theoretical mass function to mass proxies, accounting for observational uncertainties on cluster observable properties.

3.3.1 Press-Schechter mass function

Based on the spherical collapse model described in Sect. 1.8.3, Press & Schechter (1974) provided an analytical form for the halo mass function. In the Press-Schechter (PS) formalism, the probability distribution of the filtered density contrast δ_M, defined in Sect. 2.1, is assumed to be Gaussian. Thus, such probability is expressed as follows:

$$P(\delta_M) = \frac{1}{\sqrt{2\pi\sigma_M^2}} \exp\left(-\frac{\delta_M^2}{2\sigma_M^2}\right), \tag{3.25}$$

where σ_M^2 is the mass variance, defined in Eq. (2.15). As we detailed in Sect. 2.2, $P(\delta_M)$ is Gaussian only in the linear regime of perturbations. Consistent with this assumption on $P(\delta_M)$, the selected threshold for the filtered density contrast that characterises collapsed structures is approximately $\delta_c \simeq 1.68$. This value is derived in Eq. (1.126) by extrapolating the linear theory within the framework of the spherical collapse approximation. The probability of δ_M to exceed δ_c is then

$$P_{>\delta_c}(M) = \int_{\delta_c}^{\infty} P(\delta_M)\, \mathrm{d}\delta_M\,, \tag{3.26}$$

which depends on both mass and redshift as σ_M varies with time. Only isolated collapsed structures are considered, having the following probability distribution:

$$P_{>\delta_c}(M) - P_{>\delta_c}(M + \mathrm{d}M)\,. \tag{3.27}$$

However, this approach completely disregards the so-called *cloud-in-cloud problem*, which considers the possibility that a specific overdensity, initially isolated, might end up being encompassed within a larger object on a higher mass scale at a later time. This analysis exclusively focuses on recently collapsed objects, namely those that have just attained the threshold δ_c. Then, the PS mass function, $n(M)$, is expressed as follows:

$$n(M)M\mathrm{d}M = 2\rho_\mathrm{m}\left[P_{>\delta_c}(M) - P_{>\delta_c}(M + \mathrm{d}M)\right] = 2\rho_\mathrm{m}\left|\frac{\mathrm{d}P_{>\delta_c}}{\mathrm{d}\sigma_M}\right|\left|\frac{\mathrm{d}\sigma_M}{\mathrm{d}M}\right|\mathrm{d}M, \tag{3.28}$$

where ρ_m is the mean cosmic background density, and the factor 2 accounts for mass accretion from underdense regions onto the dense ones. Indeed, in Eq. (3.27), underdense regions are neglected, thus half of the mass is excluded due to the symmetry of the Gaussian distribution. Bond et al. (1991) demonstrated that, within the framework of random walk theory, the factor of 2 stems from overcounting due to the cloud-in-cloud effects. From Eq. (3.28), we obtain the following general expression for the PS mass function:

$$n(M)\mathrm{d}M = \sqrt{\frac{2}{\pi}}\frac{\rho_\mathrm{m}}{M^2}\frac{\delta_c}{\sigma_M}\left|\frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M}\right|\exp\left(-\frac{\delta_c^2}{2\sigma_M^2}\right)\mathrm{d}M\,. \tag{3.29}$$

Therefore, the PS mass function is a power-law with an exponential cut-off at high masses. This shape is reproduced also by the halo mass functions derived from N-body simulations, as described in the next section.

3.3.2 Mass functions from N-body simulations

The primary concern with the PS mass function, detailed in Sect. 3.3.1, lies in the assumption of spherical collapse. To address this issue, Sheth et al. (2001) introduced a more generalized description involving ellipsoidal collapse, and calibrated the free parameters of the mass function using numerical simulations. In particular, the generic functional form of the mass function adopted in literature has the following expression

$$\frac{\mathrm{d}n(M,z)}{\mathrm{d}M} = f(\sigma_M, z)\frac{\rho_\mathrm{m}}{M}\frac{\mathrm{d}\ln\sigma_M^{-1}}{\mathrm{d}M}\,, \tag{3.30}$$

where $\mathrm{d}n(M,z)$ is the differential volume density, and $f(\sigma_M, z)$ is the multiplicity function. The latter has a different functional form depending on modelling assumptions, and it is characterised by an exponential cut-off depending on σ_M. By comparison, the multiplicity function in the PS formalism, $f_\mathrm{PS}(\sigma_M, z)$, is expressed as follows

$$f_\mathrm{PS}(\sigma_M, z) = \sqrt{\frac{2}{\pi}}\frac{\delta_\mathrm{c}}{\sigma_M}\exp\left(-\frac{\delta_\mathrm{c}^2}{2\sigma_M^2}\right)\,. \tag{3.31}$$

Several mass function calibrations from N-body simulations have been proposed over the past decades (see, e.g., Tinker et al. 2008; Crocce et al. 2010; Bhattacharya et al. 2011; Watson et al. 2013; Despali et al. 2016), spanning the mass and redshift ranges covered by current surveys. Haloes are identified by either using spherical overdensity (SO) or friend-of-friends (FoF) halo definitions. SO is based on finding spherical regions enclosing a given overdensity Δ (Lacey & Cole 1994). Therefore, SO mass functions are consistent with mass estimates obtained from observations, for which the spherical symmetry of galaxy clusters is assumed (see Sect. 3.2). FoF finders, instead, identify haloes by connecting pairs of particles whose distance is smaller than l times the average inter-particle distance (Davis et al. 1985), with $l \simeq 0.2$ (Frenk et al. 1988). FoF haloes are known to be biased toward higher masses for haloes with a low number of particles (Warren et al. 2006). In addition, FoF finders tend to link close and different objects to each other (see, e.g., Tinker et al. 2008; Watson et al. 2013, and references therein). On the other hand, SO finders do not properly characterise non-spherical haloes.

One of the most frequently adopted mass function models in literature is the one by Tinker et al. (2008), based on a SO halo finder. This model is also used in the cosmological analyses presented in Chapts. 4 and 5. To et al. (2021b) showed that the Tinker et al. (2008) model predictions are in excellent agreement with halo mass function emulators. The same holds for the halo bias by Tinker et al. (2010). In addition, several parameterisations of the Tinker et al. (2008) mass function have been proposed, in order to account for its systematic uncertainty as a function of mass (see, e.g., Costanzi et al. 2019; To et al. 2021a).

We stress that the N-body simulations used to calibrate the aforementioned halo mass functions include only dark matter particles. However, models derived from hydrodynamical simulations would be ideal, as the impact of baryon physics may be not negligible for galaxy clusters (Henson et al. 2017; Lee et al. 2018; Giri & Schneider 2021). In fact, the impact of baryons on the halo mass function has been extensively investigated in literature (see, e.g., Cui et al. 2012; Velliscig et al. 2014; Bocquet et al. 2016; Castro et al. 2021). Nevertheless, due to the incomplete understanding of hydrodynamical processes like AGN feedback and gas cooling, the outcomes from various simulations remain unsettled.

3.3.3 Dependency on redshift and cosmological parameters

Since the halo mass function models are based on the assumption of the hierarchical structure formation, the mass function normalisation decreases as the redshift increases, as shown in

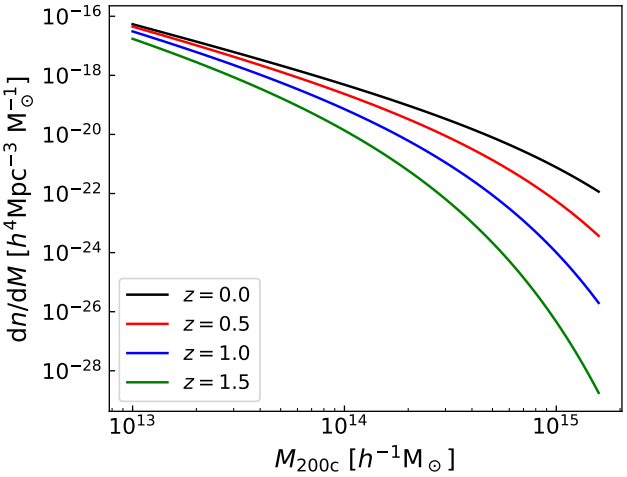

Figure 3.1: Mass function by Tinker et al. (2008), given the cosmological parameters by Planck Collaboration VI (2020, Table 2, TT,TE,EE+lowE+lensing), computed at $z = 0$ (black line), $z = 0.5$ (red line), $z = 1$ (blue line), and $z = 1.5$ (green line), as a function of M_{200c}.

Fig. 3.1. In addition, the general expression of the mass function, presented in Eq. (3.30), shows a clear dependency on fundamental cosmological parameters and redshift. In particular, the presence of ρ_m implies a direct reliance on the density parameter Ω_m, which, as discussed in Sect. 1.7, is one of the six fundamental parameters of the ΛCDM model. In addition, the growth factor f, expressed in Eq. (1.107), mainly depends on Ω_m. Specifically, f enters the mass function through σ_M, which is an integral of the power spectrum. In Fig. 3.2, we show that a decrease in Ω_m results in a lower normalisation of the mass function. In addition, thanks to the presence of σ_M, the mass function directly depends on σ_8, which is the mass standard deviation on scales of 8 h^{-1}Mpc. As detailed in Sect. 2.2, σ_8 is related to the amplitude of the primordial power spectrum, A_s, which is another fundamental parameter of the ΛCDM model. In Fig. 3.2, we can see that σ_8 variations mostly impact on the mass function exponential cut-off. Moreover, σ_8 has an impact on the mass function normalisation which is similar to that induced by Ω_m. In turn, σ_8 and Ω_m are slightly degenerate, and $S_8 := \sigma_8(\Omega_m/0.3)^\alpha$ is commonly used as a summary parameter, with $\alpha = 0.5$. Specifically, S_8 captures the direction perpendicular to the σ_8-Ω_m degeneracy. A fit of α based on the data may yield tighten constraints on S_8 (Kilbinger et al. 2013; Asgari et al. 2021), as this parameter depends on the physical scales involved in the analysis.

Besides σ_8 and Ω_m, due to the presence of σ_M, the mass function depends on the spectral index, n_s, on the other density parameters, and on the dark energy equation of state parameter, w_Λ. However, the dependency on such parameters is weaker compared to σ_8 and Ω_m, as they appear within an integral in the expression of σ_M. Additionally, we recall that the dark energy density, ρ_Λ, is expressed as follows:

$$\rho_\Lambda = \rho_{0,\Lambda}(1+z)^{3(1+w_\Lambda)}\,. \tag{3.32}$$

Consequently, an increase of w_Λ may imply an increase of Ω_Λ, causing a delay of the structure growth in a flat universe, as shown in Eq. (1.107). Nonetheless, the increase of Ω_Λ is not guaranteed as it also depends on the critical density, Eq. (1.22), which, in turn, depends the Hubble parameter, $H(z)$. From the expression of $H(z)$, in Eq. (1.64), we can see that if w_Λ increases, also $H(z)$ and the critical density increase. This would counteract the rise in ρ_Λ, and affect the final value of Ω_Λ. Furthermore, increases in $H(z)$ lead to a reduction in cosmic distances, subsequently altering the count of galaxy clusters at a specific redshift.

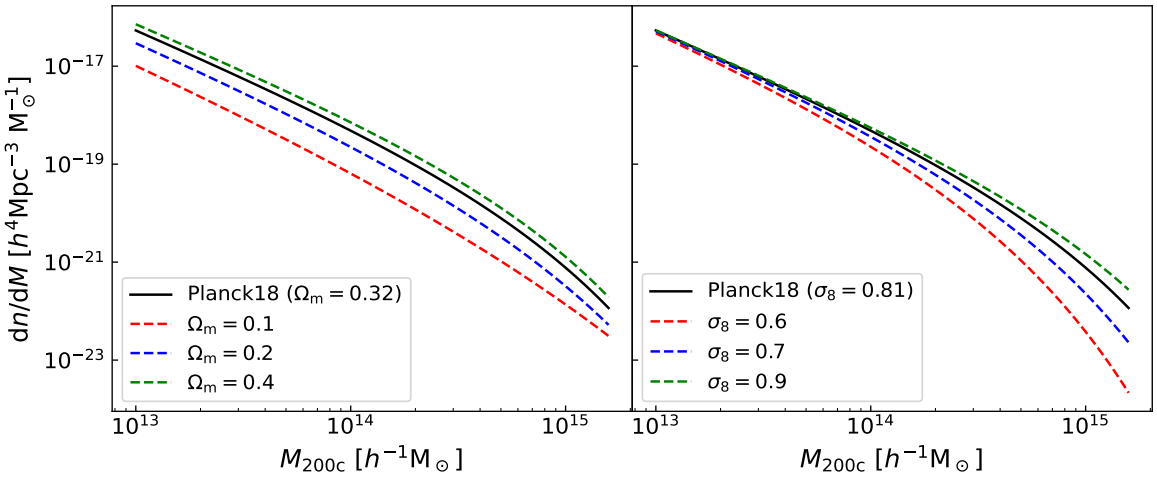

Figure 3.2: Mass function by Tinker et al. (2008) at $z = 0$, as a function of M_{200c}. The black solid lines represent the mass function with cosmological parameters by Planck Collaboration VI (2020, Table 2, TT,TE,EE+lowE+lensing). In the left and right panels, different values of Ω_m and σ_8, respectively, are assumed.

This implies a degeneracy between Ω_m and w_Λ, especially considering that the radiation density parametry is currently negligible. Another degeneracy involving w_Λ and Ω_m becomes apparent in flat universes due to the direct dependency of the mass function on ρ_m, which diminishes as the Ω_Λ value rises. Finally, cosmological analyses based on the mass function can provide constraints on the growth factor exponent (see Eq. 1.107), which is yielded by the assumption of General Relativity. In turn, the mass function provides a test for the gravitation theory. Based on forecasts of *Euclid* surveys (Laureijs et al. 2011), Sartoris et al. (2016) showed that the most robust constraints derived from cluster abundance concern σ_8 and Ω_m, with a precision of about 10^{-3}. In addition, Sartoris et al. (2016) derived competitive constraints on the massive neutrino density parameter, Ω_ν, the primordial non-Gaussianity, and the time evolution of the dark energy equation of state, assuming a CPL parameterisation for w_Λ (see Sect. 1.7).

3.3.4 Modelling of observed counts

Given a theoretical model for the mass function, constraints on cosmological parameters can be derived from galaxy cluster observations. As we previously discussed, cluster mass measurements play a crucial role in cosmological analyses, since parameters like σ_8 and Ω_m exhibit degeneracies with masses. As we detailed in Sect. 3.2, achieving precise mass measurements is not possible for every cluster observed in large surveys, because the weak-lensing or X-ray signal might be not sufficiently good. Consequently, scaling relations between observed mass proxies and masses are necessary, obtained by either stacking the observed signal or considering a representative subsample of galaxy clusters.

We remark that the mass function is defined as a volume density. Thus, cluster counts are usually modelled, as their measurement does not require the assumption of a cosmological model. The expectation value of the counts in a given bin of observed mass proxy, $\Delta\mathcal{O}_{\mathrm{ob},i}$,

and of observed redshift, $\Delta z_{\mathrm{ob},j}$, can be expressed as follows:

$$\langle N(\Delta\mathcal{O}_{\mathrm{ob},i},\Delta z_{\mathrm{ob},j})\rangle = w(\Delta\mathcal{O}_{\mathrm{ob},i},\Delta z_{\mathrm{ob},j})\;\Omega\int_0^\infty \mathrm{d}z_{\mathrm{tr}}\;\frac{\mathrm{d}V}{\mathrm{d}z_{\mathrm{tr}}\mathrm{d}\Omega}\times$$
$$\times\int_0^\infty \mathrm{d}M_{\mathrm{tr}}\;\frac{\mathrm{d}n(M_{\mathrm{tr}},z_{\mathrm{tr}})}{\mathrm{d}M_{\mathrm{tr}}}\;\int_0^\infty \mathrm{d}\mathcal{O}_{\mathrm{tr}}\;P(\mathcal{O}_{\mathrm{tr}}|M_{\mathrm{tr}},z_{\mathrm{tr}})\times$$
$$\times\int_{\Delta z_{\mathrm{ob},j}} \mathrm{d}z_{\mathrm{ob}}\;P(z_{\mathrm{ob}}|z_{\mathrm{tr}})\int_{\Delta\mathcal{O}_{\mathrm{ob},i}} \mathrm{d}\mathcal{O}_{\mathrm{ob}}\;P(\mathcal{O}_{\mathrm{ob}}|\mathcal{O}_{\mathrm{tr}})\,, \tag{3.33}$$

where V is the comoving volume, Ω is the survey effective area, while z_{tr}, M_{tr}, and $\mathcal{O}_{\mathrm{tr}}$ are the true redshift, true mass, and true mass proxy, respectively. The quantities z_{ob} and $\mathcal{O}_{\mathrm{ob}}$ are the observed redshift and mass proxy, respectively. In addition, in Eq. (3.33), $\mathrm{d}n(M_{\mathrm{tr}},z_{\mathrm{tr}})/\mathrm{d}M_{\mathrm{tr}}$ is the mass function, while $w(\Delta\mathcal{O}_{\mathrm{ob},i},\Delta z_{\mathrm{ob},j})$ is a weight factor accounting for the purity and completeness of the sample, derived from mock catalogues in the case of photometrically selected clusters (see, e.g., Lesci et al. 2022a). Analytical alternatives to $w(\Delta\mathcal{O}_{\mathrm{ob},i},\Delta z_{\mathrm{ob},j})$ can be adopted for X-ray and SZ cluster samples, as we discussed in Sect. 3.2. The probability distributions $P(z_{\mathrm{ob}}|z_{\mathrm{tr}})$ and $P(\mathcal{O}_{\mathrm{ob}}|\mathcal{O}_{\mathrm{tr}})$ account for the uncertainties on cluster redshifts and mass proxies, respectively, and can be assessed by running the cluster finding algorithm on mock galaxy catalogues. While $P(z_{\mathrm{ob}}|z_{\mathrm{tr}})$ is usually well described by a Gaussian, $P(\mathcal{O}_{\mathrm{ob}}|\mathcal{O}_{\mathrm{tr}})$ may show departures from Gaussianity in photometric cluster samples, due to percolation and projection effects which may cause biases in richness (Costanzi et al. 2019). Furthermore, in Eq. (3.33), $P(\mathcal{O}_{\mathrm{tr}}|M_{\mathrm{tr}},z_{\mathrm{tr}})$ is the probability distribution describing the relation between mass proxy and mass. It is commonly considered to follow a log-normal distribution, whose mean is given by the scaling relation between mass proxy and mass, while its standard deviation corresponds to the intrinsic scatter, σ_{intr}. The scaling relation between mass proxy and mass usually accounts for redshift evolution and can be written as follows

$$\log\frac{\mathcal{O}}{\mathcal{O}_{\mathrm{piv}}} = \alpha + \beta\log\frac{M_{\mathrm{tr}}}{M_{\mathrm{piv}}} + \gamma\log\frac{E(z_{\mathrm{tr}})}{E(z_{\mathrm{piv}})} \tag{3.34}$$

where the parameters α, β, and γ are constrained through weak-lensing or X-ray observations, $E(z)\equiv H(z)/H_0$, while $\mathcal{O}_{\mathrm{piv}}$, M_{piv}, and z_{piv} are pivot quantities, chosen to minimise the correlation between the scaling relation parameters. The last term in Eq. (3.34) accounts for deviations in the redshift evolution from what is predicted in the self-similar growth scenario (Sereno & Ettori 2015). Also the intrinsic scatter, σ_{intr}, is derived from observations, and it is expected to depend on both mass proxy and redshift. In addition, measured masses, M_{ob}, and mass proxies, $\mathcal{O}_{\mathrm{ob}}$, can be considered as scattered proxies of the latent variable M_{tr} (see Sereno et al. 2020, and references therein). The corresponding intrinsic scatters, namely between M_{ob} and M_{tr}, and between $\mathcal{O}_{\mathrm{ob}}$ and M_{tr}, are derived from simulations.

3.4 Cluster clustering

As discussed in Chapt. 2, the clustering of observed tracers of the matter density field provides robust cosmological information. Among the candidates which are suitable for such cosmological investigations, we find galaxy clusters. Large galaxy cluster samples have been exploited to provide strong cosmological constraints from both second-order and third-order statistics (see, e.g., Estrada et al. 2009; Veropalumbo et al. 2014, 2016; Marulli et al. 2018, 2021; Moresco et al. 2021; Lindholm et al. 2021, and references therein). These constraints are even

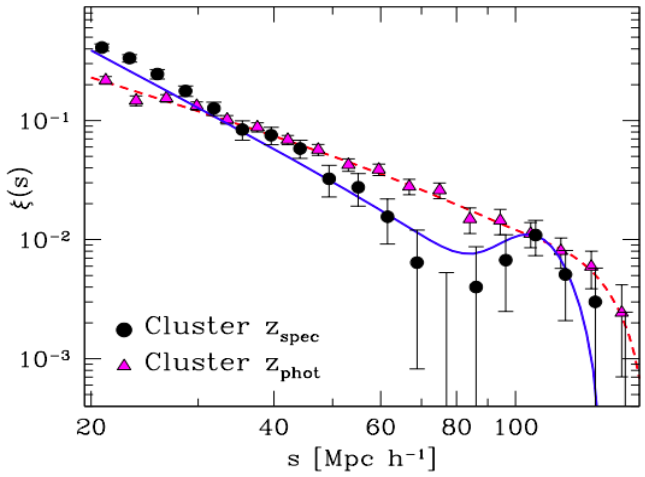

Figure 3.3: Comparison between the measured redshift-space 3D 2PCF of spectroscopic (black dots) and photometric (magenta triangles) SDSS galaxy cluster samples. The lines show the corresponding best-fit models. Image by Veropalumbo et al. (2014).

more robust when cluster clustering is combined with cluster number counts (see, e.g., Mana et al. 2013; Sartoris et al. 2016).

Galaxy clusters are more clustered than galaxies, with a clustering signal that is progressively stronger for richer systems (see, e.g., Moscardini et al. 2001, and references therein). Furthermore, clusters are relatively unaffected by nonlinear dynamics at small scales, and the redshift-space distortions on large comoving scales have a minor impact on the cluster clustering compared to galaxies because the tracer bias is larger (see, e.g., Kaiser 1987; Hamilton 1992). Having smaller theoretical uncertainties over the description of nonlinear dynamics and redshift-space distortions, the modelling of the cluster clustering signal is potentially less affected by systematics than in the galaxy clustering case. Therefore, galaxy clusters are excellent tracers of the BAO peak, especially when cluster spectroscopic redshifts are available. In Fig. 3.3, we show the 3D 2PCF signal derived from a spectroscopic cluster catalogue, where the BAO peak is evident. On the other hand, as also shown in Fig. 3.3, photometric redshift (photo-z) errors erase the BAO signal. In addition, photo-z errors imply a damping of $\xi(s)$ at low scales, as discussed in Sect. 2.4. Nevertheless, the effect of cluster photo-z errors can be modelled with high accuracy, allowing for the derivation of robust constraints on σ_8 and $\Omega_{\rm m}$ from the 3D 2PCF, as we shall discuss in Chapt. 5. This is not feasible for galaxies, as the large galaxy photo-z uncertainties erase too strongly the clustering signal along the line of sight. Cluster photo-zs, instead, have smaller uncertainties compared to galaxies, being derived for example by averaging the photo-zs of cluster members.

Another important advantage of cluster clustering lies in the theoretical calibration of the tracer bias, namely $b(M, z)$, depending on the mass and redshift of the tracers. As we detailed in Sect. 2.1, the calibration of b is currently not possible for galaxies, due to the large theoretical uncertainties in the modelling of baryonic physics. Since b can be predicted from theory for galaxy clusters, redshift-space distortions in the cluster clustering signal are excellent probes of the growth rate of perturbations. Indeed, we remark that the redshift-space distortion terms in the power spectrum depend on b and f, as we detailed in Sect. 2.4. Since clustering measurements rely on tracers with different masses and redshifts, the so-called effective bias, $b_{\rm eff}$, is considered in power spectrum models. As we shall detail in

Chapts. 4 and 5, $b_{\rm eff}$ for an ensamble of galaxy clusters can be written as follows

$$b_{\rm eff} = \frac{1}{N}\sum_{j=1}^{N}\int_0^\infty {\rm d}z_{\rm tr}\int_0^\infty {\rm d}\mathcal{O}_{\rm tr}\int_0^\infty {\rm d}M_{\rm tr}\; b(M_{\rm tr}, z_{\rm tr})\, P(M_{\rm tr}|\mathcal{O}_{\rm tr}, z_{\rm tr})\;\times$$
$$\times\, P(z_{{\rm ob},j}|z_{\rm tr})\, P(\mathcal{O}_{{\rm ob},j}|\mathcal{O}_{\rm tr}), \tag{3.35}$$

where the subscripts "tr" and "ob" refer to true and observed quantities, respectively, $\mathcal{O}$ is a mass proxy, N is the total number of galaxy clusters in the sample, j is the cluster index, while $P(z_{{\rm ob},j}|z_{\rm tr})$ and $P(\mathcal{O}_{{\rm ob},j}|\mathcal{O}_{\rm tr})$ are described in Sect. 3.3.4. The probability distribution $P(M_{\rm tr}|\mathcal{O}_{\rm tr}, z_{\rm tr})$ is analogous to $P(O_{\rm tr}|\mathcal{M}_{\rm tr}, z_{\rm tr})$ described in Sect. 3.3.4 (see Chapt. 5). We remark that Eq. (3.35) is an arithmetic average of the expected values of b derived for each cluster in the sample. As an alternative, $b_{\rm eff}$ can be expressed as an average of b weighted by the halo mass function (see, e.g., To et al. 2021b). In the latter case, however, reliable estimates of the sample completeness must be included in the model, similarly to the cluster counts models. Thus, as we shall see in the following, using Eq. (3.35) allows for including low-mass and high-redshift cluster samples in clustering analyses, for which reliable completeness estimates are not available.

Chapter 4

Planck clusters: clustering

Since cosmological parameters are inferred with high precision in current cluster statistical analyses, accurate cluster mass calibrations are of critical importance. Indeed, an incomplete assessment of systematic uncertainties affecting the derived masses may lead to significant biases in the cosmological constraints (Planck Collaboration XXIV 2016; Abbott et al. 2020). As we discussed in Sect. 3.2, simulations show that X-ray masses are typically 10-15% underestimated due to the assumption of hydrostatic equilibrium, for which bulk motions and turbulence in the intra-cluster medium are neglected (Nagai et al. 2007; Meneghetti et al. 2010; Rasia et al. 2012; Le Brun et al. 2014). Weak-lensing mass estimates can also be biased by the inaccuracy of density profile models (Oguri & Hamana 2011), baryonic effects influencing the halo concentration (Henson et al. 2017; Shirasaki et al. 2018; Beltz-Mohrmann & Berlind 2021), halo orientation (Becker & Kravtsov 2011; Dietrich et al. 2014; Zhang et al. 2023), and projections (Simet et al. 2017; Melchior et al. 2017). As the biases in the weak-lensing mass estimates are theoretically better understood, weak-lensing observations are exploited for calibrating the main bias affecting X-ray masses, called hydrostatic bias, $b_{\rm h}$ (von der Linden et al. 2014; Hoekstra et al. 2015; Planck Collaboration XXIV 2016; Smith et al. 2016; Sereno et al. 2017). In particular, the relation between the X-ray mass, M_X, and the true mass, $M_{\rm tr}$, is usually expressed as follows:

$$M_X = (1 - b_{\rm h}) M_{\rm tr} \,. \tag{4.1}$$

In this Chapter, based on the work by Lesci et al. (2023), we present the analysis of the mass bias of the SZ selected Planck clusters (Planck Collaboration XXIV 2016; Planck Collaboration XXVII 2016), which is referred to as the Planck mass bias, $b_{\rm SZ}$. Planck cluster masses are expected to be biased low as they are derived from a scaling relation based on X-ray observations of 20 relaxed clusters at $z < 0.2$ (Arnaud et al. 2010; Planck Collaboration XX 2014). Indeed, as we detailed in Sect. 3.2.2, X-ray and SZ mass proxies are closely related, and this allows for building up SZ mass scaling relations based on X-ray observations. Specifically, we obtained an estimate of $b_{\rm SZ}$ that is independent of lensing observations by exploiting the monopole of the 3D 2PCF of the galaxy clusters in the sample provided by Planck Collaboration XXVII (2016). We assumed a standard ΛCDM cosmological model, adopting the CMB constraints on cosmological parameters from Planck Collaboration VI (2020) as priors. In addition, we adopted the same priors on $(1 - b_{\rm SZ})$ as were used in the Planck cluster count analysis carried out by Planck Collaboration XXIV (2016) in order to constrain the matter density parameter, $\Omega_{\rm m}$, and the amplitude of the matter power spectrum, σ_8.

The statistical analyses presented in this Chapter were performed with the CosmoBolognaLib[1] (CBL; Marulli et al. 2016), a set of free software C++/Python numerical librar-

[1] https://gitlab.com/federicomarulli/CosmoBolognaLib/

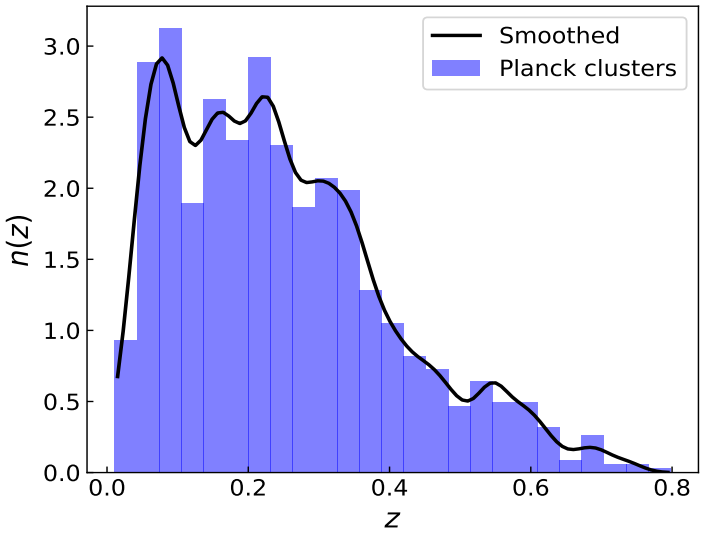

Figure 4.1: Redshift distribution $n(z)$ of the galaxy clusters considered in the analysis. The blue histogram shows the observed binned $n(z)$, and the black curve represents $n(z)$ smoothed with a Gaussian kernel (with an rms equal to 0.02) that was used to build up the random catalogue.

ies for cosmological calculations. Specifically, the measurements and the statistical Bayesian analyses were both performed with the CBL v6.1.

This Chapter is organised as follows. In Section 4.1 we describe the data set and the methods we used to estimate the 2PCF of the sample. In Section 4.2 we describe the 2PCF model, focusing on the dependence of the effective bias on the mass-observable scaling relation. In Section 4.3 we show our constraints on $b_{\rm SZ}$ and detail the cosmological analysis, while in Sect. 4.4 we draw our conclusions.

4.1 Data set and 2PCF measurement

4.1.1 The Planck cluster sample

Following Planck Collaboration XXIV (2016), we based our analysis on the cosmological sample consisting of detections by the matched multi-filters 3 (MMF3) detection algorithm (Melin et al. 2006, 2012) derived from the general Planck full-mission Sunyaev–Zeldovich catalogue (PSZ2; Planck Collaboration XXVII 2016). We considered clusters with a confirmed counterpart in external data sets and an assigned redshift estimate (see Table 9 in Planck Collaboration XXVII 2016), with a redshift limit $z \leq 0.8$, for a total of 920 clusters. We applied this redshift cut to exclude 5 clusters that are isolated with respect to the bulk of the redshift distribution because they hindered the derivation of a reliable smoothed redshift distribution, which is necessary for the construction of the random sample (see Section 4.1.2). In addition, differently from Planck Collaboration XXIV (2016), we did not apply any cut in the S/N. This does not imply any potential problems due to the reliability of the selection function at low S/N because our model does not rely on assumptions on the sample completeness (see Section 4.2.2).

4.1.2 Random catalogue

The random catalogue used for the 2PCF measurement is 100 times larger than the Planck cluster sample. We smoothed the observed redshift distribution, $n(z)$, with a Gaussian kernel

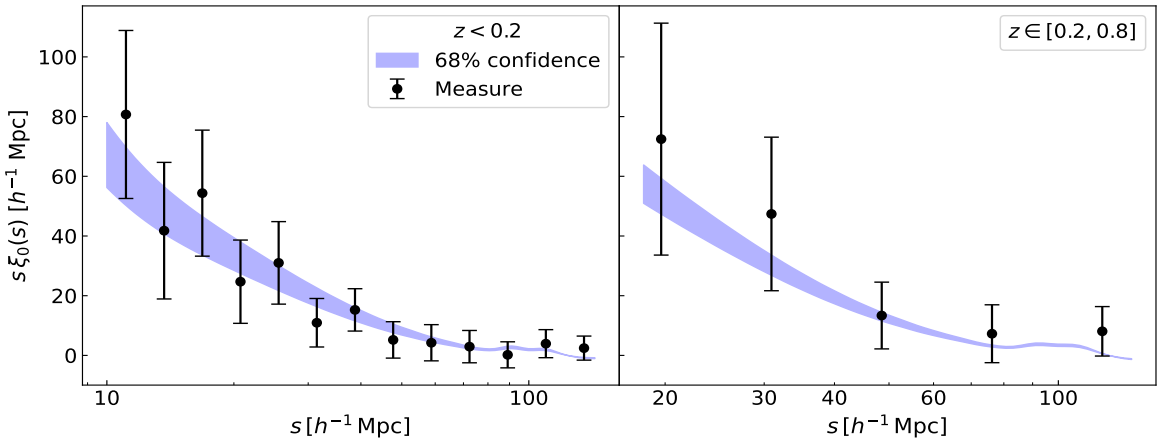

Figure 4.2: Redshift-space 2PCF monopole (black dots) of the Planck clusters in the spatial range $s \in [10, 150]$ h^{-1}Mpc for $z < 0.2$ (left plot), and $s \in [15, 150]$ h^{-1}Mpc for $z \in [0.2, 0.8]$ (right plot). The blue bands represent the model 68% confidence level derived from the posterior of the free parameters considered in the analysis described in Section 4.3.1.

with an rms equal to 0.02 (see Fig. 4.1). Then we extracted random redshifts from this distribution. Random R.A.-Dec pairs were extracted by following the sample angular selection function. It consists of the combination of the MMF3 survey mask[2], namely $\mathcal{M}_{\rm s}$, the hole mask excluding contaminated regions (e.g. by stars or large galaxies), $\mathcal{M}_{\rm h}$, and the error function completeness. Both $\mathcal{M}_{\rm s}$ and $\mathcal{M}_{\rm h}$ are equal to 0 if the region is masked, otherwise, they are equal to 1. The error function completeness is defined as (Planck Collaboration XXIX 2014)

$$P(d|Y_{500}, \sigma_{Yi}(\theta_{500}), q) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{Y_{500} - q\sigma_{Yi}(\theta_{500})}{\sqrt{2}\sigma_{Yi}(\theta_{500})}\right)\right], \tag{4.2}$$

where d is the Boolean detection state, $\mathrm{erf}(x)$ is the Gauss error function, and Y_{500} and θ_{500} are the observed SZ signal and the detection angular scale within a critical radius R_{500}, respectively. The latter is defined as the distance from the cluster centre inside which the mean density is 500 times the critical density of the Universe at the given redshift, z. M_{500} is defined as the mass enclosed within R_{500}. In Eq. (4.2), σ_{Yi} is the standard deviation of pixels for a given patch i, computed by following Melin et al. (2006), and q is the S/N threshold. As we did not apply any S/N cut to the sample, q corresponds to the minimum threshold adopted by Planck Collaboration XXVII (2016) in the detection process, namely $q = 4.5$. In Eq. (4.2), we assumed the sample mean values of Y_{500} and θ_{500}. We verified that using the median values of these quantities does not introduce significant variations in the final results. Then we extracted random angular positions, for each of which we sampled a number in the range $[0, 1]$. When this number was higher than the product of $\mathcal{M}_{\rm s}$, $\mathcal{M}_{\rm h}$ and $P(d|Y_{500}, \sigma_{Yi}, q)$, the random angular position was rejected. As an alternative to the error function completeness in Eq. (4.2), we weighted the pairs in the 2PCF estimator (described in Section 4.1.3) by $1/\sigma_{Yi}^{\rm norm}$, where $\sigma_{Yi}^{\rm norm}$ is equal to σ_{Yi} divided by its minimum value, namely $\sigma_{Yi}^{\rm norm} = \sigma_{Yi}/\min(\sigma_{Yi})$. We verified that this approach provides results that fully agreed with what was derived from the application of the error function completeness.

[2] https://irsa.ipac.caltech.edu/data/Planck/release_2/ancillary-data/HFI_Products.html

4.1.3 Clustering measurement

We estimated the redshift-space 2PCF monopole using the Landy & Szalay (1993, LS) estimator, detailed in Sect. 2.1. We recall that the LS estimator has the following expression:

$$\xi(s) = \frac{N_{RR}}{N_{DD}} \frac{DD(s)}{RR(s)} - 2\frac{N_{RR}}{N_{DR}} \frac{DR(s)}{RR(s)} + 1\,, \tag{4.3}$$

where $DD(s)$, $RR(s)$, and $DR(s)$ are the number of data-data, random-random, and data-random pairs with separation $s \pm \Delta s$, respectively, while N_{DD}, N_{RR} , and N_{DR} are the total number of data-data, random-random, and data-random pairs, respectively. To convert the observed coordinates into the comoving ones, we assumed the cosmological parameters by Planck Collaboration VI (2020), TT, TE, EE+lowE+lensing (referred to as Planck18 hereafter). The LS estimator is extensively used in clustering analyses as it is unbiased with minimum variance for an infinitely large random sample and when $|\xi| \ll 1$ (Hamilton 1992; Kerscher et al. 2000; Labatie et al. 2010; Keihänen et al. 2019).

Specifically, we measured the 2PCF considering two redshift bins, namely $z < 0.2$ and $z \in [0.2, 0.8]$, containing 407 and 513 galaxy clusters, respectively. We considered the cluster-centric radial range $s \in [10, 150]$ h^{-1}Mpc, excluding from the analysis the 2PCF measure at $s < 15$ h^{-1}Mpc in the second redshift bin due to the lack of data-data pairs. Moreover, we integrated the 2PCF measurements over larger s bins in $z \in [0.2, 0.8]$, compared to those adopted for the lower redshift bin. In this way, we compensated for the reduction of cluster pairs with small s distances, caused by the lower cluster density at high redshifts. We estimated the covariance matrix, including the cross-covariance between radial and redshift bins, through a bootstrap procedure. In particular, we considered 200 angular regions and two redshift regions, corresponding to the redshift bins, and resampled the observed and random catalogues 2000 times. We corrected the inverted covariance matrix following Hartlap et al. (2007). In Fig. 4.2 we show the measured 2PCF monopole, ξ_0. We did not include the other multipoles in the analysis, as we verified that their contribution is negligible.

4.2 Modelling

We modelled the 2PCF of Planck clusters by accounting for geometric and redshift-space distortions. In addition, different to what was done in the cluster count analysis by Planck Collaboration XXIV (2016), our model does not rely on assumptions on the sample completeness. We show in Section 4.3 that this approach leads to constraints on $b_{\rm SZ}$ and cosmological parameters that fully agree with those derived by Planck Collaboration XXIV (2016) and Planck Collaboration VI (2020).

4.2.1 Two-point correlation function model

The lth order 2PCF multipole, $\xi_l(s)$, can be expressed as follows:

$$\xi_l(s) = i^l \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi^2}\, k^2 P_l(k) j_l(ks), \tag{4.4}$$

where j_l is the spherical Bessel function of order l, and P_l is the redshift-space matter power spectrum multipole of order l,

$$P_l(k) = \frac{2l+1}{2\alpha_\perp^2 \alpha_\parallel} \int_{-1}^{1} \mathrm{d}\mu\, P(k', \mu') L_l(\mu). \tag{4.5}$$

In Eq. (4.5), L_l is the Legendre polynomial of order l, and μ is the line of sight cosine. Moreover, in Eq. (4.5), we accounted for the Alcock & Paczynski (1979, AP) geometric distortions (described in Section 2.5) caused by the assumption of a fiducial cosmology used to convert the cluster observed coordinates into comoving ones in Eq. (4.3). Specifically, k' and μ' have the following functional forms (Beutler et al. 2014):

$$k' = \frac{k}{\alpha_\perp}\left[1+\mu^2\left(\frac{\alpha_\perp^2}{\alpha_\parallel^2}-1\right)\right]^{1/2}, \tag{4.6}$$

$$\mu' = \mu\frac{\alpha_\perp}{\alpha_\parallel}\left[1+\mu^2\left(\frac{\alpha_\perp^2}{\alpha_\parallel^2}-1\right)\right]^{-1/2}, \tag{4.7}$$

where $\alpha_\perp$ and $\alpha_\parallel$ are expressed as

$$\alpha_\parallel = \frac{H^{\rm fid}(z)r_{\rm s}^{\rm fid}(z_d)}{H(z)r_{\rm s}(z_d)}, \tag{4.8}$$

$$\alpha_\perp = \frac{D_A(z)r_{\rm s}^{\rm fid}(z_d)}{D_A^{\rm fid}(z)r_{\rm s}(z_d)}. \tag{4.9}$$

Here, $H^{\rm fid}(z)$ and $D_A^{\rm fid}(z)$ are the fiducial values for the Hubble constant and angular diameter distance, respectively, and $r_{\rm s}^{\rm fid}(z_{\rm d})$ is the fiducial sound horizon at the drag redshift, $z_{\rm d}$. We stress that the AP correction takes place only in the cosmological analysis described in Section 4.3.3. To derive $b_{\rm SZ}$, as detailed in Section 4.3.1, we fixed the cosmological parameters to the fiducial ones. In Eq. (4.5), $P(k',\mu')$ is the redshift-space dark matter power spectrum expressed as (Taruya et al. 2010):

$$P(k',\mu') = D_{\rm FoG}(k',\mu',f,\sigma_{\rm v})\Big[b_{\rm eff}^2P_{\delta\delta}(k') + 2fb_{\rm eff}\mu'^2P_{\delta\theta}(k') + \\ + f^2\mu'^4P_{\theta\theta}(k') + b_{\rm eff}^3A(k',\mu',f) + b_{\rm eff}^4B(k',\mu',f)\Big], \tag{4.10}$$

where $P_{\delta\delta}$, $P_{\theta\theta}$, and $P_{\delta\theta}$ are the real-space auto power spectra of the density and velocity divergence and their cross power spectrum, respectively. These spectra are estimated in the standard perturbation theory (SPT), consisting of expanding the statistics as a sum of infinite terms, corresponding to the n-loop corrections (see e.g. Gil-Marín et al. 2012). Considering corrections up to the first-loop order, the power spectrum can be modelled as follows:

$$P^{\rm SPT}(k) = P_{\rm L}(k) + P^{(1)}(k) = P_{\rm L}(k) + 2P_{13}(k) + P_{22}(k), \tag{4.11}$$

where the leading-order term, $P_{\rm L}(k)$, is the linear matter power spectrum, computed with `CAMB`[3] (Lewis & Challinor 2011), while the one-loop correction terms are computed with the `CPT Library`[4] (Taruya & Hiramatsu 2008). In Eq. (4.10), $D_{\rm FoG}(k',\mu',f,\sigma_{\rm v})$ is a Gaussian damping function representing the fingers-of-God effect, having the following functional form:

$$D_{\rm FoG}(k',\mu',f,\sigma_{\rm v}) = e^{-k'^2\mu'^2f^2\sigma_{\rm v}^2}, \tag{4.12}$$

where f is the linear growth rate, and $\sigma_{\rm v}^2$ is the linear velocity dispersion, computed as (Taruya et al. 2010):

$$\sigma_{\rm v}^2 = \frac{1}{3}\int\frac{{\rm d}^3\boldsymbol{k}}{(2\pi)^3}\frac{P_{\rm L}(k)}{k^2}. \tag{4.13}$$

[3] https://camb.info/

[4] http://www2.yukawa.kyoto-u.ac.jp/~atsushi.taruya/cpt_pack.html

In Eqs. (4.10) and (4.13), $P_{\rm L}(k)$ is computed at the mean redshift of the cluster sub-sample in the given redshift bin. In addition, in Eq. (4.10), $b_{\rm eff}$ is the effective bias, defined in Section 4.2.2, and the functions $A(k', \mu', f)$ and $B(k', \mu', f)$ are correction terms derived from SPT (Taruya et al. 2010; de la Torre & Guzzo 2012; García-Farieta et al. 2020).

4.2.2 Effective bias

The effective bias, $b_{\rm eff}$, has the following functional form:

$$b_{\rm eff} = \frac{1}{N_{\rm cl}} \sum_{j=1}^{N} b(Y^{\rm ob}_{500,j}, z^{\rm ob}_j), \tag{4.14}$$

where $N_{\rm cl}$ is the number of clusters in the sample, $Y^{\rm ob}_{500,j}$ and $z^{\rm ob}_j$ are the observed SZ signal and redshift, respectively, of the jth cluster, and $b(Y^{\rm ob}_{500,j}, z^{\rm ob}_j)$ is expressed as

$$\begin{aligned} b(Y^{\rm ob}_{500,j}, z^{\rm ob}_j) = & \frac{1}{n(Y^{\rm ob}_{500,j}, z^{\rm ob}_j)} \\ & \times \int_0^\infty {\rm d}M_{500} \, \frac{{\rm d}n(M_{500}, z^{\rm ob}_j)}{{\rm d}M_{500}} \, b(M_{500}, z^{\rm ob}_j) \\ & \times \int_0^\infty {\rm d}Y_{500} \, P(Y_{500}|M_{500}, z^{\rm ob}_j) \, P(Y_{500}|Y^{\rm ob}_{500,j}) \, , \end{aligned} \tag{4.15}$$

where $b(M_{500}, z)$ is the halo bias, for which the model by Tinker et al. (2010) was assumed, and $P(Y_{500}|Y^{\rm ob}_{500,j})$ is a Gaussian whose mean is $Y^{\rm ob}_{500,j}$ and its rms is given by the error on $Y^{\rm ob}_{500,j}$. In addition, $P(Y_{500}|M_{500}, z)$ is a log-normal whose mean is given by the mass-observable scaling relation, and its rms is given by the intrinsic scatter, $\sigma_{\ln Y}$,

$$P(\ln Y_{500}|M_{500}, z) = \frac{1}{\sqrt{2\pi}\sigma_{\ln Y}} e^{-\ln^2(Y_{500}/\bar{Y}_{500})/(2\sigma^2_{\ln Y})} \, . \tag{4.16}$$

Specifically, following Planck Collaboration XXIV (2016), we assumed $\sigma_{\ln Y}$ to be independent of Y_{500} and redshift, and the expected value of SZ signal, $\bar{Y}_{500}$, can be expressed as

$$E^{-\beta}(z) \left[\frac{D^2_{\rm A}(z) \bar{Y}_{500}}{10^{-4}\,{\rm Mpc}^2} \right] = Y_* \left[\frac{h}{0.7} \right]^{-2+\alpha} \left[\frac{(1-b_{\rm SZ})\, M_{500}}{6 \times 10^{14}\,{\rm M}_\odot} \right]^{\alpha} , \tag{4.17}$$

where $E(z) \equiv H(z)/H_0$, with $H(z)$ being the Hubble function and H_0 the Hubble constant, $D_A(z)$ is the angular diameter distance, $h \equiv H_0/100$, $b_{\rm SZ}$ is the Planck mass bias, and Y_*, α, and β are the scaling relation parameters. In addition, $n(Y^{\rm ob}_{500,j}, z^{\rm ob}_j)$ in Eq. (4.15) is expressed as

$$\begin{aligned} n(Y^{\rm ob}_{500,j}, z^{\rm ob}_j) = & \int_0^\infty {\rm d}M_{500} \, \frac{{\rm d}n(M_{500}, z^{\rm ob}_j)}{{\rm d}M_{500}} \\ & \times \int_0^\infty {\rm d}Y_{500} \, P(Y_{500}|M_{500}, z^{\rm ob}_j) \, P(Y_{500}|Y^{\rm ob}_{500,j}) \, , \end{aligned} \tag{4.18}$$

where ${\rm d}n(M_{500}, z)/{\rm d}M_{500}$ is the halo mass function, for which the model by Tinker et al. (2008) was assumed.

Table 4.1: Free parameters considered in the analysis detailed in Section 4.3.1.

Parameter	Description	Prior	Posterior
$b_{\rm SZ}$	Planck mass bias	[-2, 0.9]	$0.38^{+0.14}_{-0.11}$
$\log Y_*$	Normalisation of the mass-observable relation	$\mathcal{N}(-0.19, 0.02)$	—
α	Slope of the mass-observable relation	$\mathcal{N}(1.79, 0.08)$	—
β	Redshift evolution of the mass-observable relation	$\mathcal{N}(0.66, 0.50)$	—
$\sigma_{\ln Y}$	Intrinsic scatter of the mass-observable relation	$\mathcal{N}(0.173, 0.023)$	—

Notes. In the third column, the priors on the parameters are listed. In particular, a range between square brackets represents a uniform prior, and $\mathcal{N}(\mu, \sigma)$ stands for a Gaussian prior with mean μ and standard deviation σ. In the fourth column, we show the median values of the 1D marginalised posteriors, along with the 16th and 84th percentiles. The posterior distributions of $\log Y_*$, α, β, and $\sigma_{\ln Y}$ are not shown since these parameters are not constrained in our analysis.

4.2.3 Likelihood

For the Bayesian analysis performed in this work, a standard Gaussian likelihood was considered,

$$\mathcal{L} \propto \exp(-\chi^2/2)\,, \tag{4.19}$$

with

$$\chi^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\xi_i^d - \xi_i^m\right) C_{i,j}^{-1} \left(\xi_j^d - \xi_j^m\right)\,, \tag{4.20}$$

where N is the number of comoving separation bins in which the 2PCF is computed, d and m indicate data and model, respectively, and $C_{i,j}^{-1}$ is the inverse of the covariance matrix. As detailed in Section 4.1.3, $C_{i,j}$ was derived through a bootstrap resampling.

4.3 Results

Based on the methods outlined in Sections 4.1 and 4.2, we carried out an analysis of the redshift-space 2PCF monopole of the Planck cluster sample (Planck Collaboration XXVII 2016). Specifically, we detail in Section 4.3.1 the derivation of the $(1 - b_{\rm SZ})$ constraint, performed by assuming the Planck18 cosmological results as priors. In Section 4.3.3 we present the constraints on cosmological parameters, obtained by assuming priors on $b_{\rm SZ}$ from external data sets.

4.3.1 Constraint on $b_{\rm SZ}$

In order to derive a constraint on the Planck mass bias, $b_{\rm SZ}$, we fixed the cosmological parameters to the Planck18 median values. We also assumed the priors on the mass-observable scaling relation parameters in Eq. (4.17), namely Y_*, α, β, and $\sigma_{\ln Y}$, adopted by Planck Collaboration XXIV (2016). In particular, this scaling relation was derived from X-ray observations of 20 relaxed clusters at $z < 0.2$ (Arnaud et al. 2010; Planck Collaboration XX 2014). Finally, we assumed a large flat prior on $b_{\rm SZ}$. In Table 4.1 we summarise the priors used for this analysis, along with the result on the mass bias, namely $(1 - b_{\rm SZ}) = 0.62^{+0.14}_{-0.11}$. The corresponding effective bias estimates are $b_{\rm eff} = 4.61^{+0.39}_{-0.36}$ and $b_{\rm eff} = 6.46^{+0.35}_{-0.37}$ for $z < 0.2$ and $z \in [0.2, 0.8]$, respectively. The obtained value of $(1 - b_{\rm SZ})$ is lower than what was predicted by numerical simulations (Nagai et al. 2007; Piffaretti & Valdarnini 2008; Meneghetti

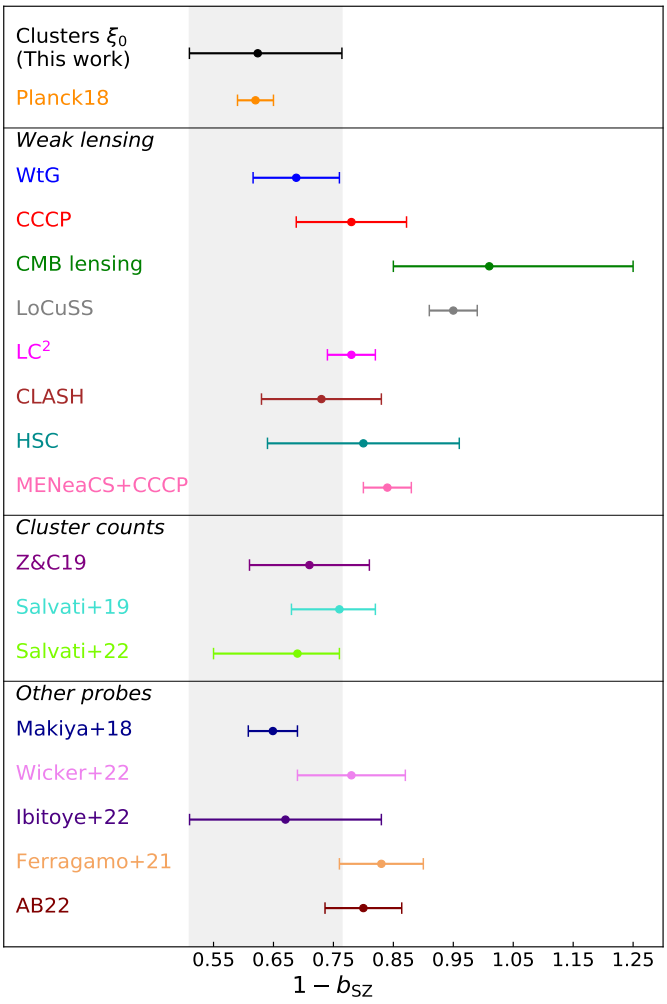

Figure 4.3: Comparison of the results on $(1 - b_{\rm SZ})$ with the literature. The median, 16th and 84th percentiles are shown. The black dot shows the constraint derived in this work. Then, in order from top to bottom, the following results are shown: Planck Collaboration VI (2020) (orange), von der Linden et al. (2014) (blue) Hoekstra et al. (2015) (red), Planck Collaboration XXIV (2016) (dark green), Smith et al. (2016) (grey), Sereno et al. (2017) (magenta), Penna-Lima et al. (2017) (brown), Medezinski et al. (2018a) (cyan), Herbonnet et al. (2020) (pink), Zubeldia & Challinor (2019) (purple), Salvati et al. (2019) (turquoise), Salvati et al. (2022) (green), Makiya et al. (2018) (dark blue), Wicker et al. (2022) (violet), Ibitoye et al. (2022) (indigo), Ferragamo et al. (2021) (orange), and Aguado-Barahona et al. (2022) (dark brown).

et al. 2010; Rasia et al. 2012; Le Brun et al. 2017; Henson et al. 2017; Gianfagna et al. 2023), but in line with what was found by Planck Collaboration VI (2020). We remark that even though the constraints on $b_{\rm eff}$ derived from the two redshift bins have similar associated uncertainties, our result on $(1 - b_{\rm SZ})$ is dominated by the 2PCF signal measured at low redshift. We obtained $(1 - b_{\rm SZ}) = 0.67^{+0.22}_{-0.13}$ for $z < 0.2$ and $(1 - b_{\rm SZ}) = 0.58^{+0.55}_{-0.31}$ for $z \in [0.2, 0.8]$. This is expected because $b_{\rm eff}$ is directly derived from the mass-observable relation, and the number of clusters is comparable in the considered redshift bins. On the other hand, the constraints on $(1 - b_{\rm SZ})$ depend on the 2PCF measurements, for which we obtained larger uncertainties in the high-redshift bin. In both redshift bins, we found $\chi^2_{\rm red} \sim 0.9$, where $\chi^2_{\rm red}$ is the reduced χ^2. Specifically, we considered three effective free model parameters because the change in the model within the assumed priors on β and $\sigma_{\ln Y}$ is not statistically significant.

In Fig. 4.3 we show a comparison between our constraint on $(1 - b_{\rm SZ})$ and the results obtained from the literature. In presence of systematic uncertainties, we added them in quadrature to the statistical ones. By combining primary CMB likelihood and cluster counts, Planck Collaboration VI (2020) derived $(1-b_{\rm SZ}) = 0.62 \pm 0.03$ (orange dot in Fig. 4.3), which fully agrees with our result. Regarding the Planck mass estimates derived from galaxy weak lensing, we found a 1σ agreement with Weighting the Giants (WtG; von der Linden et al. 2014), the Canadian Cluster Comparison Project (CCCP; Hoekstra et al. 2015), the Literature Catalogs of weak-lensing Clusters of galaxies ($\rm LC^2$; Sereno et al. 2017), the Cluster Lensing And Supernova survey with Hubble (CLASH; Penna-Lima et al. 2017), and the Subaru Hyper Suprime-Cam (HSC; Medezinski et al. 2018a). We found an agreement of only 2σ with the results from the Local Cluster Substructure Survey (LoCuSS; Smith et al. 2016), the Multi Epoch Nearby Cluster Survey (MENeaCS) combined with updated mass weak-lensing estimates in CCCP (MENeaCS+CCCP; Herbonnet et al. 2020), and with the result obtained from CMB lensing by Planck Collaboration XXIV (2016). When comparing our results to other analyses based on cluster counts, we found an agreement of 1σ with Zubeldia & Challinor (2019), Salvati et al. (2019), and Salvati et al. (2022). Concerning the results derived from the power spectra of the Planck thermal Sunyaev–Zeldovich effect, our constraint agrees within 1σ with Makiya et al. (2018) and Ibitoye et al. (2022). We also found a good agreement with the constraint by Wicker et al. (2022) based on measurements of the cluster gas mass fraction. The hydrostatic bias estimates from dynamical masses by Ferragamo et al. (2021) and Aguado-Barahona et al. (2022) agree to within 1σ with our result. In Section 4.3.2 we discuss the impact of the adopted modelling choices on our result, finding that the derived constraint on $b_{\rm SZ}$ is robust with respect to the investigated systematic uncertainties.

As many observational studies claimed a redshift dependence of the hydrostatic bias (Smith et al. 2016; Sereno et al. 2017; Salvati et al. 2019, 2022; Wicker et al. 2022), we investigated this possibility by expressing $b_{\rm SZ}$ as follows:

$$b_{\rm SZ} = \eta \left(\frac{1+z}{1+z_{\rm piv}} \right)^{\zeta} , \tag{4.21}$$

where $z_{\rm piv} = 0.25$ is the mean redshift of the sample, η is the normalisation, and ζ parametrises the redshift dependence of the mass bias. Our analysis does not constrain ζ, implying that it is not necessary to explain our data. We stress that the redshift dependence of $b_{\rm SZ}$ was derived from cluster statistics only in the case of a strong prior on the total value of $b_{\rm SZ}$, with a significant dependence on the sample (Salvati et al. 2019, 2022; Wicker et al. 2022).

4.3.2 Assessment of systematics on $b_{\rm SZ}$

To assess the robustness of the constraint on $b_{\rm SZ}$ derived in Section 4.3.1, we included the power spectrum damping due to redshift uncertainties in the analysis. No redshift errors are quoted in Planck data products, and therefore, we expressed this damping by means of a free parameter. Specifically, we replaced Eq. (4.13) by the following expression:

$$\sigma_{\rm v,tot} = \sqrt{\sigma_{\rm v}^2 + \sigma_{{\rm v},z}^2} , \tag{4.22}$$

where $\sigma_{\rm v}$ is defined in Eq. (4.13), while $\sigma_{{\rm v},z}$ is the velocity dispersion caused by redshift errors, having the following functional form:

$$\sigma_{{\rm v},z} \equiv \frac{c\,\sigma_z (1+\bar{z})}{H(\bar{z})} . \tag{4.23}$$

In this equation, $\bar{z}$ is the mean redshift of the sub-sample in a given redshift bin, c is the speed of light, $H(z)$ is the Hubble function, and σ_z is the typical redshift uncertainty of the sample. By assuming a flat prior on σ_z, namely $\sigma_z \in [0, 0.1]$, we derived $\sigma_z = 0.003^{+0.002}_{-0.002}$, which is in line with the fact that most of the cluster redshifts are spectroscopic, and $(1-b_{\rm SZ}) = 0.65^{+0.15}_{-0.12}$, which fully agrees with our previous result.

In addition, we analysed the 2PCF monopole of the Planck union catalogue, containing the clusters detected with the three detection algorithms adopted by Planck Collaboration XXVII (2016). By assuming the same sample selections and bins of redshift and radius described in Section 4.1, we found $(1 - b_{\rm SZ}) = 0.59^{+0.12}_{-0.09}$, which is in line with the constraint derived in Section 4.3.1. This implies that our result is independent of the adopted cluster detection algorithm. We also performed the analysis by considering the clusters in the MMF3 sample with $S/N > 6$, and for which the `COSMO` entry in the union catalogue is set to `'T'`, meaning that these clusters are part of the cosmological sample, following Planck Collaboration XXIV (2016). Because the statistics are poorer in this case, we analysed the 2PCF in a single redshift bin including clusters with $z \leq 0.8$ for a total of 430 objects. As the modelling provides reduced χ^2 estimates that are not close to 1, we conclude that in this case, the 2PCF signal does not allow a reliable constraint on $b_{\rm SZ}$.

In order to further assess the robustness of our results on $(1 - b_{\rm SZ})$, we computed the 2PCF model at the sample median redshifts for each redshift bin, instead of adopting the mean redshift, as discussed in Section 4.2. In this way, we derived a shift of the median $(1 - b_{\rm SZ})$ of 0.006σ. In addition, the reduction of the 2PCF radial range to $s \in [15, 150]$ h^{-1}Mpc or to $s \in [10, 90]$ h^{-1}Mpc implies comparable results, namely shifts of the median $(1 - b_{\rm SZ})$ lower than 0.6σ, and variations in the 1σ interval extension lower than 50%. We also checked the impact of a change in the definition of the effective bias, $b_{\rm eff}$, assuming the median of the halo bias distribution instead of considering its mean, as done in Eq. (4.14). In this case, we obtained a shift of the median $(1 - b_{\rm SZ})$ corresponding to 0.5σ. As the tests described above showed shifts in the median $b_{\rm SZ}$ that are within 1σ of the constraint presented in Section 4.3.1, we can conclude that our results are robust with respect to the investigated modelling choices. Lastly, Salvati et al. (2020) showed that the impact of the halo mass function calibration has a subdominant impact on current cluster count analyses. It will become a relevant source of systematic errors in upcoming surveys, for instance those from Euclid (Laureijs et al. 2011) and the Large Synoptic Survey Telescope (LSST Science Collaboration et al. 2009), because of the larger cluster statistics they will provide. As cluster clustering has a lower constraining power compared to counts, we expect the mass function calibration to provide a negligible impact on the $b_{\rm SZ}$ constraints derived in this work. The same is expected for the halo bias calibration.

4.3.3 Constraints on cosmological parameters

To further investigate the consistency of our modelling choices with those adopted by Planck Collaboration XXIV (2016), we performed a cosmological analysis aiming at constraining σ_8 and $\Omega_{\rm m}$ simultaneously by assuming the same Gaussian priors on $b_{\rm SZ}$ as were considered by Planck Collaboration XXIV (2016). Specifically, we assumed large flat priors for σ_8 and $\Omega_{\rm m}$, while for the other cosmological parameters, we assumed the same values from Planck18 that were used in the previous section. We also assumed the same Gaussian priors as adopted in Section 4.3.1 on the scaling relation parameters, namely Y_*, α, β, and $\sigma_{\ln Y}$. We found that σ_8 is not constrained through this analysis, while we found $\Omega_{\rm m} = 0.28^{+0.05}_{-0.04}$ with the WtG $b_{\rm SZ}$ prior, $\Omega_{\rm m} = 0.28^{+0.04}_{-0.03}$ with the CCCP prior, and $\Omega_{\rm m} = 0.27^{+0.04}_{-0.03}$ with the CMB lensing prior (see Fig. 4.4). These results are fully consistent and competitive, in terms of uncertainties, with those derived by Planck Collaboration XXIV (2016). Similar uncertainties on $\Omega_{\rm m}$ were

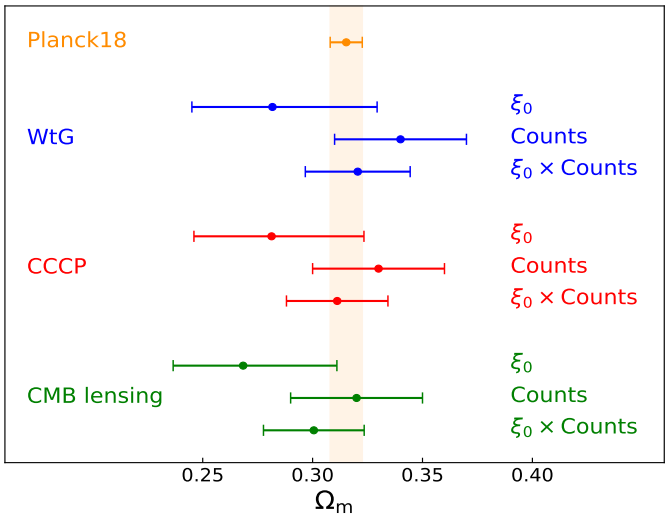

Figure 4.4: Comparison of the results on Ω_m obtained by assuming flat priors on Ω_m and σ_8, along with external priors on b_{sz}, namely WtG (blue; von der Linden et al. 2014), CCCP (red; Hoekstra et al. 2015) and CMB lensing (green; Planck Collaboration XXIV 2016). For each b_{sz} prior, the result at the top was derived from the cluster clustering measurements presented in this Chapter, the middle result refers to the cluster counts analysis by Planck Collaboration XXIV (2016), and the bottom result represents the combination of Planck cluster clustering and counts. The constraint from Planck18 is shown in orange.

obtained by Marulli et al. (2018) and Lindholm et al. (2021), who modelled the 2PCF of 182 and 1892 X-ray selected galaxy clusters, respectively. In addition, Marulli et al. (2018) found that σ_8 could not be accurately constrained, in agreement with our results. In general, recent analyses of cluster clustering showed that the constraints on Ω_m are significantly more robust than those derived on σ_8 (Lindholm et al. 2021; Lesci et al. 2022b). In addition, we note that significant changes in the value of b_{sz} do not imply significant variations in the Ω_m posteriors, similar to what was found by Planck Collaboration XXIV (2016).

We also derived an estimate of Ω_m from the combination of cluster clustering and counts by assuming them to be statistically independent. Hurier & Lacasa (2017) showed that cluster counts and thermal SZ power spectrum are not significantly correlated. The thermal SZ power spectrum mainly depends on massive haloes, while for the number counts, the main contribution comes from lower-mass haloes. The differences between the two galaxy cluster populations imply a weak correlation of these probes. As the clustering of SZ selected clusters encloses the same information contained in the SZ power spectrum, we expect a similar behaviour for its correlation with counts. With respect to the analysis based on counts only, we found that the uncertainty on Ω_m is reduced by a factor of $\sim 25\%$-30% if clustering is included in the likelihood. This confirms the importance of including cluster clustering in cosmological analyses (see also Sartoris et al. 2016; Euclid Collaboration: Fumagalli et al. 2022; Garrel et al. 2022) in order to fully exploit the cluster statistics information.

4.4 Summary

In the work by Lesci et al. (2023), presented in this Chapter, we analysed the 3D 2PCF monopole of the galaxy clusters detected by Planck Collaboration XXVII (2016), focusing on the estimate of the Planck mass bias, b_{sz}. Following Planck Collaboration XXIV (2016), we based our analysis on the cosmological sample consisting of detections by the MMF3 matched

filter (Melin et al. 2006, 2012), considering clusters with a confirmed counterpart in external data sets and having an assigned redshift estimate, with a redshift limit $z \leq 0.8$, for a total of 920 clusters. Differently from Planck Collaboration XXIV (2016), we did not apply any cut in S/N to the sample. This does not imply any potential problems due to the reliability of the selection function at low S/N because our model does not rely on assumptions on the sample completeness.

By analysing the 2PCF in the redshift bins $z < 0.2$ and $z \in [0.2, 0.8]$ within the cluster-centric radial range $r \in [10, 150]$ h^{-1}Mpc, we derived $(1 - b_{\rm SZ}) = 0.62^{+0.14}_{-0.11}$. This result fully agrees with what was found by Planck Collaboration VI (2020) by combining primary CMB likelihood and Planck cluster counts. Thus, we confirmed that Planck cluster statistics provides values of $b_{\rm SZ}$ that are lower than what was predicted by numerical simulations (Nagai et al. 2007; Piffaretti & Valdarnini 2008; Meneghetti et al. 2010; Rasia et al. 2012; Le Brun et al. 2017; Henson et al. 2017; Gianfagna et al. 2023). As redshift errors are not quoted in Planck data products, we also included the power spectrum damping due to redshift uncertainties by means of a free parameter representing the typical redshift error, namely σ_z. Thus, we simultaneously calibrated σ_z and $b_{\rm SZ}$, finding no significant changes in $b_{\rm SZ}$ and $\sigma_z = 0.003^{+0.002}_{-0.002}$, which is in line with the fact that most of the cluster redshifts are spectroscopic. In addition, from the analysis of the Planck union catalogue of clusters, we showed that our result does not depend on the adopted cluster detection algorithm. We also found that a redshift evolution of $b_{\rm SZ}$ is not necessary to describe our clustering measurements.

By adopting priors on $b_{\rm SZ}$ from external data sets, we found results on $\Omega_{\rm m}$ that fully agree and are competitive, in terms of uncertainties, with those derived from cluster counts by Planck Collaboration XXIV (2016), while σ_8 is not constrained. By assuming cluster clustering and counts to be statistically independent, we found that their combination provides a reduction of up to $\sim 30\%$ in the $\Omega_{\rm m}$ uncertainty derived from counts. Future stage IV CMB experiments (Abazajian et al. 2016) will detect about 10^5 galaxy clusters through the SZ effect, which will significantly enhance the cluster statistical analyses. This will improve the calibration of the hydrostatic mass bias from cluster clustering, and might shed light on the degeneracy between σ_8 and mass bias. In fact, this degeneracy cannot be investigated with current clustering data because σ_8 is not constrained, as we detailed in Section 4.3.3.

Chapter 5

AMICO clusters in KiDS-DR3: counts and clustering

As discussed in Chapt. 3, cluster counts provide the primary cosmological probe among the ones given by the statistics of these tracers. The constraints from cluster counts can be improved by at least 20% when cluster clustering is included in the analysis (see, e.g., Mana et al. 2013; Sartoris et al. 2016; Fumagalli et al. 2023), as we showed also in Chapt. 4. Although it is possible to predict with great accuracy the abundance of dark matter haloes as a function of mass for a given cosmological model, the cluster masses cannot be easily derived from observational data. Currently, the most reliable mass measurements are provided by weak gravitational lensing (see, e.g., Bardeau et al. 2007; Okabe et al. 2010b; Hoekstra et al. 2012; Melchior et al. 2015; Schrabback et al. 2018; Stern et al. 2019), which accounts for both the dark and baryonic matter components. As opposed to the other methods to estimate cluster masses based on the properties of the gas and member galaxies, such as the ones exploiting X-ray emission, galaxy velocity dispersion, or the Sunyaev-Zeldovich effect on the CMB, the gravitational lensing method does not rely on any assumption pertaining to the dynamical state of the cluster. However, weak-lensing mass measurements of individual galaxy clusters are only feasible if the S/N of the shear profiles is sufficiently high: this requires either a massive structure or deep observations. Thus, in cosmological studies of cluster statistics, it is often necessary to stack the weak-lensing signal produced by a set of objects with similar properties, from which a mean value of their mass is estimated. These mean mass values can be linked to a direct observable or mass proxy, which can be used to define a mass-observable scaling relation.

In addition to the mass scaling relation, it is crucial to properly account for the purity and completeness of the cluster catalogue. As we detailed in Chapts. 3 and 4, cluster clustering theoretical models can be conveniently expressed as independent of the sample completeness estimates. This makes cluster clustering a purity-limited probe. On the other hand, cluster counts strongly depend on the sample completeness, whose effect must be included in the model predictions, as we shall discuss in the following.

In this Chapter, we present the work by Lesci et al. (2022a) on cluster counts and by Lesci et al. (2022b) on cluster clustering, both based on the Data Release 3 of the Kilo Degree Survey (KiDS-DR3; de Jong et al. 2017), which is a photometric survey of galaxies. The cluster sample, built up through the use of the Adaptive Matched Identifier of Clustered Objects (AMICO; Bellagamba et al. 2018), was presented in Maturi et al. (2019). In addition, both cluster counts and clustering analyses are based on the weak-lensing measurements provided by Bellagamba et al. (2019). Assuming a flat ΛCDM framework, we constrain the cosmological parameters σ_8, Ω_m, and $S_8 \equiv \sigma_8(\Omega_\mathrm{m}/0.3)^{0.5}$ from cluster counts and clustering

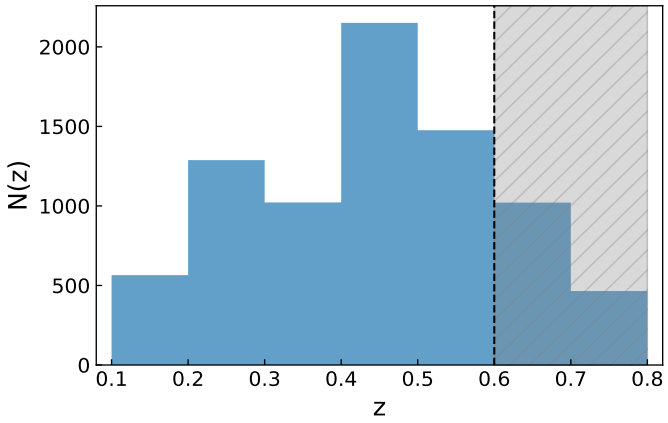

Figure 5.1: Distribution of the AMICO KiDS-DR3 clusters as a function of redshift. Objects with $z > 0.6$, not considered in the cluster counts and clustering analyses, are covered by the shaded grey area.

measurements, as well as the mass-richness scaling relation parameters, including its intrinsic scatter, $\sigma_{\rm intr}$. The cosmological analyses presented in this Chapter are based on the CBL numerical libraries (Marulli et al. 2016) v5.4.

This Chapter is organised as follows. In Sect. 5.1 we present the galaxy cluster catalogue based on KiDS-DR3 data, along with the weak-lensing cluster mass measurements. The sample selections and the results of the cosmological analyses of counts and clustering are detailed in Sects. 5.2 and 5.3, respectively.

5.1 Data set

5.1.1 AMICO KiDS-DR3 cluster catalogue

The catalogue of galaxy clusters the works by Lesci et al. (2022a) and Lesci et al. (2022b) are based on, named AMICO KiDS-DR3 (Maturi et al. 2019), is derived from the third data release of KiDS (de Jong et al. 2017). Such observations were carried out with the OmegaCAM wide-field imager (Kuijken 2011) mounted at the VLT Survey Telescope, which is a 2.6 m telescope situated at the Paranal Observatory. The 2 arcsec aperture photometry in u, g, r, and i bands is provided, as well as the photometric redshifts for all galaxies down to the 5σ limiting magnitudes of 24.3, 25.1, 24.9, and 23.8 for the aforementioned four bands, respectively. To construct the galaxy cluster catalogue, Maturi et al. (2019) considered only the galaxies with a magnitude $r < 24$, for a total of about 32 million galaxies. For the weak-lensing analysis, instead, no limits in magnitude have been imposed for the lensed sources in order to exploit the whole data set available. Such data set provides the shear measurements for about 15 million galaxies (Hildebrandt et al. 2017).

Galaxy clusters have been detected thanks to the application of the AMICO algorithm (Bellagamba et al. 2018), which identifies galaxy overdensities by exploiting a linear matched optimal filter. In particular, the detection process adopted for this study solely relies on the angular coordinates, magnitudes, and photometric redshifts (photo-zs from now on) of galaxies. Unlike other algorithms used in literature for cluster identification, AMICO does not use any direct information coming from colours, as for example the so-called red sequence. For this reason, AMICO is also expected to be accurate at higher redshifts, where the red sequence may not be prominent yet. The excellent performances of AMICO have been

recently confirmed by the analysis made in Euclid Collaboration et al. (2019), where the purity and completeness of the cluster catalogues extracted by applying six different algorithms on realistic mock catalogues reproducing the expected characteristics of the future *Euclid* photometric surveys (Laureijs et al. 2011; Euclid Collaboration: Scaramella et al. 2022) have been compared. As the result of this challenge, AMICO is one of the two algorithms for cluster identification that has been officially adopted by the *Euclid* mission.

The KiDS-DR3 sample covers a total area of 438 deg^2, but all the galaxies located in the regions affected by image artefacts, or falling in the secondary and tertiary halo masks used for the weak-lensing analysis have been rejected (de Jong et al. 2015). This yields a final effective area of 377 deg^2, containing all the cluster detections with S/N$>$ 3.5 and within the redshift range $z \in [0.1, 0.8]$, for a total of 7988 objects. Due to the low S/N of the shear profiles for $z > 0.6$, which is not sufficient to perform a stacked weak-lensing analysis, we decided to exclude the redshift bin $z \in [0.6, 0.8]$ from the analysis. In Fig. 5.1 we show the redshift distribution of the AMICO KiDS-DR3 cluster sample.

5.1.2 Mass proxy

As we shall detail in the following, we exploit the cluster shear signal through a stacked weak-lensing analysis to estimate the mean cluster masses in bins of intrinsic richness and redshift. The intrinsic richness, λ^*, is defined as follows (Maturi et al. 2019):

$$\lambda^*_j := \sum_{i=1}^{N_{\rm gal}} P_i(j) \quad \text{with} \quad \begin{cases} m_i < m^*(z_j) + 1.5 \\ R_i(j) < R_{\rm max}(z_j) \end{cases} , \tag{5.1}$$

where $P_i(j)$ is the probability assigned by AMICO to the i-th galaxy of being a member of a given detection j. The intrinsic richness thus represents the sum of the membership probabilities, that is the weighted number of visible galaxies belonging to a detection, under the conditions given by Eq. (5.1). The sum of the membership probabilities is an excellent estimator of the true number of member galaxies, as shown in Bellagamba et al. (2018) by running the AMICO algorithm on mock catalogues (see Fig. 8 in the reference). In Eq. (5.1), z_j is the redshift of the j-th detected cluster, m_i is the magnitude of the i-th galaxy, and R_i corresponds to the distance of the i-th galaxy from the centre of the cluster. The parameter $R_{\rm max}(z_j)$ represents the radius enclosing a mass $M_{200} = 10^{14} M_\odot/h$, such that the corresponding mean density is 200 times the critical density of the Universe at the given redshift z_j. In the following analysis, indeed, we consider the masses evaluated as M_{200}. Lastly, m^* is the typical magnitude of the Schechter function in the cluster model assumed in the AMICO algorithm. We use the term intrinsic richness as opposed to apparent richness, defined in Maturi et al. (2019). In particular, since the threshold in absolute magnitude is always lower than the survey limit, thanks to its redshift dependence, λ^* does not depend on the survey limit. Conversely, the apparent richness is a quantity that includes all visible galaxies and is therefore related to how a cluster is observed given a certain apparent magnitude limit.

5.1.3 Weak-lensing masses

To estimate the mean masses of the observed galaxy clusters, we followed the same stacked weak-lensing procedure described in Bellagamba et al. (2019), based on KiDS-DR3 data. The clusters selected for the weak-lensing analysis lie in the redshift range $z \in [0.1, 0.6]$, over an effective area of 360.3 deg^2. This area is slightly smaller compared to that considered for the counts' analysis since it was derived from the masking described in Hildebrandt et al. (2017). Despite the availability of galaxy clusters up to $z = 0.8$ in the AMICO KiDS-DR3 catalogue,

Table 5.1: Cluster binning used for the weak-lensing analysis.

z range	$z_{\rm eff}$	λ^* range	$\lambda^*_{\rm eff}$	$\log \bar{M}_{200}$	$N_{\rm cl}$	$z_{\rm s,eff}$
[0.10, 0.30]	0.189 ± 0.001	[0, 15]	10.20 ± 0.09	-0.73 ± 0.07	1246	0.849 ± 0.002
[0.10, 0.30]	0.212 ± 0.002	[15, 25]	18.88 ± 0.12	-0.38 ± 0.07	684	0.867 ± 0.002
[0.10, 0.30]	0.222 ± 0.004	[25, 35]	29.02 ± 0.21	0.05 ± 0.07	209	0.879 ± 0.002
[0.10, 0.30]	0.228 ± 0.007	[35, 45]	39.75 ± 0.32	0.32 ± 0.06	82	0.877 ± 0.005
[0.10, 0.30]	0.222 ± 0.008	[45, 150]	56.59 ± 2.20	0.54 ± 0.06	44	0.890 ± 0.013
[0.30, 0.45]	0.374 ± 0.001	[0, 20]	15.10 ± 0.11	-0.41 ± 0.08	1113	0.948 ± 0.003
[0.30, 0.45]	0.387 ± 0.002	[20, 30]	24.08 ± 0.11	-0.07 ± 0.07	767	0.944 ± 0.004
[0.30, 0.45]	0.389 ± 0.002	[30, 45]	35.91 ± 0.27	0.21 ± 0.06	320	0.941 ± 0.005
[0.30, 0.45]	0.390 ± 0.005	[45, 60]	50.88 ± 0.50	0.41 ± 0.08	87	0.950 ± 0.015
[0.30, 0.45]	0.379 ± 0.006	[60, 150]	73.60 ± 2.09	0.68 ± 0.07	45	0.946 ± 0.012
[0.45, 0.60]	0.498 ± 0.001	[0, 25]	19.71 ± 0.11	-0.33 ± 0.09	1108	0.958 ± 0.001
[0.45, 0.60]	0.514 ± 0.002	[25, 35]	29.23 ± 0.12	-0.07 ± 0.07	761	0.961 ± 0.004
[0.45, 0.60]	0.523 ± 0.003	[35, 45]	39.25 ± 0.18	0.21 ± 0.07	299	0.961 ± 0.006
[0.45, 0.60]	0.513 ± 0.004	[45, 150]	55.12 ± 0.76	0.36 ± 0.07	197	0.960 ± 0.004

Notes. The computation of $z_{\rm eff}$ and $\lambda^*_{\rm eff}$ and their uncertainties are described in Bellagamba et al. (2019). For the logarithm of the measured mean masses, $\log \bar{M}_{200}$, expressed in units of $(10^{14}\, h^{-1}{\rm M}_\odot)$, we quote the mean and the standard deviation of the posterior probability distribution. We note that $N_{\rm cl}$ is the number of clusters in the bin. In the last column, $z_{\rm s,eff}$ is the effective redshift of the lensed sources, obtained by following the procedure described in Giocoli et al. (2021). Quoted masses refer to a flat ΛCDM model with $\Omega_{\rm m} = 0.3$ and $h = 0.7$.

the S/N of the stacked shear profiles is too low to perform the stacking for $z > 0.6$. Therefore we base our analysis on the redshift bins $z \in [0.1, 0.3]$, $z \in [0.3, 0.45]$, and $z \in [0.45, 0.6]$, deriving the estimated mean masses in a flat ΛCDM cosmology with $\Omega_{\rm m} = 0.3$ and $h = 0.7$.

With a Markov chain Monte Carlo (MCMC) analysis, we sampled the posterior distributions of the base 10 logarithm of the estimated mean cluster masses, $\log \bar{M}_{200}$, in 14 bins of intrinsic richness and redshift, considering $\lambda^* \geq 0$ for a total of 6962 objects (see Table 5.1). To test the reliability of this cut in λ^*, we also performed the cosmological analysis by imposing $\lambda^* \geq 20$ for the weak-lensing masses, deriving results fully in agreement with those obtained by assuming $\lambda^* \geq 0$, as detailed in Section 5.2.3. As we discuss in Section 5.2.2, we account for the systematic errors affecting the weak lensing mass estimates by relying on the results found in Bellagamba et al. (2019). Specifically, we consider the systematics due to background selection, photo-zs, and shear measurements, affecting the measured stacked cluster profiles. Such errors are then propagated into the mass estimates. The sum in quadrature of such contributions to systematic errors, along with those produced by the halo model, orientation, and projections, is equal to 7.6%. The description of the modelling, including a more extensive discussion on the statistical and systematic uncertainties, is detailed in Bellagamba et al. (2019).

The $\log \bar{M}_{200}$ posteriors are marginalised over the other parameters entering the modelling, that is the concentration parameter, c_{200}, the fraction of haloes belonging to the miscentred population, $f_{\rm off}$, and the rms of the distribution of the halo misplacement on the plane of the sky, $\sigma_{\rm off}$. In particular, we derived the posteriors for c_{200}, $f_{\rm off}$, and $\sigma_{\rm off}$ in each bin, assuming the following flat priors: $c_{200} \in [1, 20]$, $f_{\rm off} \in [0, 0.5]$, and $\sigma_{\rm off} \in [0\,{\rm Mpc}\,h^{-1}, 0.5\,{\rm Mpc}\,h^{-1}]$. Such parameters are not constrained by the data, that is their posteriors are statistically

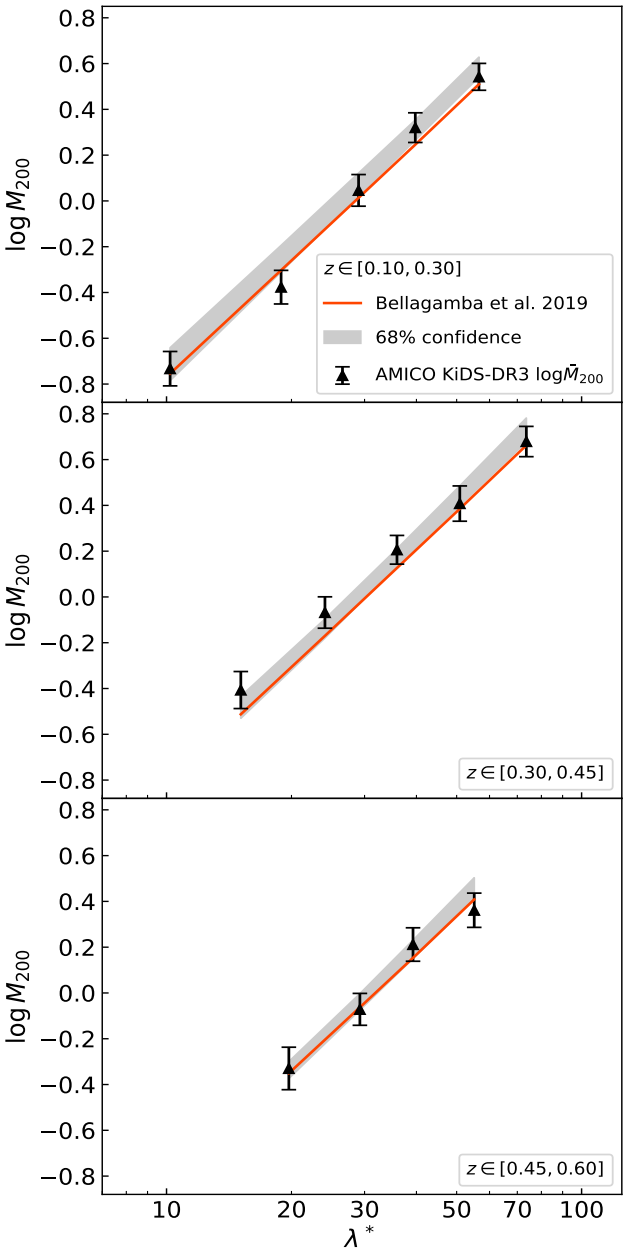

Figure 5.2: Logarithm of the masses in units of $(10^{14}\,h^{-1}\mathrm{M}_\odot)$, $\log \bar{M}_{200}$, from the AMICO KiDS-DR3 cluster catalogue as a function of the intrinsic richness λ^*, in the redshift bins $z \in [0.10, 0.30]$, $z \in [0.30, 0.45]$, and $z \in [0.45, 0.60]$, from top to bottom. The black triangles represent the mean values of $\log \bar{M}_{200}$, given by the mean of the marginalised posterior obtained in the weak-lensing analysis, while the error bars are given by 1σ of the posterior distribution. The orange lines represent the median scaling relation obtained by modelling only $\log \bar{M}_{200}$, following the procedure described in Bellagamba et al. (2019). The grey bands represent the 68% confidence level derived from the multivariate posterior of all the free parameters considered in the cosmological analysis described in Section 5.2.2.

consistent with the priors. For what concerns the miscentring parameters, they are related to the possible difference between the centre defined by the AMICO algorithm using the galaxy overdensities and the mass centre related to the weak-lensing signal. In fact, the uncertainty due to the use of a grid in AMICO indeed only impacts small scales not used in this analysis

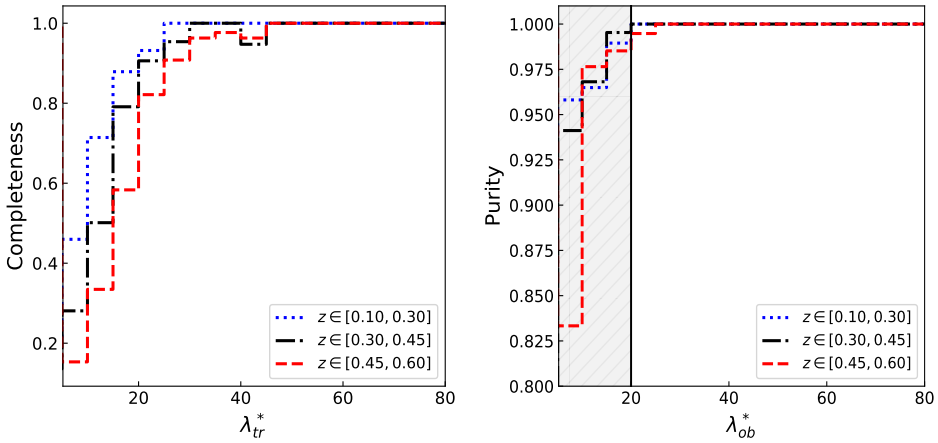

Figure 5.3: Completeness (left panel) and purity (right panel) of the AMICO KiDS-DR3 cluster catalogue as a function of the redshift, z, and of the intrinsic richness, λ^*. The completeness is a function of the true λ^*, i.e. $\lambda^*_{\rm tr}$, while the purity is defined as a function of the observed intrinsic richness, $\lambda^*_{\rm ob}$. The shaded area in the right panel highlights the bins excluded from the number counts' analysis, namely for $\lambda^*_{\rm ob} < 20$.

and thus they can be neglected.

The logarithm of the estimated mean mass values for different bins of intrinsic richness and redshift are listed in Table 5.1. In Fig. 5.2 we show the median value of the mass-intrinsic richness scaling relation, described in Section 5.2.2, obtained by only performing the modelling of the weak lensing masses as in Bellagamba et al. (2019), along with the 68% confidence level obtained from the analysis of cluster counts and weak lensing masses, detailed in Section 5.2.2. It turns out that the $\log M_{200} - \log \lambda^*$ relation is reasonably linear and with an intrinsic scatter of 0.1, as we discuss in Section 5.2.3, indicating the reliability of λ^* as a mass proxy.

5.1.4 Selection function

In order to estimate the selection function of the AMICO KiDS-DR3 cluster catalogue, we make use of the mock catalogue described in Maturi et al. (2019). The construction of the mock clusters is based on the original galaxy data set, thus all the properties of the survey are properly taken into account, such as masks, photo-z uncertainties, and the clustering of galaxies. In this way the assumptions necessary to build up the mock catalogue are minimised. In particular, regarding the photo-z uncertainties, the galaxies are drawn from the survey sample and selected from bins of richness and redshift, using a Monte Carlo sampling based on the cluster membership probability. The probability that each galaxy is included in a given redshift bin is driven by its own photo-z probability distribution function, which includes the contribution of the photometric noise. In this way, the selection of the simulated cluster member galaxies mimics the real uncertainties of the photometric redshifts in the photometric catalogue. Then, to derive the selection function, the AMICO code was run on the mock catalogue, consisting of 9018 clusters distributed over a total area of 189 deg^2. Only the detections with an S/N$>$ 3.5 are considered, with this being the threshold applied to the real dataset.

In Fig. 5.3 we show the purity and completeness of the dataset, which define the selection function. The completeness is defined as the number of detections correctly identified as clusters over the total number of mock clusters as a function of redshift and intrinsic richness. Thus it provides a measure of how many objects are lost in the detection procedure. On the

other hand, the purity is a measure of the contamination level of the cluster sample. It is defined as the fraction of detections matching the clusters in the mock catalogue, over the total number of detections, in a given bin of redshift and intrinsic richness. As shown in Fig. 5.3, it turns out that the catalogue is highly pure, with a purity approaching 100% over the whole redshift range for $\lambda^* \geq 20$.

In order to account for the selection function in the modelling of cluster counts, we built a new dataset by applying the purity and the completeness to the real cluster catalogue. This dataset will be used to derive the multiplicative weights that will be considered in the cluster counts' model, as we detail in the following. Since we define the purity as a function of the observed intrinsic richness, $\lambda^*_{\rm ob}$, we assigned each object in the real catalogue to a bin of observed intrinsic richness, in which we computed the purity. Subsequently, we extracted a uniform random number between 0 and 1, and if it is lower than the purity corresponding to the bin, the object is considered in the aforementioned new dataset. Otherwise, it is rejected. In this way, the final sample will statistically take the effects of impurities into account. On the other hand, since the completeness is defined in bins of true intrinsic richness, $\lambda^*_{\rm tr}$, it is required to implement a method that assigns a value of completeness to an observed value of intrinsic richness. For this purpose, we derived several probability distributions from the mock catalogue describing the probability to obtain a true value of λ^*, given a range of observed intrinsic richness defined by $\lambda^{*\,\rm low}_{\rm ob}$ and $\lambda^{*\,\rm up}_{\rm ob}$, namely $P(\lambda^*_{\rm tr}|\lambda^{*\,\rm low}_{\rm ob}, \lambda^{*\,\rm up}_{\rm ob})$. We find that these distributions are reasonably Gaussian. Then, given a galaxy cluster in our dataset with a value of $\lambda^*_{\rm ob}$ in a given range, we performed a Gaussian Monte Carlo extraction from $P(\lambda^*_{\rm tr}|\lambda^{*\,\rm low}_{\rm ob}, \lambda^{*\,\rm up}_{\rm ob})$ through which we obtain a value of $\lambda^*_{\rm tr}$. Given the extracted true value of intrinsic richness, we assigned a completeness value to the considered object. Having this new catalogue corrected for the purity and the completeness, we constructed a weight factor defined as the ratio between the uncorrected counts and the corrected ones for bins in intrinsic richness (denoted by $\Delta\lambda^*_{{\rm ob},i}$) and redshift (labelled as $\Delta z_{{\rm ob},j}$). These weight factors, $w(\Delta\lambda^*_{{\rm ob},i}, \Delta z_{{\rm ob},j})$, will be used to weigh the cluster count model as described in Section 5.2.2. The value of $w(\Delta\lambda^*_{{\rm ob},i}, \Delta z_{{\rm ob},j})$ in the first bins of intrinsic richness amounts to 0.87, 0.76, and 0.64 in the redshift bins $z \in [0.10, 0.30]$, $z \in [0.30, 0.45]$, and $z \in [0.45, 0.60]$, respectively, while we derived no correction for the other bins (i.e. in these bins the weights are equal to 1).

5.2 AMICO KiDS-DR3 cluster counts

In this section, we detail the analysis of AMICO KiDS-DR3 galaxy cluster counts presented in Lesci et al. (2022a), based on the data set discussed in Sect. 5.1. In Sect. 5.2.1 we detail the sample selection and the measurements of cluster counts. In Sect. 5.2.2 we detail the theoretical expected values of cluster weak-lensing masses and counts, along with the likelihood function adopted for the Bayesian analysis. In Sect. 5.2.3 we present our constraints on cosmological parameters and on the mass-richness scaling relation, while in Sect. 5.2.4 we discuss our results.

5.2.1 Sample selection and measurements

We set $\lambda^* = 20$ as the threshold for the cluster count analysis in order to exclude the bins affected by detection impurities and severe incompleteness. Thus the final sample of galaxy clusters considered for the analysis contains 3652 objects, with $\lambda^* \geq 20$, and in the redshift bins $z \in [0.1, 0.3]$, $z \in [0.3, 0.45]$, and $z \in [0.45, 0.6]$. With regard to the binning in intrinsic richness, we adopt four logarithmically spaced bins in the range $\lambda^* \in [20, 137]$ for each redshift

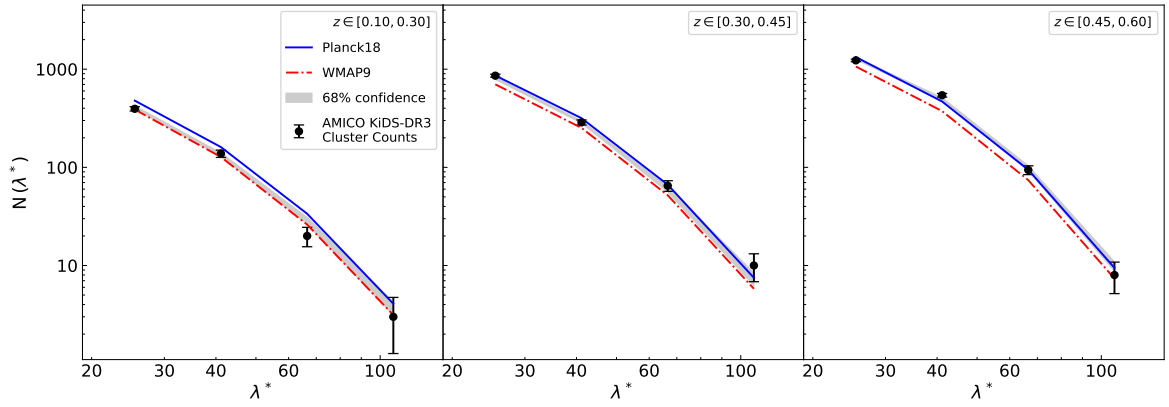

Figure 5.4: Number counts from the AMICO KiDS-DR3 cluster catalogue as a function of the intrinsic richness λ^*, in the redshift bins $z \in [0.10, 0.30]$, $z \in [0.30, 0.45]$, and $z \in [0.45, 0.60]$, from left to right. The black dots represent the counts directly retrieved from the catalogue, where the error bars are given by the Poissonian noise. The solid blue lines represent the model computed by assuming the cosmological parameters obtained by Planck Collaboration VI (2020, Table 2, TT, TE, and EE+lowE), while the red dashed lines show the results based on the WMAP cosmological parameters (Hinshaw et al. 2013, Table 3, WMAP-only Nine-year). Both in the Planck and WMAP cases, the scaling relation parameters and the intrinsic scatter have been fixed to the median values listed in Table 5.2, retrieved from the modelling. The grey bands represent the 68% confidence level derived from the multivariate posterior of all the free parameters considered in the cosmological analysis.

bin. To test the robustness of our results with respect to this binning choice, we repeated the analysis assuming different numbers of λ^* bins and obtained negligible differences in the final results, that is far below the 1σ of the posterior distributions, with values of reduced χ^2 always consistent with 1. The measured counts of the AMICO KiDS-DR3 clusters are shown in Fig. 5.4, along with the 68% confidence level derived in Section 5.2.3.

5.2.2 Modelling

Model for the weak-lensing masses

We modelled the scaling relation between the estimated cluster mean masses and the intrinsic richnesses using the following functional form:

$$\log \frac{M_{200}}{10^{14} M_\odot / h} = \alpha + \beta \int_0^\infty \mathrm{d}\lambda^* \; P(\lambda^* | \lambda^*_{\rm eff}) \; \log \frac{\lambda^*}{\lambda^*_{\rm piv}} + \\ + \gamma \int_0^\infty \mathrm{d}z \; P(z | z_{\rm eff}) \; \log \frac{E(z)}{E(z_{\rm piv})} \; , \tag{5.2}$$

where $E(z) \equiv H(z)/H_0$, while $z_{\rm eff}$ and $\lambda^*_{\rm eff}$ are the lensing-weighted effective redshift and richness, respectively, whose computation is described in Bellagamba et al. (2019). The probability distributions $P(\lambda^*|\lambda^*_{\rm eff})$ and $P(z|z_{\rm eff})$ are assumed to be Gaussian, with a mean equal to the values of $\lambda^*_{\rm eff}$ and $z_{\rm eff}$ listed in Table 5.1, and an rms given by the uncertainties on $\lambda^*_{\rm eff}$ and $z_{\rm eff}$, respectively. The last term in Eq. (5.2) accounts for deviations in the redshift evolution from what is predicted in the self-similar growth scenario (Sereno & Ettori 2015). Following Bellagamba et al. (2019), we set $\lambda^*_{\rm piv} = 30$ and $z_{\rm piv} = 0.35$. In Eq. (5.2) the observables are the estimated mean mass values, $\log \bar{M}_{200}$, shown in Table 5.1, along with the effective values of redshift, $z_{\rm eff}$, and of intrinsic richness, $\lambda_{\rm eff}$, in the given bin. Furthermore,

since $\log \bar{M}_{200}$ depends on cosmological parameters, we adopted the rescaling described in Sereno (2015), that is

$$\bar{M}_{200,\,\mathrm{new}} = \bar{M}_{200,\,\mathrm{ref}} \frac{\left[D_d^{-\frac{3\delta\gamma}{2-\delta\gamma}} \left(\frac{D_{ds}}{D_s}\right)^{-\frac{3}{2-\delta\gamma}} H(z)^{-\frac{1+\delta\gamma}{1-\delta\gamma/2}} \right]_{\mathrm{new}}}{\left[D_d^{-\frac{3\delta\gamma}{2-\delta\gamma}} \left(\frac{D_{ds}}{D_s}\right)^{-\frac{3}{2-\delta\gamma}} H(z)^{-\frac{1+\delta\gamma}{1-\delta\gamma/2}} \right]_{\mathrm{ref}}}, \tag{5.3}$$

where ref indicates the assumed reference cosmology, that is $\Omega_{\mathrm{m}} = 0.3$ and $h = 0.7$, while the subscript new refers to that of the test. We set the slope to $\delta\gamma = 0$, corresponding to the case of a singular isothermal profile, with this being a good approximation in general (as discussed in Sereno 2015). For example, assuming $M_{200} \simeq 10^{15}\ M_{\odot}$ and $c_{200} \simeq 3$, $\delta\gamma \simeq -0.1$ is obtained. Thus we varied $\delta\gamma$ in the reasonable range $[-0.2, 0.2]$ and verified that this does not have an impact on the final results. The terms D_{s}, D_{d}, and D_{ds} are the source's, the lens', and the lens-source's angular diameter distances, respectively. In D_{d}, D_{ds}, and in the Hubble parameter, $H(z)$, we assumed the effective redshift values, z_{eff}, listed in Table 5.1. With regard to the redshifts of the sources, we considered the lensed source effective redshifts, $z_{\mathrm{s,eff}}$, listed in Table 5.1, obtained by following the procedure described in Giocoli et al. (2021). In particular, we obtained $z_{\mathrm{s,eff}}$ by weighting the redshift of each source by the corresponding source density for each derived radial bin for each cluster. We then considered the mean value of source redshift in bins of cluster richness per redshift. We verified that we can neglect the uncertainty on $z_{\mathrm{s,eff}}$ in our analysis. In such mass rescaling, the relative uncertainty on $\log \bar{M}_{200,\,\mathrm{new}}$ is constant, corresponding to the relative errors on $\log \bar{M}_{200,\,\mathrm{ref}}$.

Model for the cluster counts

The specific characteristics of the data set must be included in the model and in the covariance matrix of the likelihood function. We describe the expectation value of the counts in a given bin of intrinsic richness, $\Delta\lambda^*_{\mathrm{ob},i}$, and of observed redshift, $\Delta z_{\mathrm{ob},j}$, as

$$\begin{aligned}\langle N(\Delta\lambda^*_{\mathrm{ob},i}, \Delta z_{\mathrm{ob},j})\rangle = w(\Delta\lambda^*_{\mathrm{ob},i}, \Delta z_{\mathrm{ob},j})\ \Omega \int_0^\infty \mathrm{d}z_{\mathrm{tr}}\ \frac{\mathrm{d}V}{\mathrm{d}z_{\mathrm{tr}}\mathrm{d}\Omega} \times \\ \times \int_0^\infty \mathrm{d}M_{\mathrm{tr}}\ \frac{\mathrm{d}n(M_{\mathrm{tr}}, z_{\mathrm{tr}})}{\mathrm{d}M_{\mathrm{tr}}} \int_0^\infty \mathrm{d}\lambda^*_{\mathrm{tr}}\ P(\lambda^*_{\mathrm{tr}}|M_{\mathrm{tr}}, z_{\mathrm{tr}}) \times \\ \times \int_{\Delta z_{\mathrm{ob},j}} \mathrm{d}z_{\mathrm{ob}}\ P(z_{\mathrm{ob}}|z_{\mathrm{tr,\,corr}}) \int_{\Delta\lambda^*_{\mathrm{ob},i}} \mathrm{d}\lambda^*_{\mathrm{ob}}\ P(\lambda^*_{\mathrm{ob}}|\lambda^*_{\mathrm{tr}})\,,\end{aligned} \tag{5.4}$$

where z_{tr} is the true redshift, V is the co-moving volume, Ω is the survey effective area, M_{tr} is the true mass, and $\mathrm{d}n(M_{\mathrm{tr}}, z_{\mathrm{tr}})/\mathrm{d}M_{\mathrm{tr}}$ is the mass function, for which the model by Tinker et al. (2008) is assumed. The term $w(\Delta\lambda^*_{\mathrm{ob},i}, \Delta z_{\mathrm{ob},j})$ is the weight factor described in Section 5.1.4, accounting for the purity and completeness of the sample. The probability distribution $P(z_{\mathrm{ob}}|z_{\mathrm{tr,\,corr}})$, assessed through the mock catalogue described in Section 5.1.4, is a Gaussian accounting for the uncertainties on the redshifts. The mean of such distribution, $z_{\mathrm{tr,\,corr}}$, is the true redshift corrected by the redshift bias, and it is expressed as

$$z_{\mathrm{tr,\,corr}} = z_{\mathrm{tr}} + \Delta z_{\mathrm{bias}}\,(1 + z_{\mathrm{tr}})\,, \tag{5.5}$$

where $\Delta z_{\mathrm{bias}}\,(1+z_{\mathrm{tr}})$ is the redshift bias term discussed in Maturi et al. (2019), with $\Delta z_{\mathrm{bias}} = 0.02$. In particular, this bias corresponds to what was found in de Jong et al. (2017) by

comparing the KiDS photo-zs to the GAMA spectroscopic redshifts (see their Table 8). In order to assess the impact of its uncertainty, we included $\Delta z_{\rm bias}$ as a free parameter in the model, assuming a Gaussian prior with a mean equal to 0.02 and an rms equal to 0.02, which is similar to the rms of the sample and much larger than the rms of the mean (see Fig. 7 in Maturi et al. 2019). As we verified, such uncertainty on $\Delta z_{\rm bias}$ does not significantly impact our final results. Conversely, AMICO provides unbiased estimates of redshift (see Maturi et al. 2019), thus we modelled $P(z_{\rm ob}|z_{\rm tr})$ by keeping the mean of such distributions fixed to the central value of $\Delta z_{\rm tr}$. In particular, in the mock catalogue, we measured $P(z_{\rm ob}|z_{\rm tr})$ in several bins of $z_{\rm tr}$, namely $\Delta z_{\rm tr}$, and we performed the statistical MCMC analysis assuming a common flat prior on the rms in all the $\Delta z_{\rm tr}$ bins. The resulting rms of $P(z_{\rm ob}|z_{\rm tr})$ is equal to 0.025. AMICO also provides unbiased estimates for λ^*, thus following the same procedure adopted for $P(z_{\rm ob}|z_{\rm tr})$, we derived an uncertainty of 17% on $\lambda^*_{\rm ob}$, defining the rms of the Gaussian distribution $P(\lambda^*_{\rm ob}|\lambda^*_{\rm tr})$, whose mean is equal to $\lambda^*_{\rm tr}$. We neglected the uncertainties on the rms of $P(z_{\rm ob}|z_{\rm tr})$ and $P(\lambda^*_{\rm ob}|\lambda^*_{\rm tr})$, amounting to 1%, since we verified their negligible effect on the final results.

Furthermore, $P(\lambda^*_{\rm tr}|M_{\rm tr}, z_{\rm tr})$ is a probability distribution that weights the expected counts according to the shape of the mass-richness scaling relation, and it is expressed as follows:

$$P(\lambda^*_{\rm tr}|M_{\rm tr}, z_{\rm tr}) = \frac{P(M_{\rm tr}|\lambda^*_{\rm tr}, z_{\rm tr})\, P(\lambda^*_{\rm tr}|z_{\rm tr})}{P(M_{\rm tr}|z_{\rm tr})}, \tag{5.6}$$

where the distribution $P(M_{\rm tr}|\lambda^*_{\rm tr}, z_{\rm tr})$ is a log-normal one whose mean is given by the mass-richness scaling relation and the standard deviation is given by the intrinsic scatter, $\sigma_{\rm intr}$, set as a free parameter of the model:

$$P(\log M_{\rm tr}|\lambda^*_{\rm tr}, z_{\rm tr}) = \frac{1}{\sqrt{2\pi}\sigma_{\rm intr}} \exp\left(-\frac{x^2(M_{\rm tr}, \lambda^*_{\rm tr}, z_{\rm tr})}{2\sigma^2_{\rm intr}}\right), \tag{5.7}$$

where

$$x(M_{\rm tr}, \lambda^*_{\rm tr}, z_{\rm tr}) = \log\frac{M_{\rm tr}}{10^{14} M_\odot/h} - \left(\alpha + \beta \log\frac{\lambda^*_{\rm tr}}{\lambda^*_{\rm piv}} + \gamma \log\frac{E(z_{\rm tr})}{E(z_{\rm piv})}\right) \tag{5.8}$$

and

$$\sigma_{\rm intr} = \sigma_{\rm intr,0} + \sigma_{\rm intr,\lambda^*} \log\frac{\lambda^*_{\rm tr}}{\lambda^*_{\rm piv}}. \tag{5.9}$$

The $P(M_{\rm tr}|\lambda^*_{\rm tr}, z_{\rm tr})$ distribution indeed accounts for the intrinsic uncertainty that affects a scaling relation between the intrinsic richness and the mass: given an infinitely accurate scaling relation, represented by the mean, the cluster mass provided by a value of intrinsic richness is scattered from the true value. Furthermore, $P(\lambda^*_{\rm tr}|\Delta z_{\rm tr})$ in Eq. (5.6) is a power law with an exponential cut-off, derived from the mock catalogue by considering the objects with $\lambda^*_{\rm tr} \gtrsim 20$. Specifically, similar to other literature analyses (see, e.g., Murata et al. 2019; Costanzi et al. 2019; Abbott et al. 2020), $P(\lambda^*_{\rm tr}|M_{\rm tr}, z_{\rm tr})$ is assumed to be cosmology-independent. Thus we assume that the ratio $P(\lambda^*_{\rm tr}|z_{\rm tr})/P(M_{\rm tr}|z_{\rm tr})$ is cosmology-independent, where $P(M_{\rm tr}|z_{\rm tr})$ acts as a normalisation of $P(\lambda^*_{\rm tr}|M_{\rm tr}, z_{\rm tr})$:

$$P(M_{\rm tr}|z_{\rm tr}) = \int_0^\infty {\rm d}\lambda^*_{\rm tr}\, P(M_{\rm tr}|\lambda^*_{\rm tr}, z_{\rm tr})\, P(\lambda^*_{\rm tr}|z_{\rm tr})\,. \tag{5.10}$$

Halo mass function systematic uncertainties

As mentioned in Section 5.2.2, we assume the Tinker et al. (2008) halo mass function to model the observed cluster counts. Following Costanzi et al. (2019), in order to characterise

the systematic uncertainty in the halo mass function in dark matter only simulations, we related the Tinker et al. (2008) mass function to the true mass function via

$$n(M,z) = n(M,z)^{\rm Tinker}(s\log(M/M^*)+q), \tag{5.11}$$

where $\log M^* = 13.8\ h^{-1}{\rm M}_\odot$ is the pivot mass, while q and s are free parameters of the model with a Gaussian prior having the following covariance matrix:

$$C(\bar{s},\bar{q}) = \begin{pmatrix} 0.00019 & 0.00024 \\ 0.00024 & 0.00038 \end{pmatrix}, \tag{5.12}$$

and with $\bar{s} = 0.037$ and $\bar{q} = 1.008$ as the mean values. Diagonalising the matrix (5.12), we obtained the following 1D Gaussian priors: $\mathcal{N}(\bar{s},\sigma_{\rm s}) = \mathcal{N}(0.037, 0.014)$, and $\mathcal{N}(\bar{q},\sigma_{\rm q}) = \mathcal{N}(1.008, 0.019)$, where $\mathcal{N}(\mu,\sigma)$ stands for a Gaussian distribution with mean μ and standard deviation σ.

Likelihood

Our likelihood function encapsulates the description of cluster counts and weak lensing masses. We base the likelihood term describing the counts, $\mathcal{L}_{\rm counts}$, on the functional form given by Lacasa & Grain (2019), that is to say a convolution of a Poissonian likelihood describing the counts, and a Gaussian distribution accounting for the super-sample covariance (SSC):

$$\mathcal{L}_{\rm counts} = \int {\rm d}\boldsymbol{\delta}_{\rm b}^{n_{\rm z}} \left[\prod_{i,j} {\rm Poiss}\left(N_{i,j}|\bar{N}_{i,j} + \frac{\partial N_{i,j}}{\partial\delta_{{\rm b},j}}\delta_{{\rm b},j}\right)\right] \mathcal{N}(\boldsymbol{\delta}_{\rm b}|0,S)\,. \tag{5.13}$$

In the equation above, $\mathcal{N}(\boldsymbol{\delta}_{\rm b}|0,S)$ is the Gaussian function describing the SSC effects on cluster count measurements, which is a function of the matter density contrast fluctuation, $\boldsymbol{\delta}_{\rm b}$, and has a covariance matrix S.[1] In particular, $n_{\rm z}$ is the number of redshift bins considered in the modelling procedure, and it defines the dimension of the integration variable, $\boldsymbol{\delta}_{\rm b} = \{\delta_{{\rm b},1},...,\delta_{{\rm b},n_{\rm z}}\}$, and of the S matrix, whose dimension is $n_{\rm z}\times n_{\rm z}$. Thus each $\delta_{{\rm b},j}$ represents the fluctuation of the measured matter density contrast, with respect to the expected one, in a given bin of redshift. The indices i and j are the labels of the bins of intrinsic richness and redshift, while $N_{i,j} \equiv N(\Delta\lambda^*_{{\rm ob},i}, \Delta z_{{\rm ob},j})$ is the observed cluster number counts in a bin of intrinsic richness and redshift, and $\bar{N}_{i,j}$ is the model defined in Eq. (5.4). The term $\partial N_{i,j}/\partial\delta_{{\rm b},j}$ is the response of the counts, that is the measure of how the counts vary with changes of the background density, and it is expressed as follows:

$$\begin{aligned}\frac{\partial N_{i,j}}{\partial\delta_{{\rm b},j}} = w(\Delta\lambda^*_{{\rm ob},i},\Delta z_{{\rm ob},j})\ \Omega \int_0^\infty {\rm d}z_{\rm tr}\ \frac{{\rm d}V}{{\rm d}z_{\rm tr}{\rm d}\Omega}\times \\ \times \int_0^\infty {\rm d}M_{\rm tr}\ \frac{{\rm d}n(M_{\rm tr},z_{\rm tr})}{{\rm d}M_{\rm tr}}\ b(M_{\rm tr},z_{\rm tr}) \int_0^\infty {\rm d}\lambda^*_{\rm tr}\ P(\lambda^*_{\rm tr}|M_{\rm tr},z_{\rm tr})\times \\ \times \int_{\Delta z_{{\rm ob},j}} {\rm d}z_{\rm ob}\ P(z_{\rm ob}|z_{{\rm tr},{\rm corr}}) \int_{\Delta\lambda^*_{{\rm ob},i}} {\rm d}\lambda^*_{\rm ob}\ P(\lambda^*_{\rm ob}|\lambda^*_{\rm tr})\,,\end{aligned} \tag{5.14}$$

[1] For the computation of the S matrix, we implemented a code similar to the one at https://github.com/fabienlacasa/PySSC developed by Lacasa & Grain (2019).

Table 5.2: Parameters considered in the joint analysis of cluster counts and stacked weak-lensing data.

Parameter	Description	Prior	Posterior
Ω_m	Total matter density parameter	[0.09, 1]	$0.24^{+0.03}_{-0.04}$
σ_8	Amplitude of the matter power spectrum	[0.4, 1.5]	$0.86^{+0.07}_{-0.07}$
S_8	Cluster normalisation parameter	—	$0.78^{+0.04}_{-0.04}$
α	Normalisation of the mass-richness relation	[-2, 2]	$0.04^{+0.04}_{-0.03}$
β	Slope of the mass-richness relation	[0, 5]	$1.72^{+0.08}_{-0.08}$
γ	Redshift evolution of the mass-richness relation	[-4, 4]	$-2.37^{+0.37}_{-0.56}$
$\sigma_{\mathrm{intr},0}$	Normalisation of σ_intr	[0.05, 1]	$0.18^{+0.08}_{-0.10}$
$\sigma_{\mathrm{intr},\lambda^*}$	λ^* evolution of σ_intr	[-5, 5]	$0.11^{+0.19}_{-0.22}$
Ω_b	Baryon density parameter	$\mathcal{N}(0.0486, 0.0017)$	—
τ	Thomson scattering optical depth at re-ionisation	$\mathcal{N}(0.0544, 0.0365)$	—
n_s	Primordial power spectrum spectral index	$\mathcal{N}(0.9649, 0.0210)$	—
h	Normalised Hubble constant	$\mathcal{N}(0.7, 0.1)$	—
s	Slope correction to the halo mass function	$\mathcal{N}(0.037, 0.014)$	—
q	Amplitude correction to the halo mass function	$\mathcal{N}(1.008, 0.019)$	—
$\boldsymbol{\delta}_\mathrm{b}$	Fluctuation of the mean matter density	Gaussian	—

Notes. In the third column, the priors on the parameters are listed, and in particular a range represents a uniform prior, while $\mathcal{N}(\mu, \sigma)$ stands for a Gaussian prior with mean μ and standard deviation σ. The Gaussian prior on $\boldsymbol{\delta}_\mathrm{b}$ is cosmology-dependent: both the mean and the standard deviation change at each MCMC step. In the fourth column, we show the median values of the 1D marginalised posteriors, along with the 16th and 84th percentiles. The posterior distributions of Ω_b, τ, n_s, h, s, q, and $\boldsymbol{\delta}_b$ are not shown since these nuisance parameters are not constrained in our analysis.

in other words, the response is similar to the model described in Eq. (5.4), in which we also include the contribution of the linear bias $b(M, z)$.

For computational purposes, we consider in the analysis an alternative form of the likelihood, $\mathcal{L}'_\mathrm{counts}$, that is the integrand in Eq. (5.13), of which we computed the natural logarithm:

$$\ln \mathcal{L}'_\mathrm{counts} = \ln \left[\prod_{i,j} \mathrm{Poiss}\left(N_{i,j} | \bar{N}_{i,j} + \frac{\partial N_{i,j}}{\partial \delta_{\mathrm{b},j}} \delta_{\mathrm{b},j} \right) \cdot \mathcal{N}(\boldsymbol{\delta}_\mathrm{b} | 0, S) \right] . \tag{5.15}$$

Here, we set $\boldsymbol{\delta}_\mathrm{b} = \{\delta_{\mathrm{b},1}, ..., \delta_{\mathrm{b},n_z}\}$ as free parameters of the model, with a multivariate Gaussian prior having S as the covariance matrix. Due to the dependence on cosmological parameters of the S matrix, the values of its elements change at every step of the MCMC. In turn, a variation of S implies the change of the prior on $\boldsymbol{\delta}_\mathrm{b}$. At the end of the MCMC, we marginalised over $\boldsymbol{\delta}_\mathrm{b}$ to derive the posteriors of our parameters of interest.

With regard to the likelihood describing the weak lensing masses, $\mathcal{L}_\mathrm{lens}$, we assumed a log-normal functional form and then we considered its natural logarithm:

$$\ln \mathcal{L}_\mathrm{lens} \propto \sum_{k=1}^{N_\mathrm{bin}} \sum_{l=1}^{N_\mathrm{bin}} [\log \bar{M}^k_\mathrm{ob} - \log \bar{M}^k_\mathrm{mod}] \, \boldsymbol{C}^{-1}_{M,kl} \, [\log \bar{M}^l_\mathrm{ob} - \log \bar{M}^l_\mathrm{mod}] \, , \tag{5.16}$$

where N_bin corresponds to the number of bins in which the mean mass $\bar{M}_{200}$, that is $\bar{M}_\mathrm{ob}$, was measured, through the weak-lensing analysis described in Section 5.1.3. Furthermore, $\bar{M}_\mathrm{mod}$ represents the mass obtained from the scaling relation model described in Eq. (5.2),

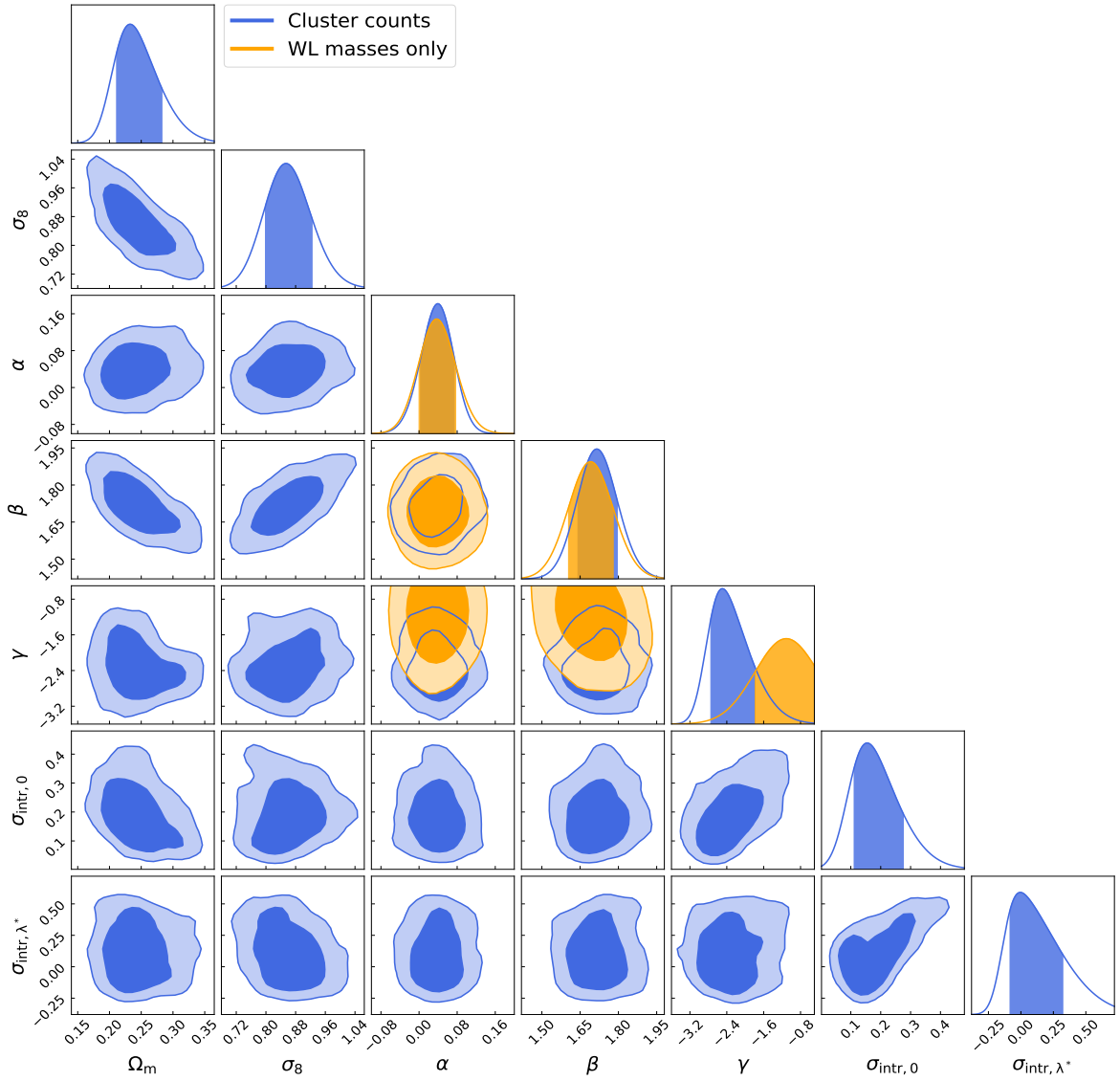

Figure 5.5: Constraints on Ω_m, σ_8, α, β, γ, $\sigma_{\mathrm{intr},0}$, and $\sigma_{\mathrm{intr},\lambda^*}$, derived in a flat ΛCDM universe by combining the redshift bins $z \in [0.1, 0.3]$, $z \in [0.3, 0.45]$, and $z \in [0.45, 0.6]$ and assuming a minimum intrinsic richness $\lambda^*_{\mathrm{ob,\,min}} = 20$ for cluster counts. The shown posteriors were also marginalised over Ω_b, τ, n_s, h, s, q, and $\boldsymbol{\delta}_\mathrm{b}$. The blue contours represent the results obtained from the joint analysis of cluster counts and weak lensing masses, while the orange contours show the posteriors on α, β, and γ, derived from the analysis including only the weak lensing masses as in Bellagamba et al. (2019). The confidence ellipses correspond to 68% and 95%, while the bands over the 1D marginalised posteriors represent the 68% of confidence.

where we assumed the effective redshift and intrinsic richness values z_eff and λ^*_eff listed in Table 5.1. The covariance matrix $\boldsymbol{C}_M$ in Eq. (5.16) has the following form:

$$\boldsymbol{C}_{M,kl} = \delta_{kl} E_k^2 + [\sigma_\mathrm{sys}/\ln(10)]^2 + \delta_{kl}(\sigma_\mathrm{intr}/\sqrt{N_\mathrm{cl}})^2\,, \tag{5.17}$$

where E_k represents the statistical error on $\log \bar{M}_\mathrm{ob}$ derived from the posterior distribution of $\log \bar{M}_\mathrm{ob}$, where we stress that the relative uncertainties are constants after the rescaling described in Section 5.1.3. The term $\sigma_\mathrm{sys} = 0.076$ is the sum in quadrature of the uncertainties on the background selection, photo-zs, shear measurements, halo model, orientation, and projections, obtained in Bellagamba et al. (2019), and N_cl is the number of clusters in the

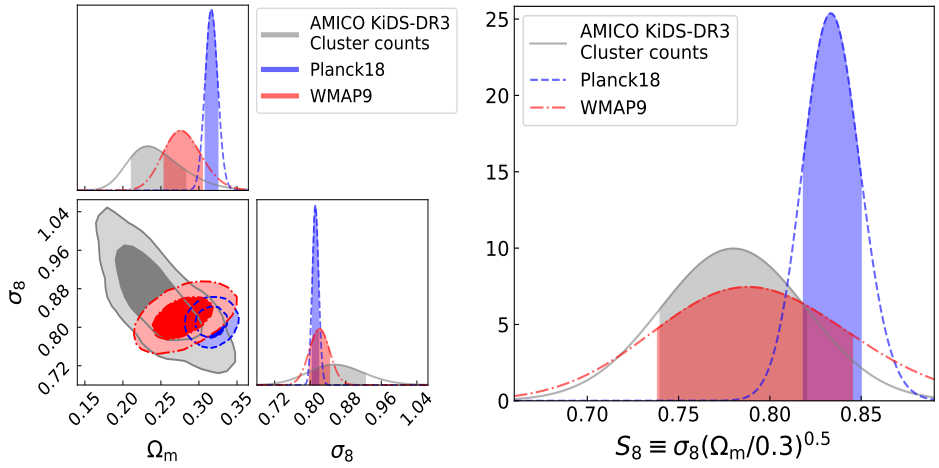

Figure 5.6: Comparison with WMAP and Planck results. In the left panel, we show the Ω_m-σ_8 parameter space, along with the 1D marginalised posteriors with the relative intervals between the 16th and 84th percentiles, in the case of the cluster counts' analysis in the AMICO KiDS-DR3 catalogue (solid grey lines). In the same panel, we also display the results from WMAP (Hinshaw et al. 2013) (Table 3, WMAP-only Nine-year; red dash-dotted lines) and Planck (Planck Collaboration VI 2020) (Table 2, TT, TE, and EE+lowE; blue dashed lines). In the right panel, we show the posteriors for the parameter S_8, where the bands show the intervals between 16th and 84th percentiles. The symbols are the same as in the left panel.

bin of intrinsic richness and redshift in which the mean mass has been derived. By dividing $\sigma_{\rm intr}$ by $\sqrt{N_{\rm cl}}$, we neglected the cluster clustering contribution to the last term of Eq. (5.17).

Thus, the logarithm of the joint likelihood, $\ln \mathcal{L}$, is given by

$$\ln \mathcal{L} = \ln \mathcal{L}'_{\rm counts} + \ln \mathcal{L}_{\rm lens} \,. \tag{5.18}$$

5.2.3 Results

We performed a cosmological analysis of cluster number counts and stacked weak lensing based on the assumption of a flat ΛCDM model. The aim is to constrain the matter density parameter, Ω_m, and the square root of the mass variance computed on a scale of 8 Mpc/h, namely σ_8, with both being provided at $z = 0$, along with the parameters defining the scaling relation between masses and intrinsic richnesses, α, β, and γ in Eq. (5.2), and the intrinsic scatter, $\sigma_{\rm intr}$. Therefore we set Ω_m, σ_8, α, β, γ, and $\sigma_{\rm intr}$ as free parameters of the model, Eq. (5.4), as well as the baryon density, Ω_b, the optical depth at re-ionisation, τ, the primordial spectral index, n_s, the normalised Hubble constant, h, the Tinker mass function correction parameters, q and s, described in Section 5.2.2, and the fluctuation of the mean density of matter due to super-survey modes, $\boldsymbol{\delta}_b$. We assumed flat priors for Ω_m, σ_8, α, β, γ, and $\sigma_{\rm intr}$, while we set Gaussian priors on the other parameters (see Table 5.2). In particular, for the Gaussian prior distributions of Ω_b, τ, n_s, and h, we considered the values obtained by the Planck Collaboration VI (2020) (Table 2, TT, TE, and EE+lowE+lensing), assuming the same mean values and imposing a standard deviation equal to 5σ for all the aforementioned parameters, but h, for which we assumed a standard deviation equal to 0.1. In this baseline cosmological model, we also assumed three neutrino species, approximated as two massless states and a single massive neutrino of mass $m_\nu = 0.06$ eV, following Planck Collaboration VI (2020). Finally, we assumed a multivariate Gaussian prior for $\boldsymbol{\delta}_b$, as described in Section 5.2.2.

In our analysis, we constrained the value of the cluster normalisation parameter, $S_8 \equiv$

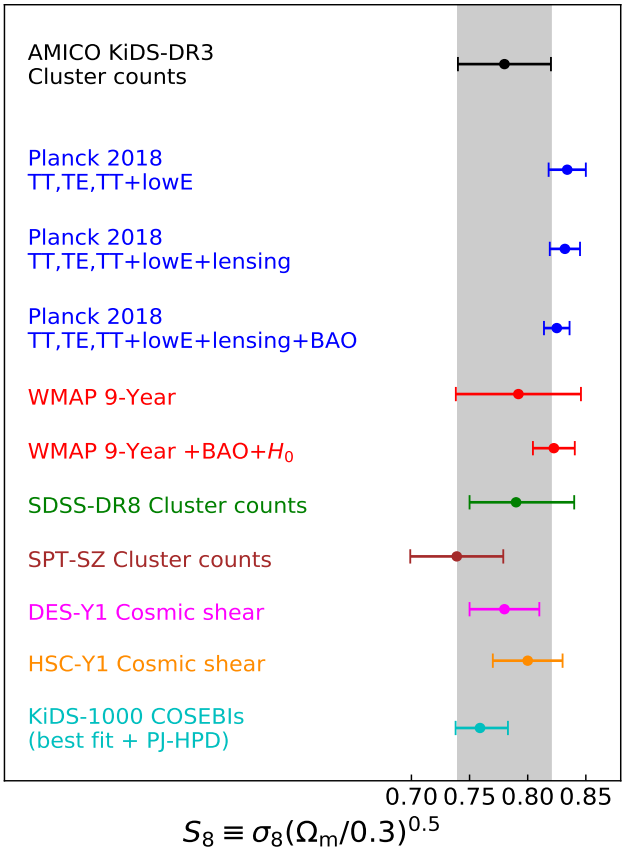

Figure 5.7: Comparison of the constraints on $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$ obtained, from top to bottom, from the joint analysis of cluster counts and weak lensing in the AMICO KiDS-DR3 catalogue (black dot), from the results obtained by Planck Collaboration VI (2020) (blue dots), Hinshaw et al. (2013) (red dots), Costanzi et al. (2019) (green dot), Bocquet et al. (2019) (brown dot), Troxel et al. (2018) (magenta dot), Hikage et al. (2019) (orange dot), and Asgari et al. (2021) (cyan dot). The median, as well as the 16th and 84th percentiles are shown.

$\sigma_8(\Omega_m/0.3)^{0.5}$. The significance of this parameter is rooted in the degeneracy between σ_8 and Ω_m, being defined along the $\sigma_8 - \Omega_m$ confidence regions. Since the number of massive clusters increases with both σ_8 and Ω_m, in order to hold the cluster abundance fixed at its observed value, any increase in σ_8 must be compensated for by a decrease in Ω_m, implying that S_8 is held fixed.

From this modelling, we obtain $\Omega_m = 0.24^{+0.03}_{-0.04}$, $\sigma_8 = 0.86^{+0.07}_{-0.07}$, and $S_8 = 0.78^{+0.04}_{-0.04}$, where we quote the median, 16th and 84th percentiles, as shown in Fig. 5.5 and Table 5.2. In Fig. 5.5 we also show that the results on the mass-richness scaling relation retrieved from this analysis, that is for α, β, and γ, are in agreement within 1σ with those obtained by performing the modelling of the weak lensing data only, as carried out by Bellagamba et al. (2019). In particular, the inclusion of the cluster counts in the analysis provides tighter constraints on the slope β, also governing the slope of the cluster model at low values of λ^*. Additionally, the cluster count redshift evolution provides a more accurate estimate of γ. Lastly, we find a tight constraint on the intrinsic scatter, deriving $\sigma_{\rm intr,0} = 0.18^{+0.08}_{-0.10}$ and $\sigma_{\rm intr,\lambda^*} = 0.11^{+0.19}_{-0.22}$, which confirms the reliability of λ^* as a mass proxy. This result on $\sigma_{\rm intr}$ is consistent within 1σ with that derived in Sereno et al. (2020) from a weak lensing analysis of the sample of AMICO clusters in KiDS-DR3.

As shown in Fig. 5.6, the constraints obtained for S_8 and σ_8 are in agreement within 1σ

with WMAP results (Hinshaw et al. 2013) (Table 3, WMAP-only Nine-year), and with Planck results (Planck Collaboration VI 2020) (Table 2, TT, TE, and EE+lowE). With regard to $\Omega_{\rm m}$, we find an agreement within 1σ with WMAP and a 2σ tension with Planck. Furthermore, in Fig. 5.7 we show the comparison with the S_8 constraints obtained from additional external data sets. In particular, we find an agreement within 1σ with the results obtained from the cluster counts' analyses performed by Costanzi et al. (2019), based on SDSS-DR8 data, and by Bocquet et al. (2019), based on the 2500 deg^2 SPT-SZ survey data, as well as with the results derived from the cosmic shear analyses performed by Troxel et al. (2018) on DES-Y1 data, Hikage et al. (2019) on HSC-Y1 data, and Asgari et al. (2021) on KiDS-DR4 data. The constraint on S_8 obtained from the cluster counts and weak lensing joint analysis in DES (Abbott et al. 2020), $S_8 = 0.65^{+0.04}_{-0.04}$, not shown in Fig. 5.7, is not consistent with our result.

Following a more conservative approach, we repeated the analysis also assuming the threshold in intrinsic richness $\lambda^* \geq 20$ for the weak lensing data. This leads to $\Omega_{\rm m} = 0.27^{+0.04}_{-0.05}$, $\sigma_8 = 0.83^{+0.06}_{-0.07}$, and $S_8 = 0.78^{+0.04}_{-0.04}$, which are consistent within 1σ with the constraints derived from the analysis previously described. Also for the other free parameters of the model, the consistency within 1σ still holds.

5.2.4 Summary

In this section, we presented the work by Lesci et al. (2022a) on the joint analysis of galaxy cluster abundance and weak-lensing in the AMICO KiDS-DR3 catalogue (Maturi et al. 2019). Lesci et al. (2022a) simultaneously constrained the cosmological parameters and the cluster mass-richness scaling relation. In particular, we relied on the intrinsic richness, defined in Eq. (5.1), as the observable linked to the cluster masses. The sample exploited for cluster counts includes 3652 galaxy clusters having an intrinsic richness of $\lambda^* \geq 20$ and divided in the redshift bins $z \in [0.1, 0.3]$, $z \in [0.3, 0.45]$, and $z \in [0.45, 0.6]$. For the weak-lensing analysis, we followed the procedure developed by Bellagamba et al. (2019), not assuming any thresholds in λ^*. We assessed the incompleteness and the impurities of the cluster sample by exploiting a mock catalogue developed by Maturi et al. (2019), and we corrected our data accordingly.

We assumed a model for cluster counts, shown in Eq. (5.4), accounting for the redshift uncertainties and for the mass-richness scaling relation. In particular, the mass-richness scaling relation plays a crucial role in the $P(\lambda^*_{\rm tr}|M_{\rm tr}, z_{\rm tr})$ term given by Eq. (5.6), which also depends on the observed distribution of galaxy clusters as a function of the intrinsic richness. Furthermore, this term includes the contribution of the intrinsic scatter of the scaling relation, $\sigma_{\rm intr}$, which is considered as an unknown parameter. Subsequently, we modelled the cluster counts and the scaling relation by combining the relative likelihood functions.

Assuming a flat ΛCDM model with massive neutrinos, we found $\Omega_{\rm m} = 0.24^{+0.03}_{-0.04}$, $\sigma_8 = 0.86^{+0.07}_{-0.07}$, and $S_8 = 0.78^{+0.04}_{-0.04}$, which are competitive constraints, in terms of uncertainties, with results from state-of-the-art cluster number counts' analyses. In addition, the result on S_8 is in agreement within 1σ with the results from WMAP and Planck. We also derived results for the scaling relation that are consistent within 1σ with those obtained by only modelling the weak lensing signal as in Bellagamba et al. (2019), thus validating the reliability of our model. With regard to the intrinsic scatter, we found $\sigma_{\rm intr,0} = 0.18^{+0.08}_{-0.10}$ and $\sigma_{{\rm intr},\lambda^*} = 0.11^{+0.19}_{-0.22}$, which is a very competitive result compared to the present-day estimates in the field of galaxy clusters, outlining the goodness of the assumption of λ^* as the mass proxy.

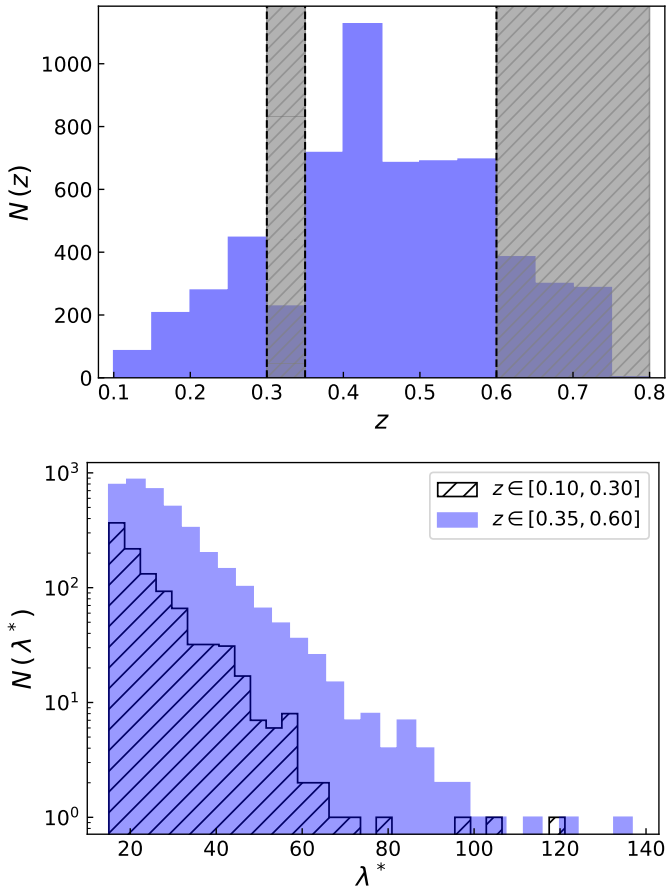

Figure 5.8: Redshift and intrinsic richness distributions of the clusters in the sample analysed by Lesci et al. (2022b). *Top panel:* Redshift distribution after the correction for the redshift bias, including only objects with $\lambda^* > 15$. The grey shaded areas represent the redshift ranges that were not used in the analysis. *Bottom panel:* λ^* distribution of the objects in the redshift ranges $z \in [0.10, 0.30]$ (black hatched histogram) and $z \in [0.35, 0.60]$ (blue histogram).

5.3 AMICO KiDS-DR3 cluster clustering

In this section, we present the analysis of AMICO KiDS-DR3 galaxy cluster 3D 2PCF monopole by Lesci et al. (2022b), based on the data set detailed in Sect. 5.1. In Sect. 5.3.1 we discuss the sample selection and the measurements of the cluster 3D 2PCF, while in Sect. 5.3.2 we present the theoretical models describing the matter power spectrum and effective bias. In Sect. 5.3.3 we discuss the results on cosmological parameters and on the mass-richness scaling relation, while in Sect. 5.3.4 we draw our conclusions.

5.3.1 Sample selection and measurements

Before measuring the cluster 2PCF, we accounted for the bias described in Maturi et al. (2019) that affects cluster redshifts, that is, we defined the corrected redshift as $z_{\rm corr} = z - 0.02\,(1+z)$. This bias corresponds to what was found in de Jong et al. (2017) by comparing the KiDS photo-zs to the GAMA spectroscopic redshifts. Furthermore, in a small redshift range around $z \sim 0.32$, the photo-z errors are higher and harder to model because of the shape of the g and r filter transmissions. In the following cosmological analysis (Section 5.3.3), we also considered the photo-z range $z \in [0.10, 0.60]$ because we assumed priors on the mass-richness

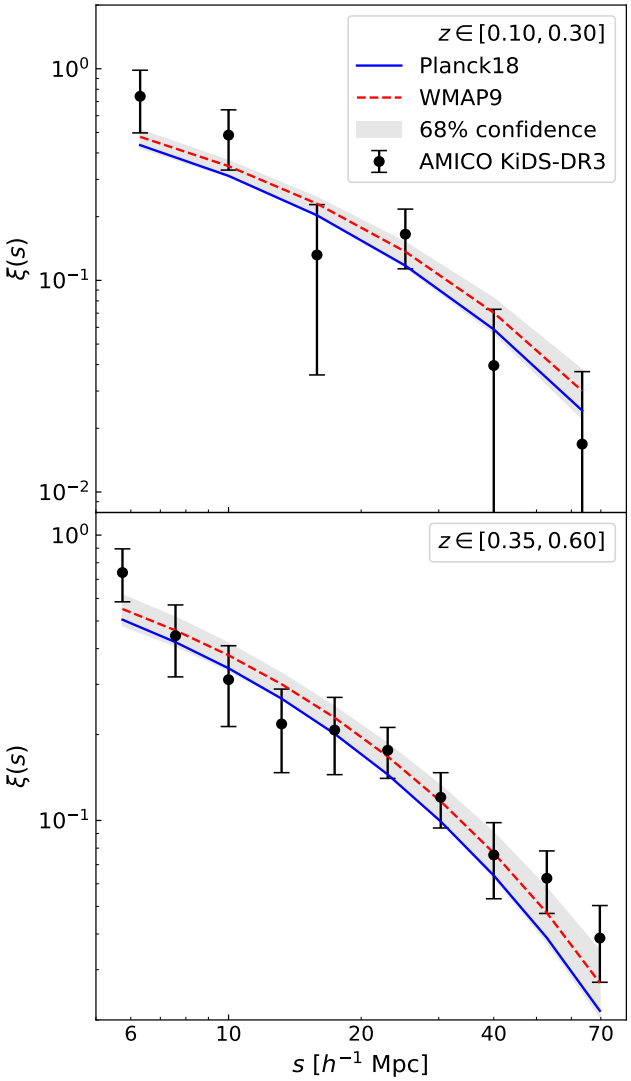

Figure 5.9: Redshift-space 2PCF monopole (black dots) of the AMICO KiDS-DR3 clusters in the spatial range $s \in [5, 80]$ h^{-1}Mpc, and redshift ranges $z \in [0.10, 0.30]$ (*top panel*) and $z \in [0.35, 0.60]$ (*bottom panel*). In both panels, the grey band represents the 68% confidence level derived from the multivariate posterior of the free parameters considered in the cosmological analysis, described in Section 5.3.3. The model computed by assuming the cosmological parameters derived by Planck (Planck Collaboration VI 2020, Table 2, TT, TE, and EE+lowE; blue lines) and WMAP (Hinshaw et al. 2013, Table 3, WMAP-only Nine-year; red lines) is represented by the solid blue lines and by the dashed red lines, respectively. In both cases, the median values of the scaling relation parameters derived by Lesci et al. (2022a) are assumed.

relation estimated from a weak-lensing analysis performed in this redshift range (Bellagamba et al. 2019). Consequently, we decided to model the 2PCF in two separate redshift bins: $z \in [0.10, 0.30]$ and $z \in [0.35, 0.60]$. We considered only the clusters with $\lambda^* > 15$, which ensures a purity higher than 97% over the whole sample (Maturi et al. 2019). Consequently, the sample consists of 1019 clusters for $z \in [0.10, 0.30]$ and 3915 clusters for $z \in [0.35, 0.60]$, for a total of 4934 objects. Fig. 5.8 shows the redshift and richness distributions of the objects considered in the analysis.

We estimated the redshift-space 2PCF monopole with the LS estimator (Landy & Szalay 1993), detailed in Sect. 2.1. In order to build up the random catalogue, we extracted random

(R.A., Dec) cluster positions within the survey tiles by accounting for the same masks as were used by Maturi et al. (2019) for the construction of the cluster catalogue. Regarding the redshifts, we shuffled the observed photo-zs. Namely, to each object of the random sample, we assigned the photo-z randomly extracted from the AMICO KiDS-DR3 catalogue. Hence, the random sample has the same redshift distribution as the real sample by construction. We verified that the observed cluster distribution does not depend on the sky position, which is expected as Maturi et al. (2019) imposed the strict magnitude cut at $r = 24$, corresponding to the depth of the shallowest tile. The generated random catalogue is 30 times larger than the AMICO KiDS-DR3 cluster sample to limit shot noise effects. The observed coordinates were then converted into comoving ones by assuming the ΛCDM model, with cosmological parameters from Planck Collaboration VI (2020). In Fig. 5.9, we show the 2PCF monopole measurements in the two redshift bins considered in the analysis, and in the spatial range $s \in [5, 80]\ h^{-1}$Mpc. At larger scales, the clustering signal starts to be weak and is dominated by the errors, while at smaller scales, the signal is negligible as the cluster size sets the minimum cluster separation.

5.3.2 Modelling

As we discussed in Sect. 2.4, the observed redshift, $z_{\rm ob}$, can be expressed as

$$z_{\rm ob} = z_{\rm c} + \frac{v_{\parallel}}{c}\,(1 + z_{\rm c}) \pm \sigma_z\,, \tag{5.19}$$

where $z_{\rm c}$ is the cosmological redshift, σ_z is the error on the redshift measurements, and $v_{\parallel}$ is the component of the peculiar velocity along the line of sight. Therefore, using $z_{\rm ob}$ to estimate the comoving distance creates distortions in the measures of the 2PCF, not only because of the error on the measurements (σ_z), but also because $z_{\rm ob}$ encodes information on the peculiar motions along the line of sight ($v_{\parallel}$). The peculiar motions cause the so-called dynamical distortions, an effect commonly known also as redshift-space distortions (RSD).

Since our whole analysis is performed on a catalogue extracted from photometric data, it is crucial to take the errors on the observed cluster redshifts into account. Following the approach presented in Sereno et al. (2015), we modelled the redshift-space 2D power spectrum as follows:

$$P(k, \mu) = P_{\rm DM}(k)\,\left(b_{\rm eff} + f\mu^2\right)^2\,\exp\left(-k^2\,\mu^2\,\sigma^2\right)\,, \tag{5.20}$$

where $P_{\rm DM}(k)$ is the dark matter power spectrum, $k = \sqrt{k_{\perp}^2 + k_{\parallel}^2}$, with $k_{\parallel}$ and $k_{\perp}$ being the wave-vector components parallel and perpendicular to the line of sight, respectively, $\mu \equiv k_{\parallel}/k$, $b_{\rm eff}$ represents the effective bias factor (detailed in the following), f is the growth rate, and the $f\mu^2$ term parametrises the coherent motions of large-scale structure, enhancing the clustering signal at all scales (Kaiser 1987). The exponential cut-off term describes the random perturbations affecting redshift measurements, caused by both nonlinear stochastic motions and redshift errors. This is a Gaussian damping term, which causes a scale-dependent effect by removing the signal over a typical scale $k \sim 1/\sigma$, where σ represents the displacement along the line of sight due to random perturbations of cosmological redshifts, defined as

$$\sigma \equiv \frac{c\,\sigma_z}{H(z_{\rm m})}\,, \tag{5.21}$$

where $H(z_{\rm m})$ is the Hubble function computed at the mean redshift of the cluster distribution in the bin, $z_{\rm m}$, and σ_z is the typical cluster redshift error, expressed as

$$\sigma_z = \sigma_{z,0}(1 + z_{\rm m}), \tag{5.22}$$

where $\sigma_{z,0}$ is a free parameter in the analysis. We derived $\sigma_{z,0}$ from the mock catalogue described in Maturi et al. (2019). In particular, we measured the conditional probability $P(z_{\rm ob}|z_{\rm tr})$, where $z_{\rm ob}$ and $z_{\rm tr}$ are the observed and true redshifts, respectively, in several bins of $z_{\rm tr}$, namely $\Delta z_{\rm tr}$. We described the standard deviation of $P(z_{\rm ob}|z_{\rm tr})$ through Eq. (5.22), where $z_{\rm m}$ is the mean value of $z_{\rm tr}$ within $\Delta z_{\rm tr}$. Given the input galaxy photo-zs, AMICO provides unbiased estimates of redshift (Maturi et al. 2019). Then we performed a statistical MCMC analysis assuming a common flat prior on $\sigma_{z,0}$ in all the $\Delta z_{\rm tr}$ bins, obtaining $\sigma_{z,0} = 0.02$ with an uncertainty of 2×10^{-4}, namely equal to 1%.

To derive the monopole of the correlation function, we integrated Eq. (5.20) over μ and computed the inverse Fourier transform of the result. The solution can be written as a function of $b_{\rm eff}$ as follows:

$$\xi(s) = b_{\rm eff}^2\, \xi'(s) + b_{\rm eff}\xi''(s) + \xi'''(s)\,, \tag{5.23}$$

where $\xi'(s)$ is the inverse Fourier transform of the monopole $P'(k)$, that is,

$$P'(k) = P_{\rm DM}(k)\frac{\sqrt{\pi}}{2k\sigma}\,{\rm erf}(k\sigma)\,, \tag{5.24}$$

and $\xi''(s)$ and $\xi'''(s)$ are the inverse Fourier transform of $P''(k)$ and $P'''(k)$, respectively,

$$P''(k) = \frac{f}{(k\sigma)^3}P_{\rm DM}(k)\left[\frac{\sqrt{\pi}}{2}\,{\rm erf}(k\sigma) - k\sigma\exp(-k^2\sigma^2)\right]\,,$$

$$P'''(k) = \frac{f^2}{(k\sigma)^5}P_{\rm DM}(k)\frac{3\sqrt{\pi}}{8}\,{\rm erf}(k\sigma) - \frac{k\sigma}{4}\left[2\,(k\sigma)^2 + 3\right]\exp(-k^2\sigma^2)\,. \tag{5.25}$$

We neglected geometric distortions, which appear when a fiducial cosmology is assumed (in our case, Planck Collaboration VI 2020, Table 2, TT, TE, and EE+lowE) to convert observed coordinates into physical ones, because their effect is negligible with respect to dynamic distortions and photo-z errors (see Marulli et al. 2012).

Effective bias and mass-richness relation

The cosmological model of structure formation and evolution predicts that the dark matter halo bias, b, primarily depends on halo mass and redshift. Specifically, at a fixed redshift, the bias increases with the tracer mass, while for a given mass, the bias is an increasing function of the redshift (see, e.g., Sheth & Tormen 1999). We derived the effective bias in the ith bin of redshift, namely Δz_i, as

$$b_{\rm eff}(\Delta z_i) = \frac{1}{N_i}\sum_{j=1}^{N_i}\int_0^\infty {\rm d}z\int_0^\infty {\rm d}\lambda^*\int_0^\infty {\rm d}M\; b(M,z)\,P(M|\lambda^*,z)\,P(z|z_{{\rm ob},j})\,P(\lambda^*|\lambda^*_{{\rm ob},j}), \tag{5.26}$$

where N_i is the number of clusters in the ith redshift bin, j is the cluster index, and b is the halo bias, for which the model by Tinker et al. (2010) is assumed. As discussed in Section 5.3.3, the results do not significantly change when the halo bias model by Sheth et al. (2001) is assumed. In addition, $P(M|\lambda^*,z)$ is a log-normal distribution expressed in Eqs. (5.7) and (5.8), where we set $\lambda^*_{\rm piv} = 30$ and $z_{\rm piv} = 0.35$ following Bellagamba et al. (2019). The rms of $P(\log M|\lambda^*,z)$ is the intrinsic scatter $\sigma_{\rm intr}$, expressed as a function of the intrinsic richness (see Eq. 5.9). In Eq. (5.26), $P(z|z_{{\rm ob},j})$ and $P(\lambda^*|\lambda^*_{{\rm ob},j})$ are Gaussian distributions, whose mean is the jth cluster's observed redshift, $z_{{\rm ob},j}$, and richness, $\lambda^*_{{\rm ob},j}$, respectively. The rms of $P(z|z_{{\rm ob},j})$ is expressed as $\sigma_{z,0}(1+z_{{\rm ob},j})$, where $\sigma_{z,0}$ was derived from the mock catalogue developed by Maturi et al. (2019) as described above. Analogously, we derived an uncertainty on λ^* amounting to 17% from the mock catalogue.

Table 5.3: Parameters considered in the cosmological analysis by Lesci et al. (2022b).

Parameter	Description	Prior	Posterior
$\Omega_{\rm m}$	Total matter density parameter	[0.09, 1]	$0.28^{+0.05}_{-0.04}$
σ_8	Amplitude of the matter power spectrum	[0.4, 1.5]	$0.82^{+0.14}_{-0.12}$
S_8	Cluster normalisation parameter	—	$0.80^{+0.08}_{-0.08}$
α	Normalisation of the mass-richness relation	$\mathcal{N}(0.04, 0.04)$	—
β	Slope of the mass-richness relation	$\mathcal{N}(1.72, 0.08)$	—
γ	Redshift evolution of the mass-richness relation	$\mathcal{N}(-2.37, 0.40)$	—
$\sigma_{\rm intr,0}$	Normalisation of $\sigma_{\rm intr}$	$\mathcal{N}(0.18, 0.09)$	—
$\sigma_{\rm intr,\lambda^*}$	λ^* evolution of $\sigma_{\rm intr}$	$\mathcal{N}(0.11, 0.20)$	—
$\sigma_{z,0}$	Factor entering the damping of the power spectrum	$\mathcal{N}(0.02, 2\times10^{-4})$	—
$\Omega_{\rm b}$	Baryon density parameter	$\mathcal{N}(0.0486, 0.0017)$	—
$n_{\rm s}$	Primordial power spectrum spectral index	$\mathcal{N}(0.9649, 0.0210)$	—
h	Normalised Hubble constant	$\mathcal{N}(0.7, 0.1)$	—

Notes. In the third column, the priors on the parameters are listed. In particular, a range represents a uniform prior, while $\mathcal{N}(\mu, \sigma)$ stands for a Gaussian prior with mean μ and standard deviation σ. In the fourth column, we show the median values of the 1D marginalised posteriors, along with the 16th and 84th percentiles. The posterior distributions of α, β, γ, $\sigma_{\rm intr,0}$, $\sigma_{\rm intr,\lambda^*}$, $\sigma_{z,0}$, Ω_b, n_s, and h are not shown since these parameters are not constrained in our analysis.

Likelihood

For the cosmological Bayesian analysis performed in this work, a standard Gaussian likelihood was considered,

$$\mathcal{L} \propto \exp(-\chi^2/2)\,, \tag{5.27}$$

with

$$\chi^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\xi_i^d - \xi_i^m\right) C_{i,j}^{-1} \left(\xi_j^d - \xi_j^m\right)\,, \tag{5.28}$$

where N is the number of comoving separation bins in which the 2PCF is computed, d and m indicate data and model, respectively, and $C_{i,j}^{-1}$ is the inverse of the covariance matrix. The covariance matrix measures the variance and the correlation between the different bins of the 2PCF. It is estimated from the data with the jackknife technique (Norberg et al. 2009),

$$C_{i,j} = \frac{N_{\rm sub}-1}{N_{\rm sub}} \sum_{k=1}^{N}\left(\xi_i^k - \bar{\xi}_i\right)\left(\xi_j^k - \bar{\xi}_j\right)\,, \tag{5.29}$$

where ξ_i^k is the value of the correlation function at the i-th bin for the k-th subsample, $\bar{\xi}_i$ is the mean value of the subsample, and $N_{\rm sub}$ is the number of resamplings of our cluster catalogue. In particular, the survey tiles are set as the subsample regions for the jackknife.

5.3.3 Results

Based on the methods outlined in the previous sections, we performed a cosmological analysis of the redshift-space 2PCF of the AMICO KiDS-DR3 cluster sample. Our analysis was based on two fully independent redshift bins, $z \in [0.10, 0.30]$ and $z \in [0.35, 0.60]$, and we

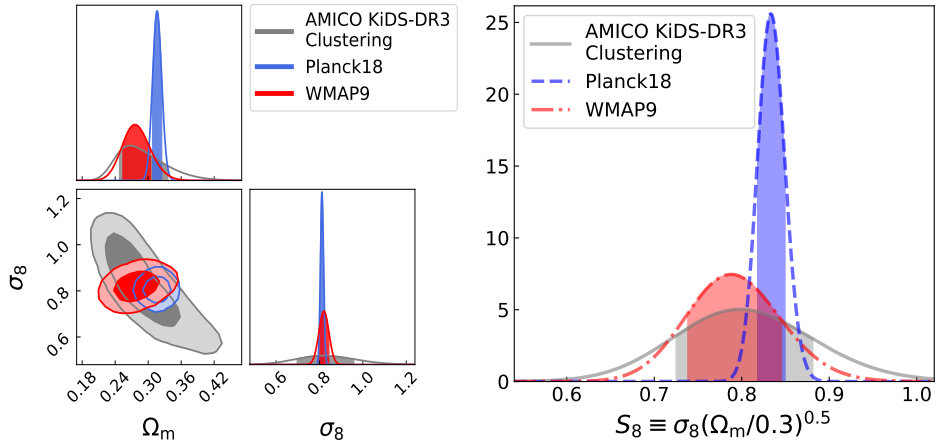

Figure 5.10: Constraints obtained from the cosmological analysis, compared to WMAP and Planck results. In the left panel, we show the 68% and 95% confidence levels in the Ω_m-σ_8 parameter space, along with the 1D marginalised posteriors with the relative intervals between the 16th and 84th percentiles, in the case of the cluster clustering analysis of the AMICO KiDS-DR3 catalogue (grey lines). In the same panel, we also display the results from WMAP (Hinshaw et al. 2013, Table 3, WMAP-only Nine-year; red lines) and Planck (Planck Collaboration VI 2020, Table 2, TT, TE, and EE+lowE; blue lines). In the right panel, we show the posteriors for the parameter S_8, where the bands show the intervals between 16th and 84th percentiles. The colours are the same as in the left panel.

considered galaxy clusters with $\lambda^* > 15$. The 2PCF of this cluster sample was estimated in the spatial range $s \in [5, 80]$ h^{-1}Mpc. The aim of this analysis is to constrain the matter density parameter, Ω_m, the square root of the mass variance computed on a scale of 8 h^{-1}Mpc, σ_8, and the cluster normalisation parameter, $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$, by assuming Gaussian priors for the parameters of the mass-richness relation. In addition, we investigated a method to infer the normalisation of the cluster mass-richness relation from cluster clustering measurements.

Constraints on cosmological parameters

We exploited the methods described in Section 5.3.2 to constrain the main parameters of the ΛCDM model. We assumed large flat priors for σ_8 and Ω_m, while for the parameters of the mass-richness relation (Eqs. 5.8, 5.8), α, β, γ, $\sigma_{\rm intr,0}$, and $\sigma_{\rm intr,\lambda^*}$, we considered Gaussian priors with the mean and standard deviation given by the posteriors derived from the joint analysis of cluster counts and weak lensing performed by Lesci et al. (2022a) (see Sect. 5.2). We also assumed a Gaussian prior on $\sigma_{z,0}$, entering the damping factor of $P(k)$ accounting for the uncertainties on the photo-zs (Eqs. 5.21 - 5.22), with the mean equal to 0.02 and the standard deviation equal to 2×10^{-4} (see Section 5.3.2). Lastly, we assumed Gaussian priors for the baryon density, Ω_b, the primordial spectral index, n_s, and the normalised Hubble constant, h, assuming the same mean values derived by Planck Collaboration VI (2020, Table 2, TT, TE, and EE+lowE). With regard to the standard deviation of such priors, for Ω_b and n_s we imposed a standard deviation equal to five times the 1σ error derived by Planck, while for h we assumed a standard deviation equal to 0.1. We verified that the results do not significantly change if we instead use the 1σ errors derived from Planck as the standard deviations for the priors on Ω_b, n_s, and h. In Table 5.3 we show the priors and the posteriors of the free parameters of the model. In Fig. 5.9 we compare our 2PCF measurements in the two selected redshift ranges to the best-fit model. The statistical analysis was performed by assuming a standard Gaussian likelihood, defined in Eq. (5.27).

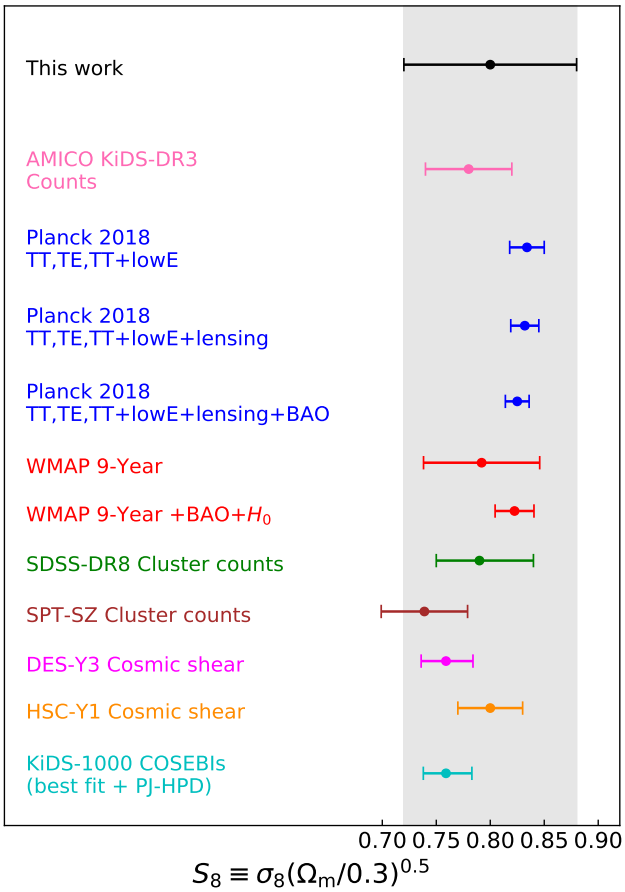

Figure 5.11: Comparison of the constraints on $S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$ obtained, from top to bottom, from the cluster clustering analysis in the AMICO KiDS-DR3 catalogue (black dot), from the joint analysis of cluster counts and weak lensing in KiDS-DR3 performed by Lesci et al. (2022a) (pink dot), from the results obtained by Planck Collaboration VI (2020) (blue dots), Hinshaw et al. (2013) (red dots), Costanzi et al. (2019) (green dot), Bocquet et al. (2019) (brown dot), Amon et al. (2022) and Secco et al. (2022) (magenta dot), Hikage et al. (2019) (orange dot), and Asgari et al. (2021) (cyan dot). The median and the 16th and 84th percentiles are shown.

We obtained $\Omega_m = 0.28^{+0.05}_{-0.04}$, $\sigma_8 = 0.82^{+0.14}_{-0.12}$, and $S_8 = 0.80^{+0.08}_{-0.08}$, where we quote the median and 16th and 84th percentiles, as shown in Fig. 5.10 and Table 5.3. These constraints agree within 1σ with the WMAP (Hinshaw et al. 2013, Table 3, WMAP-only Nine-year) and Planck results (Planck Collaboration VI 2020, Table 2, TT, TE, and EE+lowE). In addition, Fig. 5.11 shows that our constraint on S_8 agrees within 1σ with the results obtained from the joint analysis of cluster counts and weak lensing performed in KiDS-DR3 by Lesci et al. (2022a). The agreement within 1σ also holds for the cluster count analyses performed by Costanzi et al. (2019), based on SDSS-DR8 data, and by Bocquet et al. (2019), based on the 2500 deg^2 SPT-SZ survey data, as well as for the results derived from the cosmic shear analyses performed by Amon et al. (2022) and Secco et al. (2022) on DES-Y3 data, Hikage et al. (2019) on HSC-Y1 data, and Asgari et al. (2021) on KiDS-DR4 data. In addition, our result on S_8 agrees with the constraint by Lindholm et al. (2021), namely $S_8 = 0.85^{+0.10}_{-0.08}$, derived from the autocorrelation of X-ray selected CODEX clusters. We also performed the analysis by assuming the halo bias model by Sheth et al. (2001), obtaining $S_8 = 0.79^{+0.08}_{-0.08}$. This result agrees well with the result derived from the previously described analysis.

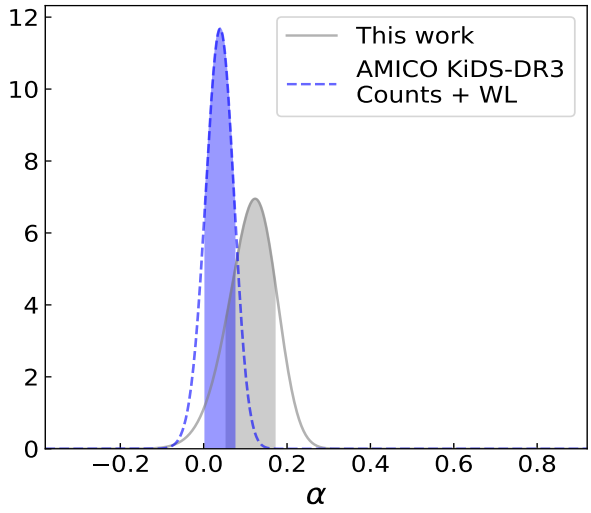

Figure 5.12: Constraint on the normalisation of the mass-richness relation, α. The result obtained from the cluster clustering analysis is shown in grey, and the constraint derived from the joint analysis of counts and weak lensing (Lesci et al. 2022a) is shown in blue. The bands show the intervals between the 16th and 84th percentiles.

Mass-richness scaling relation

Cluster clustering might provide robust constraints on the normalisation of the mass-richness relation, α, if a cosmological model is assumed (see, e.g., Chiu et al. 2020). Based on the 2PCF measures used in the cosmological analysis detailed in the previous section, we also performed the analysis by assuming a flat prior on α and Gaussian priors on β, γ, $\sigma_{\rm intr,0}$, and $\sigma_{\rm intr,\lambda^*}$, given by the posteriors derived by Lesci et al. (2022a), and the same prior on $\sigma_{z,0}$ assumed in Section 5.3.3. In addition, we fixed the cosmological parameters to the values derived by Planck Collaboration VI (2020, Table 2, TT, TE, and EE+lowE). We obtained $\alpha = 0.12^{+0.06}_{-0.06}$, which agrees within 1σ with the result obtained by Lesci et al. (2022a), as shown in Fig. 5.12. It should be noted that the median value of α corresponds within 2σ to the prior considered in Section 5.3.3. This is due to the adopted prior on S_8, whose value is higher than what was predicted by the clustering signal (see Fig. 5.11). A higher value of α is required to compensate for a higher S_8.

5.3.4 Summary

In this section, we presented the study by Lesci et al. (2022b) on the clustering properties of the AMICO KiDS-DR3 galaxy cluster catalogue (Maturi et al. 2019). The sample consists of 4934 clusters in the redshift bins $z \in [0.1, 0.3]$, $z \in [0.35, 0.6]$, with intrinsic richness $\lambda^* > 15$. We measured the monopole of the 2PCF and performed a cosmological statistical analysis. The clustering model considered includes a damping of the power spectrum to account for the uncertainties on the photo-zs. In addition, we derived the effective bias in each redshift bin (Eq. 5.26) from the mass-richness scaling relation (Eq. 5.8).

We performed a cosmological analysis by assuming flat priors on $\Omega_{\rm m}$ and σ_8, and Gaussian priors on the parameters of the mass-richness relation given by the constraints derived by Lesci et al. (2022a), from the joint analysis of cluster counts and weak lensing in KiDS-DR3. In addition, we marginalised our posteriors over the other cosmological parameters and over the damping of the power spectrum caused by the uncertainties on photo-zs. From this modelling, we derived $\Omega_{\rm m} = 0.28^{+0.05}_{-0.04}$, $\sigma_8 = 0.82^{+0.14}_{-0.12}$, and $S_8 = 0.80^{+0.08}_{-0.08}$, which are

consistent within 1σ with the results obtained from CMB experiments and from the most recent analyses of the late Universe. In addition, by fixing the cosmological parameters to the values derived by Planck Collaboration VI (2020, Table 2, TT, TE, and EE+lowE) and assuming Gaussian priors on the parameters β and γ of the mass-richness relation, we derived a robust constraint on the normalisation, α. In particular, we obtained $\alpha = 0.12^{+0.06}_{-0.06}$, which agrees within 1σ and is competitive, in terms of uncertainties, with the result obtained by Lesci et al. (2022a). This confirms the key role of clustering in constraining cluster masses, as we also detailed in Sect. 3.4 and in Chapt. 4.

For the next KiDS data release, namely KiDS-1000 (Kuijken et al. 2019), we will combine counts, clustering, and weak lensing to improve the accuracy of our results further. KiDS-1000 covers an area of 1000 square degrees, which is more than two-thirds of the final area, and photometry extends to the near-infrared ($ugriZYJHK_\mathrm{s}$), joining the data from the KiDS and VIKING (Edge et al. 2013a) surveys, thus allowing, for instance, one to improve the photometric redshift estimates (Wright et al. 2019).

Chapter 6

Galaxy colour selections with *Euclid* and ground photometry

Methods for galaxy cluster weak-lensing analyses

As discussed in Chapt. 3, weak gravitational lensing is currently one of the most reliable methods to accurately and precisely measure cluster masses (Okabe et al. 2010b; Hoekstra et al. 2012; Melchior et al. 2015; Sereno et al. 2017; Stern et al. 2019; Schrabback et al. 2021; Zohren et al. 2022). Consequently, weak-lensing cluster mass estimates are widely used in current photometric galaxy surveys, such as KiDS, DES, and HSC.

An accurate selection of lensed background galaxies is crucial to derive a reliable cluster weak-lensing signal. Including the contribution from foreground and cluster member galaxies may significantly dilute the weak-lensing signal (Broadhurst et al. 2005; Medezinski et al. 2007; Sifón et al. 2015; McClintock et al. 2019). For example, background selections with 90% purity dilute the cluster reduced shear measurements by 10% (see, e.g., Dietrich et al. 2019), in the absence of intrinsic alignments (Heymans & Heavens 2003). Highly pure background selections are required to properly account for this effect in weak-lensing measurements, in order to minimise the variance in the selection purity. Selection incompleteness, instead, impacts the weak-lensing noise and, in turn, the S/N, which depends on the density of background sources along with the intrinsic ellipticity dispersion and measurement noise (see, e.g., Schrabback et al. 2018; Umetsu 2020). The effect of low background densities can be partially mitigated by increasing the size of the cluster-centric radial bins used in the analysis, or through the stacking of the weak-lensing signal of cluster ensembles.

Background selections based on the galaxy photo-z posteriors are commonly used in the literature (Gruen et al. 2014; Applegate et al. 2014; Melchior et al. 2017; Sereno et al. 2017; Bellagamba et al. 2019), as well as galaxy colour selections (Medezinski et al. 2010; Oguri et al. 2012; Medezinski et al. 2018b). These selections can also be combined to significantly improve the background sample completeness and, in turn, the weak-lensing S/N. In fact, colour selections have been demonstrated to help identify galaxies with poorly defined photometric redshifts that would not have been classified as background sources through photo-z selection alone (Covone et al. 2014; Sereno et al. 2017; Bellagamba et al. 2019).

In this Chapter, we present the paper developed by Euclid Collaboration: Lesci et al. (2023). This work aims at deriving colour selections based on the galaxy catalogue developed by Bisigello et al. (2020), hereafter referred to as B20, and extended by Euclid Collaboration: Bisigello et al. (2023). The B20 catalogue includes simulated SDSS (Gunn et al. 1998) *griz* magnitudes and simulated *Euclid* observations in the $Y_{\rm E}J_{\rm E}H_{\rm E}$ bands. In particular, in Euclid

Collaboration: Lesci et al. (2023) we develop a method to obtain optimal colour selections, having maximal completeness given a threshold on purity, given any set of photometric filters. In addition, we provide for the first time colour selections expressed as a continuous function of the lens limiting redshift. This allows for a finer background definition compared to colour selections found in the literature (Medezinski et al. 2010; Oguri et al. 2012; Medezinski et al. 2018b), implying a significant improvement of the weak-lensing source statistics. In addition, we test the efficiency of these colour selections on real public external data and on simulations, combining them with photo-z selections.

The paper by Euclid Collaboration: Lesci et al. (2023) is part of a series presenting and discussing mass measurements of galaxy clusters using the *Euclid* combined clusters and weak-lensing pipeline `COMB-CL`. `COMB-CL` forms part of the global *Euclid* data processing pipeline and is responsible for measuring weak-lensing shear profiles and masses for photometrically detected clusters. A comprehensive description of the code structure and methods employed by `COMB-CL` will be presented in a forthcoming paper, but a brief overview of the pipeline can be found in the appendix of Euclid Collaboration: Sereno et al. (in prep.). The galaxy colour selections presented in this Chapter are already implemented in `COMB-CL`.

This Chapter is organised as follows. In Sect. 6.1, we describe the data set used for the calibration of galaxy colour selections, and in Sect. 6.2 we detail a general method to derive optimal colour selections. In Sect. 6.3, we show the selections obtained for *griz* and $griz Y_{\rm E} J_{\rm E} H_{\rm E}$ filter sets, validating them on external data sets. In Sect. 6.4 we compare the *griz* selection calibrated in this work with selections from the literature. Finally, in Sect. 6.5, we summarise the results and draw our conclusions.

6.1 Calibration sample

We base our analysis on the photometric catalogue developed by B20 and extended by Euclid Collaboration: Bisigello et al. (2023). This catalogue contains simulated *Euclid* $I_{\rm E} Y_{\rm E} J_{\rm E} H_{\rm E}$ aperture magnitudes[1], covering the spectral range 5500–20 000 Å, along with the Canada-France Imaging Survey (CFIS; Ibata et al. 2017) u band, for the galaxies contained in the COSMOS catalogue by Laigle et al. (2016, COSMOS15). Specifically, such photometry is based on 3″ fixed-aperture magnitudes. Despite the u band is already present in COSMOS15, B20 derived it using the same approach adopted for the other filters in order to avoid colour biases. B20 verified that this provides results that are consistent with the observed fluxes. Simulated SDSS *griz* magnitudes, spanning the wavelength range 4000–11 000 Å, are also provided in the catalogue, since observations in similar filters, such as those in *Vera C. Rubin Observatory* (*Rubin*/LSST; Ivezic et al. 2008) and DES, will be available to complement *Euclid* observations (Euclid Collaboration: Pocino et al. 2021; Euclid Collaboration: Scaramella et al. 2022). Corrections for photometric offsets due to flux outside the fixed-aperture, systematic offsets, and Galactic extinction, as suggested in Laigle et al. (2016), have been applied. B20 derive simulated magnitudes through two alternative approaches. The first is a linear interpolation of the 30 medium-band and broad-band filters available in the COSMOS15 catalogue, based on the effective wavelength of the filters. The second approach is based on the best theoretical template that describes the spectral energy distribution (SED) of each galaxy, assuming the COSMOS15 redshifts as the ground truth. The SED fitting is performed based on COSMOS15 bands and the template resulting in the minimum χ^2 is used to predict the expected fluxes. We refer to B20 for the details of the SED templates used,

[1] $I_{\rm E}$ band observations are supplied by the *Euclid* Visible Imager (VIS; Cropper et al. 2016), while $Y_{\rm E} J_{\rm E} H_{\rm E}$ photometry is provided by the Near-Infrared Spectrometer and Photometer (NISP; Euclid Collaboration: Schirmer et al. 2022).

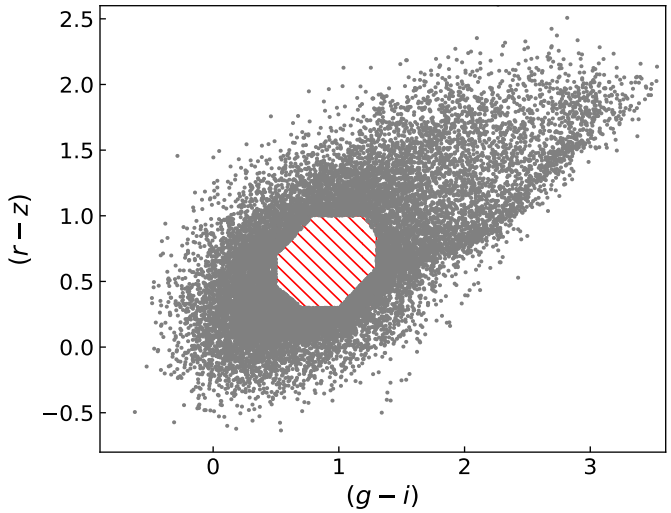

Figure 6.1: Example of an uncalibrated selection in the $(r-z)$ - $(g-i)$ colour-colour space. The grey dots represent the selected galaxy colours. Galaxies within the octagonal hatched region are excluded by applying Eq. (6.1). Specifically, $(r-z)$ and $(g-i)$ correspond to x and y in Eq. (6.1), respectively.

based on the model by Bruzual & Charlot (2003). The expected fluxes are then randomised 10 times considering a Gaussian distribution centred on the true flux and with standard deviation equal to the expected photometric uncertainities, scaled considering the depths listed in Table 1 of Euclid Collaboration: Bisigello et al. (2023). In this process, the $I_{\rm E}Y_{\rm E}J_{\rm E}H_{\rm E}$ magnitude errors expected for the Euclid Wide Survey are considered. Despite the fact that the *griz* photometry is based on SDSS filter transmissions, the corresponding uncertainties are based on depths that are consistent with those of DES and the Ultraviolet Near-Infrared Optical Northern Survey (UNIONS).[2] The *ugriz* photometry provided by LSST is expected to go from 1 to 2.5 magnitudes deeper at the end of the *Euclid* mission, depending on the photometric filter. Throughout this Chapter, we focus on the magnitudes derived from the best theoretical SED templates, as these estimates better reproduce absorption and emission lines that are not covered by COSMOS15 bands. We neglect u magnitudes since, due to the low u-band throughput, a 5σ depth of 25.6 mag will only be reached after 10 years of LSST observations[3]. In addition, the u band is not available in DES wide fields. We emphasise that the B20 catalogue contains all the galaxies present in the COSMOS15 sample, which is deeper than the shear samples derived from current surveys (see, e.g., Giblin et al. 2021; Gatti et al. 2021) and expected from the Euclid Wide Survey (Euclid Collaboration: Scaramella et al. 2022). As we shall discuss in the following, the colour selections calibrated in this study yield robust results against alternative magnitude cuts, including those that reproduce the selections adopted in current and *Euclid* cosmic shear analyses.

6.2 Method

In order to find a set of optimal galaxy colour-redshift relations that maximises the selection completeness given a threshold on the foreground contamination, we consider the colours given by any combination of photometric bands. This includes bands that are not adjacent in

[2]UNIONS is carried out with the Subaru Telescope (Iye et al. 2004), the Canada-France-Hawaii Telescope (CFHT; Gwyn 2012), and the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS; Chambers et al. 2016). More information at `https://www.skysurvey.cc/news/`.

[3]`https://www.lsst.org/scientists/keynumbers`

wavelength. Thus, for each colour-colour space, given a redshift lower limit, z_l, corresponding to the lens redshift, we consider the following set of conditions,

$$\begin{aligned} & x > c_1 \vee \\ & x < c_2 \vee \\ & y > c_3 \vee \\ & y < c_4 \vee \\ & x > s_1 y + c_5 \vee \\ & x > s_2 y + c_6 \vee \\ & x < s_3 y + c_7 \vee \\ & x < s_4 y + c_8, \end{aligned} \tag{6.1}$$

where $\vee$ is the logical "or" operator, x and y are two different colours, and c_i and s_i are colour selection parameters. Specifically, $c_1, \ldots, c_8 \in (-\infty, +\infty)$, s_1 and $s_3 \in (0, +\infty)$, while s_2 and $s_4 \in (-\infty, 0)$. The edges of the aforementioned parameter ranges are excluded, and Eq. (6.1) can potentially define an irregular octagon that contains the foreground galaxies, as we show in Fig. 6.1. As we shall see, since we only select the colour conditions that satisfy given requirements, not all the sides of the irregular octagon may be considered. In addition, since we consider the conditions in Eq. (6.1) as independent, the $c_1, \ldots, c_8$ and $s_1, \ldots, s_4$ parameters are not related to each other. In particular, for each condition in Eq. (6.1), we derive the completeness,

$$\mathcal{C}_i^{nf}(z_\mathrm{l} \,|\, \vec{p}) := \frac{N_{\mathrm{sel},i}(z_\mathrm{g} > z_\mathrm{l} \,|\, \vec{p})}{N_\mathrm{tot}(z_\mathrm{g} > z_\mathrm{l})}\,, \tag{6.2}$$

and the purity,

$$\mathcal{P}_i^{nf}(z_\mathrm{l} \,|\, \vec{p}) := \frac{N_{\mathrm{sel},i}(z_\mathrm{g} > z_\mathrm{l} \,|\, \vec{p})}{N_{\mathrm{sel},i}(z_\mathrm{g} \geq 0 \,|\, \vec{p})}\,, \tag{6.3}$$

where z_g is the galaxy redshift, $\vec{p}$ is the set of colour condition parameters, i is the ith colour condition index, $N_{\mathrm{sel},i}$ is the number of galaxies selected with the ith colour condition, N_tot is the total number of galaxies in the calibration sample, while the nf superscript represents quantities derived from colour conditions not fitted as a function of z_l. As we shall see, we do not adopt any superscripts for the quantities derived from fitted colour conditions. In Eqs. (6.2) and (6.3), we have $i = 1 ... N_\mathrm{cond}$, where N_cond is the number of all possible colour conditions, given Eq. (6.1), expressed as

$$N_\mathrm{cond} = 8\,\frac{N_\mathrm{col}!}{(N_\mathrm{col} - 2)!\,2!}, \tag{6.4}$$

where N_col is the number of colours, given by

$$N_\mathrm{col} = \frac{N_\mathrm{band}!}{(N_\mathrm{band} - 2)!\,2!}, \tag{6.5}$$

where N_band is the number of photometric bands.

We set requirements on completeness and purity to be satisfied by each colour condition in Eq. (6.1). Specifically, for a given z_l, we select the colour conditions having at least one $\vec{p}$ set providing $\mathcal{C}_i^{nf}(z_\mathrm{l} \,|\, \vec{p})$ and $\mathcal{P}_i^{nf}(z_\mathrm{l} \,|\, \vec{p})$ larger than their corresponding thresholds. We remark that $\vec{p}$ does not explicitly depend on z_l at this stage, and that z_l values are arbitrarily sampled. Setting a threshold on $\mathcal{C}_i^{nf}(z_\mathrm{l} \,|\, \vec{p})$ is important for excluding colour conditions that do not significantly contribute to the total completeness, and that may appear as optimal only due to

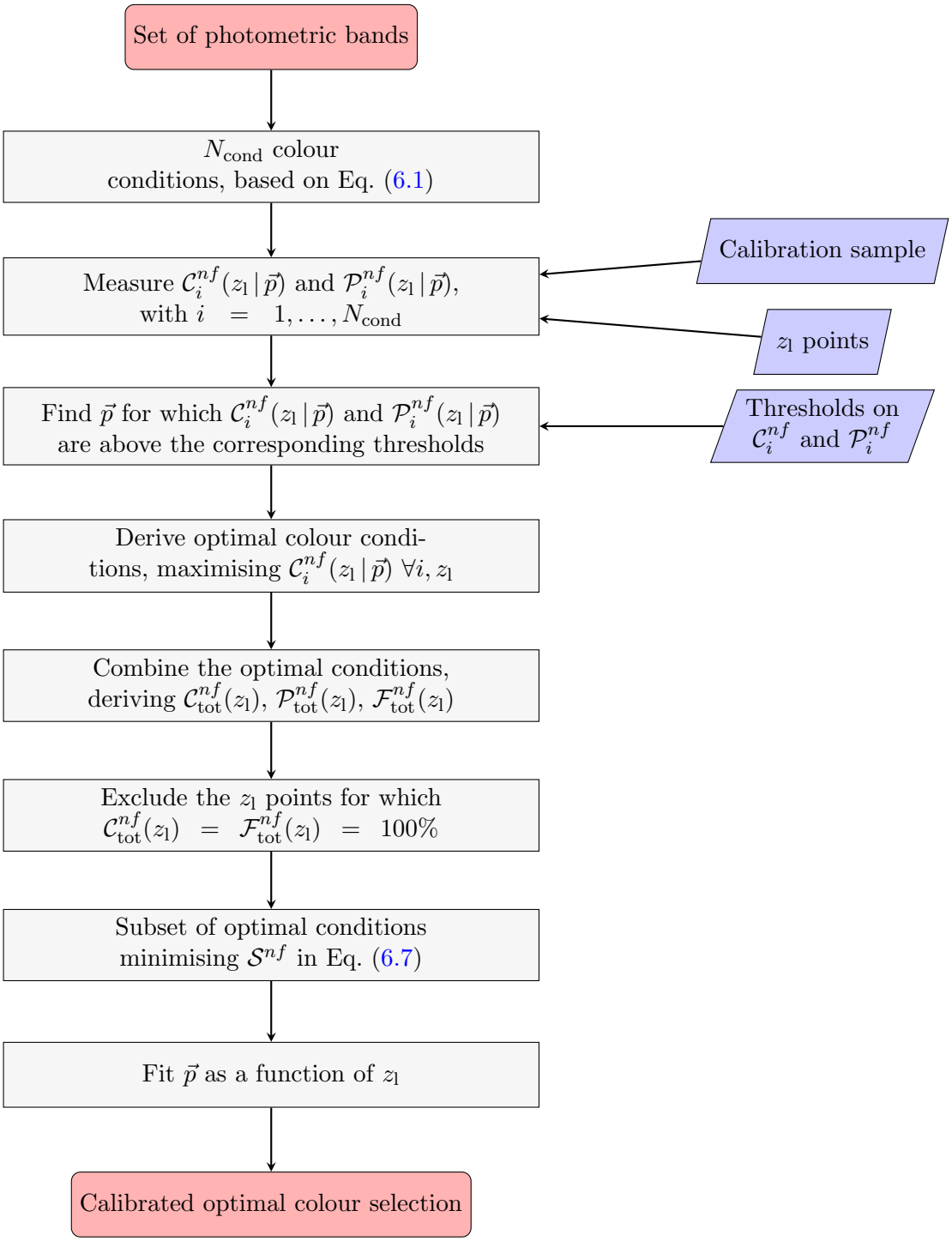

Figure 6.2: Flowchart summarising the calibration process described in Sect. 6.2. Round red rectangles represent the start and end points of the calibration process. Grey rectangles represent processing steps, while blue trapezoids correspond to the inputs.

statistical fluctuations. Thus, the threshold on $\mathcal{C}_i^{nf}(z_{\rm l}\,|\,\vec{p})$ is meant to be low compared to that on $\mathcal{P}_i^{nf}(z_{\rm l}\,|\,\vec{p})$. Indeed, as we shall detail in Sect. 6.3.6, impurities in the background selection imply systematic uncertainties in galaxy cluster reduced shear measurements. Highly pure selections are required to properly account for this effect, in order to minimise the scatter in purity. We discuss the choice of the thresholds on $\mathcal{C}_i^{nf}(z_{\rm l}\,|\,\vec{p})$ and $\mathcal{P}_i^{nf}(z_{\rm l}\,|\,\vec{p})$ in greater

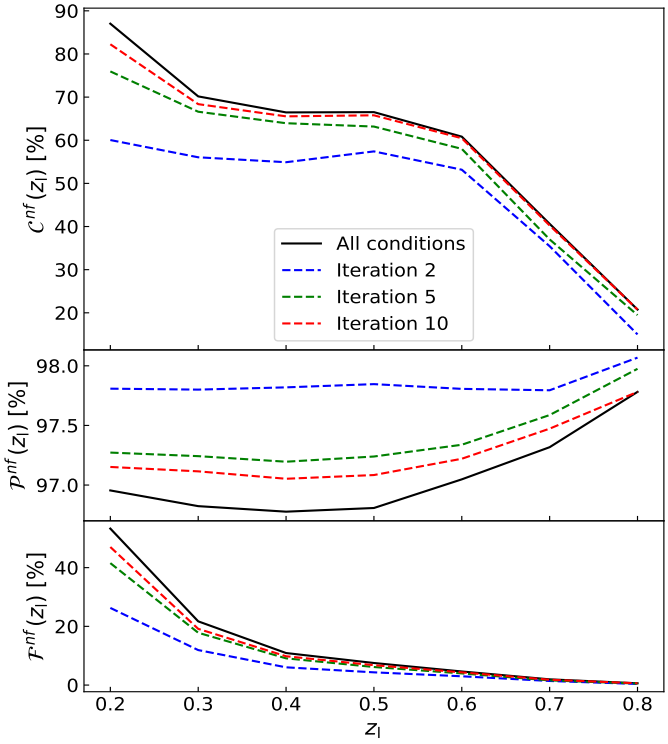

Figure 6.3: Selection completeness (top panel), purity (middle panel), and foreground failure rate (bottom panel), derived from subsets of optimal colour conditions not fitted as a function of z_l, for the case of *griz* photometry. The solid black lines represent the selection given by the full set of optimal colour conditions, while the dashed lines show the selection at different steps of the iterative process detailed in Sect. 6.2, given by subsets of optimal conditions.

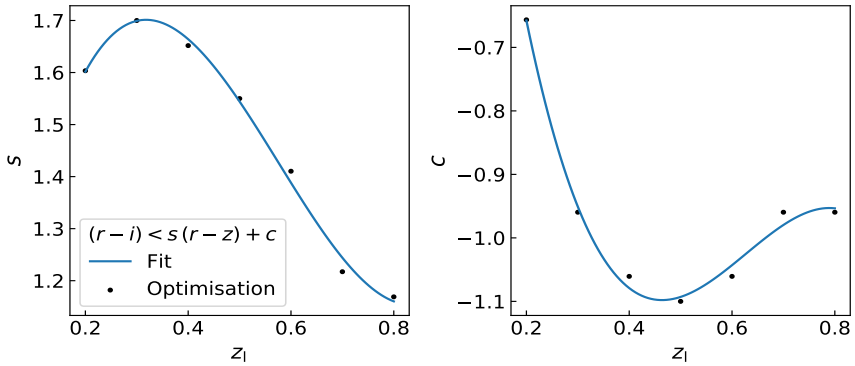

Figure 6.4: Values of s (left panel) and c (right panel) parameters, from Eq. (6.1), as a function of z_l for the colour condition quoted in the left panel legend. The black dots represent the optimal values of s and c, while the blue curves represent the polynomial fits.

deatil in Sect. 6.3.1. For each colour condition in Eq. (6.1), with parameter values for which the conditions on $\mathcal{C}_i^{nf}(z_l\,|\,\vec{p})$ and $\mathcal{P}_i^{nf}(z_l\,|\,\vec{p})$ are satisfied, we select the $\vec{p}$ set providing the highest completeness at a given z_l. In this way, we derive the set of optimal colour conditions maximising the selection completeness, given the chosen threshold on purity.

We note that the maximum z_l of the calibrated colour selections depends on the $\mathcal{C}_i^{nf}(z_l\,|\,\vec{p})$ and $\mathcal{P}_i^{nf}(z_l\,|\,\vec{p})$ limits, while the minimum z_l is derived by excluding the z_l points for which

Table 6.1: Description of the completeness functions introduced in Sect. 6.2. Analogous descriptions hold for purity and foreground failure rate.

Symbol	Description
C_i^{nf}	Completeness of the ith colour condition, given a set of sampled parameters.
$C_{\rm tot}^{nf}$	Completeness derived through the combination of all the optimal colour conditions.
C^{nf}	Completeness given by the combination of a subset of optimal colour conditions.
C	Completeness obtained from a subset of optimal colour conditions fitted as a function of $z_{\rm l}$.

Notes. The nf superscript represents quantities derived from colour conditions not fitted as a function of $z_{\rm l}$. Optimal colour conditions satisfy the thresholds on purity and completeness, and provide maximal completeness.

$\mathcal{C}_{\rm tot}^{nf}(z_{\rm l}) = \mathcal{F}_{\rm tot}^{nf}(z_{\rm l}) = 100\%$. Here, $\mathcal{C}_{\rm tot}^{nf}$ and $\mathcal{F}_{\rm tot}^{nf}$ are the completeness and the foreground failure rate given by the full set of optimal colour conditions, respectively. For simplicity, we drop the dependence on $\vec{p}$ in the text. The foreground failure rate is defined as follows:

$$\mathcal{F}_{\rm tot}^{nf}(z_{\rm l}) := \frac{N_{\rm sel}(z_{\rm g} < z_{\rm l})}{N_{\rm tot}(z_{\rm g} < z_{\rm l})} = \frac{N_{\rm sel}(z_{\rm g} > z_{\rm l})}{N_{\rm tot}(z_{\rm g} < z_{\rm l})} \frac{1 - \mathcal{P}_{\rm tot}^{nf}(z_{\rm l})}{\mathcal{P}_{\rm tot}^{nf}(z_{\rm l})}\,, \tag{6.6}$$

where $N_{\rm sel}$ is the number of galaxies selected with all the optimal colour conditions, given a condition on $z_{\rm g}$, and $\mathcal{P}_{\rm tot}^{nf}(z_{\rm l})$ is the purity given by the full set of optimal conditions. On the right-hand side of Eq. (6.6), derived from Eqs. (6.2) and (6.3), we can see that $\mathcal{F}_{\rm tot}^{nf}(z_{\rm l})$ diminishes with increasing $z_{\rm l}$ if high lower limits on purity are chosen. We stress that $\mathcal{F}_{\rm tot}^{nf}(z_{\rm l}) \leq 1$ by definition.

In the selection process described above, some colour conditions may be redundant. Thus, we iteratively search for an optimal subset of colour conditions to find the minimum number of conditions sufficient to approximately reproduce the required completeness. Specifically, at each step of this iterative process, we compute the following quantity:

$$\mathcal{S}^{nf} = \sum_{j=1}^{N} \mathcal{C}_{\rm tot}^{nf}(z_{{\rm l},j}) - \mathcal{C}^{nf}(z_{{\rm l},j}), \tag{6.7}$$

where N is the number of $z_{\rm l}$ points, $\mathcal{C}_{\rm tot}^{nf}(z_{{\rm l},j})$ is the completeness given by all optimal conditions, while $\mathcal{C}^{nf}(z_{{\rm l},j})$ is the completeness given by a subset of optimal conditions, computed at the jth $z_{\rm l}$ value. As the first step of this iterative process, we find the optimal colour condition minimising $\mathcal{S}^{nf}$. Then, at each iteration, we add the colour condition that, combined with the conditions selected in the previous steps, minimises $\mathcal{S}^{nf}$. We repeat this process until $\mathcal{S}^{nf}$ is lower than a given tolerance. We remark that the logical operator between colour conditions is $\vee$.

Lastly, we apply a nonlinear least squares analysis to find the best fit to the $\vec{p}$ parameters as a function of $z_{\rm l}$ for the subset of optimal colour conditions. We choose the fitting formulae which best reproduce the $z_{\rm l}$ dependence, namely polynomials, while aiming at minimising the number of free parameters in the fit.

In Fig. 6.2 we show a flowchart summarising the calibration process described in this section. In Fig. 6.3 we show an example of the iterative process detailed above, while Fig.

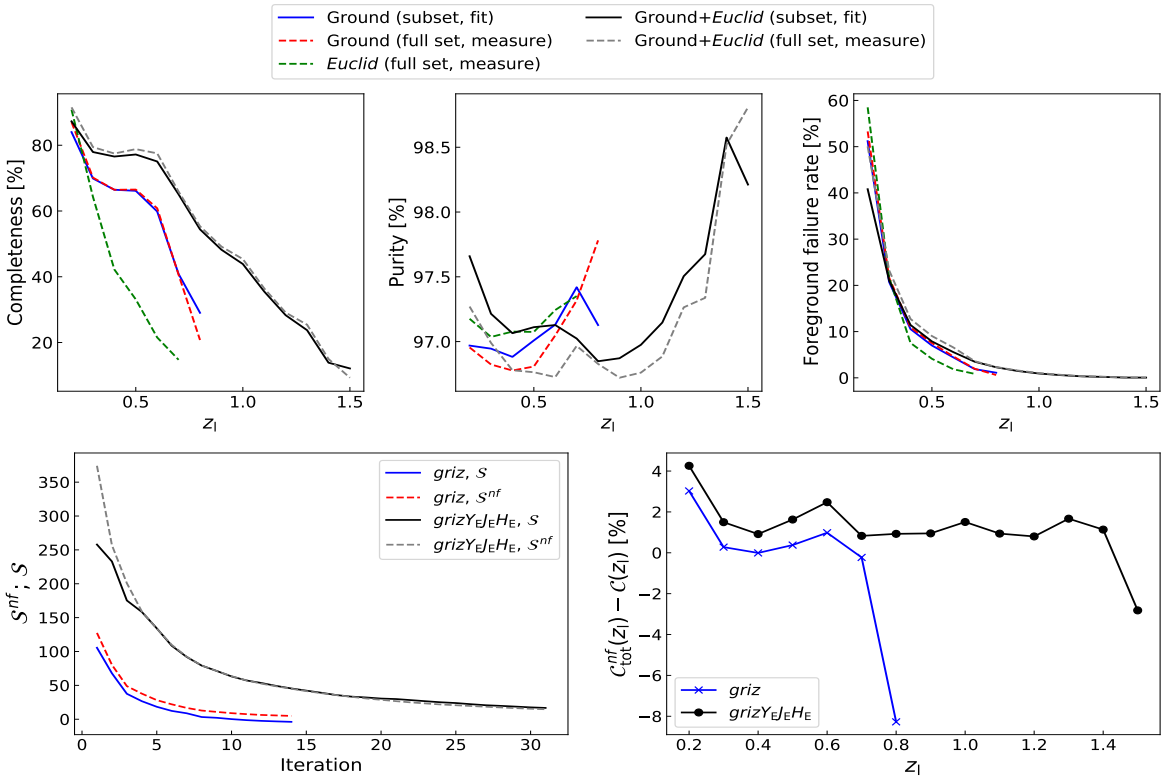

Figure 6.5: Summary of the results on the colour selection optimisation, based on the B20 galaxy sample. *Top panels*: Selection completeness (left panel), purity (central panel), and foreground failure rate (right panel), as a function of z_l. The dashed lines represent the selections derived from the full sets of optimal conditions not fitted as a function of z_l, in the case of ground-only (red), *Euclid*-only (green), and the combination of ground-based and *Euclid* bands (grey). The solid lines represent the selections obtained from the subsets of optimal conditions, with parameters fitted as a function of z_l, in the case of ground-only (blue) and for the combination of ground-based and *Euclid* bands (black). *Bottom panels*: In the left panel, $\mathcal{S}$ and $\mathcal{S}^{nf}$ are shown as a function of the iteration number. For the ground-based selection, using *griz* filters, $\mathcal{S}$ and $\mathcal{S}^{nf}$ are represented by solid blue and dashed red lines, respectively. For the selection derived from the combination of ground-based and *Euclid* filters, namely $grizY_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$, $\mathcal{S}$ and $\mathcal{S}^{nf}$ are represented by solid black and dashed grey lines, respectively. In the right panel, the difference between $\mathcal{C}^{nf}_\mathrm{tot}$ and $\mathcal{C}$ is shown, for the *griz* (blue lines) and $grizY_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ (black lines) selections.

6.4 displays an example of parameter dependence on z_l. Hereafter, we will refer to the completeness, purity, and foreground failure rate, derived from sets of fitted colour conditions, as $\mathcal{C}(z_\mathrm{l})$, $\mathcal{P}(z_\mathrm{l})$, and $\mathcal{F}(z_\mathrm{l})$, respectively. For better clarity, in Table 6.1 we summarise the symbols referring to the completeness functions introduced in this section.

6.3 Results

6.3.1 Calibration of colour selections

By applying the methods detailed in Sect. 6.2 and adopting the B20 calibration sample described in Sect. 6.1, we calibrate galaxy colour selections using ground-based and *Euclid* photometry, namely SDSS *griz* and *Euclid* $Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ filters, respectively. These selections are implemented in `COMB-CL`, and will be available for weak-lensing analyses of galaxy clusters. We consider the following cases: ground-only, *Euclid*-only, and the combination of ground-

based and *Euclid* photometry. For the cases including *Euclid* photometry, we adopt an S/N threshold for *Euclid* near-infrared observations of $(\mathrm{S/N})_{\mathrm{E}} > 3$, which corresponds to $Y_{\mathrm{E}} < 24.85$, $J_{\mathrm{E}} < 25.05$, and $H_{\mathrm{E}} < 24.95$ (Euclid Collaboration: Scaramella et al. 2022). In addition, we consider z_{l} points in the range $z_{\mathrm{l}} \in [0.1, 2.5]$, assuming a precision of $\delta z_{\mathrm{l}} = 0.1$ for the sampling. To derive the full set of optimal colour conditions, we impose $\mathcal{C}_i^{nf}(z_{\mathrm{l}} \,|\, \vec{p}) > 10\%$ for the ith colour condition. For the ground-only and *Euclid*-only cases, we impose that the purity of each colour condition is $\mathcal{P}_i^{nf}(z_{\mathrm{l}} \,|\, \vec{p}) > 99\%$. We adopt a more restrictive threshold on purity for the combination of ground-based and *Euclid* photometry, corresponding to $\mathcal{P}_i^{nf}(z_{\mathrm{l}} \,|\, \vec{p}) > 99.7\%$. This threshold is chosen as the larger number of colour combinations leads to a higher summation of impurities. We obtain $\mathcal{P}^{nf}(z_{\mathrm{l}}) > 97\%$ for any z_{l}, when combining all the optimal colour conditions, as shown in Fig. 6.5. As we shall discuss in Sect. 6.3.3, the purity derived from different real data sets is stable, showing sub-percent changes, on average.

The $\mathcal{F}^{nf}(z_{\mathrm{l}})$ decrease with increasing z_{l}, shown in Fig. 6.5, is expected, as discussed in Sect. 6.2. In addition, for any combination of photometric bands, we find $\mathcal{C}_{\mathrm{tot}}^{nf}(z_{\mathrm{l}}) = \mathcal{F}_{\mathrm{tot}}^{nf} = 100\%$ for $z_{\mathrm{l}} = 0.1$. Consequently, we set $z_{\mathrm{l}} = 0.2$ as the minimum lens redshift for the calibrated colour selections. As shown in Fig. 6.5, from *griz* photometry we derive a selection within $z_{\mathrm{l}} \in [0.2, 0.8]$, with 84% completeness at $z_{\mathrm{l}} = 0.2$, decreasing to 29% at $z_{\mathrm{l}} = 0.8$. In the *Euclid*-only case, namely $Y_{\mathrm{E}}J_{\mathrm{E}}H_{\mathrm{E}}$ and I_{E} bands, results are not competitive with those derived from *griz* photometry. On the other hand, by combining ground-based and *Euclid* photometry, the completeness significantly increases in the z_{l} range covered by the *griz* selection, by up to 25 percent points. Also the z_{l} range of the selection is significantly extended compared to the *griz* case, corresponding to $z_{\mathrm{l}} \in [0.2, 1.5]$. Specifically, in this case we exclude the *Euclid* I_{E} band, as it covers a large wavelength interval, namely $\sim$5000–10 000 Å, corresponding to the wavelength range already covered by *griz* photometry. Furthermore, the use of very broad photometric bands is not the most optimal choice for calibrating galaxy colour selections, which share similarities with photo-z estimates.

We exclude any possible redundant colour condition, as detailed in Sect. 6.2. In Table A.1 we show the subset of optimal colour conditions for the ground-only case, namely *griz* photometry, along with the corresponding parameter fits. The first condition quoted in Table A.1 corresponds to the one derived in the first step of the iterative process described in Sect. 6.2. This is analogous for the subsequent conditions. We remark that the quoted conditions have different ranges of validity in z_{l}. Analogous information is listed in Table A.2 for the combination of ground-based and *Euclid* photometry, corresponding to $grizY_{\mathrm{E}}J_{\mathrm{E}}H_{\mathrm{E}}$ filters. We neglect the optimisation and parameter fitting for the *Euclid*-only case, as we have already shown that it does not provide competitive completeness values.

In Fig. 6.5 we show the results for the selections obtained from the subsets of optimal conditions, with parameters fitted as a function of z_{l}. For both *griz* and $grizY_{\mathrm{E}}J_{\mathrm{E}}H_{\mathrm{E}}$ photometry, such fitted selections well reproduce those given by the full sets of optimal conditions. To quantify the goodness of the colour condition parameter fits, we define a parameter analogous to $\mathcal{S}^{nf}$ in Eq. (6.7), namely $\mathcal{S}$. This parameter quantifies the difference between $\mathcal{C}_{\mathrm{tot}}^{nf}$, that is the completeness given by the full set of optimal conditions not fitted as a function of z_{l}, and $\mathcal{C}$, which is the completeness given by the subset of optimal colour conditions fitted as a function of z_{l}. As shown in Fig. 6.5, $\mathcal{S}$ does not perfectly match $\mathcal{S}^{nf}$, for both *griz* and $grizY_{\mathrm{E}}J_{\mathrm{E}}H_{\mathrm{E}}$ selections. This is due to the fact that the $c_1, \ldots, c_8$, $s_1, \ldots, s_4$ parameters in Eq. (6.1) do not always show a simple dependence on z_{l}. Despite the fact that better parameter fits could be achieved by adopting an arbitrarily high order polynomial as the model, we set a 4th order polynomial as the highest-degree functional form for describing these parameters (see Tables A.1 and A.2). As shown in Fig. 6.5, $\mathcal{C}$ is underestimated by at

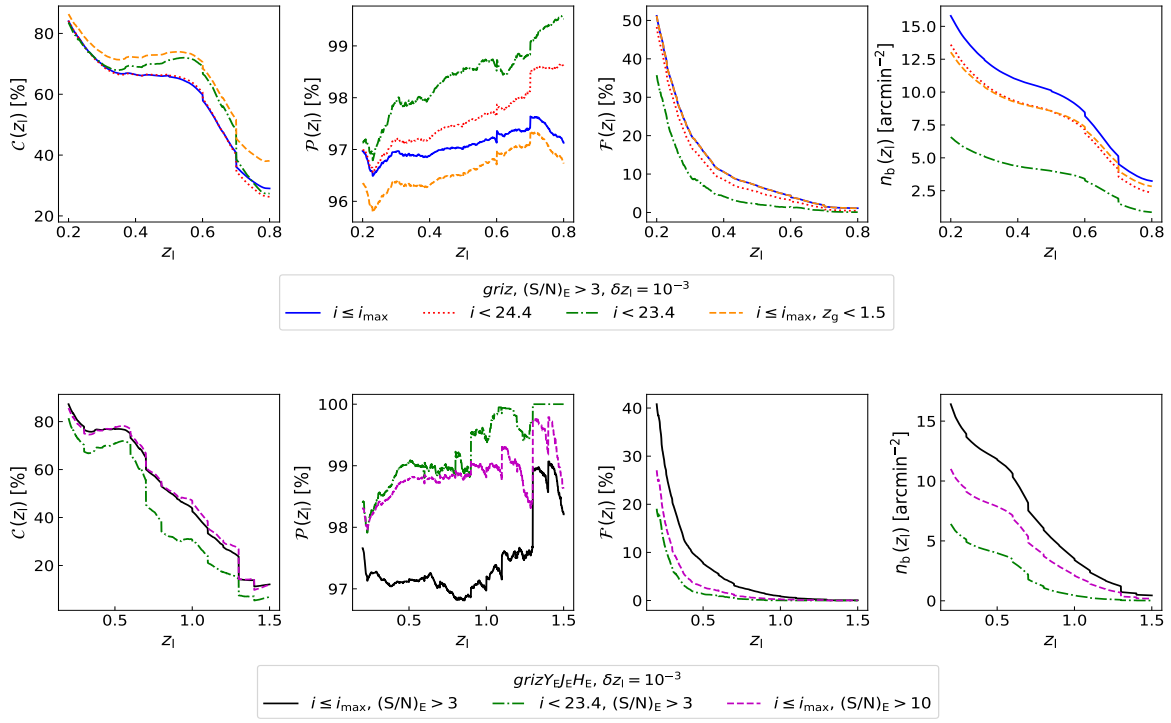

Figure 6.6: From left to right: completeness, purity, foreground failure rate, and background density as a function of z_l, from the fitted colour selections derived in Sect. 6.3.1. The assumed z_l precision is $\delta z_\mathrm{l} = 10^{-3}$. *Top panels:* efficiency of the *griz* selection, detailed in Table A.1, applied to the B20 catalogue with $(\mathrm{S/N})_\mathrm{E} > 3$ and $i \leq i_\mathrm{max}$ (blue solid lines), to its subsample including galaxies with $i < 24.4$ (red dotted lines), to the case with $i < 23.4$ (green dash-dotted lines), and to the sample with $z_\mathrm{g} < 1.5$ (orange dashed lines). *Bottom panels:* efficiency of the $grizY_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selection, detailed in Table A.2, applied to the B20 catalogue with $(\mathrm{S/N})_\mathrm{E} > 3$ and $i \leq i_\mathrm{max}$ (black solid lines), to its subsample with $i < 23.4$ (green dash-dotted lines), and to the subsample with $(\mathrm{S/N})_\mathrm{E} > 10$ (magenta dashed lines).

most 4 percent points. We verify that adding further conditions to these selections, that is, lowering the $\mathcal{S}$ threshold down to 0, provides sub-percent level improvements in the selection completeness, on average. We remark that, in order to derive colour selections not defined in z_l bins, the final selection completeness is slightly degraded compared to $\mathcal{C}^{nf}_\mathrm{tot}$ for some z_l values. In realistic cluster weak-lensing analyses, however, we expect this to statistically increase the galaxy background completeness. When colour selections are defined on finite sets of z_l points, the background galaxies are excluded based on the z_l precision adopted in the colour selection calibration.

6.3.2 Dependence on magnitude and redshift selections

To verify the robustness of the *griz* selection with respect to alternative magnitude cuts, we apply the selection $i < 24.4$, corresponding to the peak value of the i magnitude distribution in the B20 catalogue. We also investigate the selection for the subsample with $i < 23.4$, which is a threshold similar to the DES i band limit (Sevilla-Noarbe et al. 2021). In both cases, we derive higher $\mathcal{P}(z_\mathrm{l})$ and lower $\mathcal{F}(z_\mathrm{l})$, compared to what we find from the calibration sample used in Sect. 6.3.1, namely the one with $(\mathrm{S/N})_\mathrm{E} > 3$ and $i \leq i_\mathrm{max}$, where $i_\mathrm{max} = 24.9$ is the maximum i magnitude in the sample (see Fig. 6.6). In the case with $i < 24.4$, $\mathcal{C}(z_\mathrm{l})$ is close to that from the calibration sample, while for $i < 23.4$ we derive higher completeness,

on average. In addition, as the bulk of the redshift distribution in the calibration sample, described in Sect. 6.1, extends up to $z_g \sim 4$, we apply the *griz* selection to the galaxy sample with redshift $z_g < 1.5$, $(S/N)_E > 3$, and $i \leq i_{max}$. In Fig. 6.6, we can see that this redshift limit provides $\mathcal{F}(z_l)$ values that are identical to those derived from the calibration sample, which is expected since $\mathcal{F}(z_l)$ does not depend on the maximum redshift of the sample, while the completeness increases by up to 10 percent points and the purity is at most 1 percent point lower. Note that the computation of $\mathcal{C}(z_l)$ and $\mathcal{F}(z_l)$ is made relative to the sample under consideration. In other words, they refer to galaxy populations defined by given magnitude and redshift limits. We measure the aforementioned colour selections by assuming a z_l precision of $\delta z_l = 10^{-3}$. This δz_l value is one order of magnitude lower (i.e. one order of magnitude higher precision) than the typical galaxy cluster photometric redshift uncertainty in current surveys (see, e.g., Rykoff et al. 2016; Maturi et al. 2019) and *Euclid* (Euclid Collaboration et al. 2019). Consequently, the $\delta z_l = 10^{-3}$ precision ensures the reliability of the colour condition fits for galaxy cluster background selections. We remark that we assumed $\delta z_l = 0.1$ for the z_l sampling in the calibration process.

In Fig. 6.6 we show the efficiency of the $grizY_EJ_EH_E$ selection, computed by adopting $\delta z_l = 10^{-3}$, applied to the B20 calibration sample, with $(S/N)_E > 3$ and $i \leq i_{max}$. We find analogous selections from the subsample with $i < 23.4$ and from the one with $(S/N)_E > 10$. Specifically, in both cases, we derive higher $\mathcal{P}(z_l)$ and lower $\mathcal{F}(z_l)$, in agreement with what we find from the *griz* selection. In addition, the increase in the minimum *Euclid* S/N does not significantly change the completeness, while the $i < 23.4$ limit decreases $\mathcal{C}(z_l)$ by at most 18 percent points. As we obtain excellent $\mathcal{P}(z_l)$ and $\mathcal{F}(z_l)$ estimates from these tests, we conclude that both *griz* and $grizY_EJ_EH_E$ selections are stable and reliable with respect to changes in the sample limiting magnitude and redshift. In addition, we note that brighter galaxy samples provide lower foreground contamination. This is expected, as faint galaxies have more scattered colour-redshift relations.

In Fig. 6.6 we show the density of background galaxies, $n_b(z_l)$, defined as the number of selected galaxies with $z_g > z_l$ per square arcmin. For both *griz* and $grizY_EJ_EH_E$ selections, $n_b(z_l) = 16$ arcmin^{-2} at $z_l = 0.2$ for $i \leq i_{max}$ and $(S/N)_E > 3$, decreasing with increasing z_l. In both colour selections, the $i < 23.4$ limit implies the largest decrease in $n_b(z_l)$, providing $n_b(z_l) < 7$ arcmin^{-2}. In addition, for the *griz* selection, the $i < 24.4$ and $z_g < 1.5$ limits provide consistent results on $n_b(z_l)$, showing a difference of at most 3 arcmin^{-2} compared to that derived from the calibration sample. With regard to the $grizY_EJ_EH_E$ selection, the $(S/N)_E > 10$ limit implies a decrease in $n_b(z_l)$ of up to 5 arcmin^{-2} at low z_l, while $n_b(z_l)$ becomes compatible with that derived from the calibration sample for $z_l > 1$.

6.3.3 *griz* selection validation on real data

To further assess the reliability of the *griz* colour selection detailed in Sect. 6.3.1, we apply it to external data sets obtained from real observations. In particular, we consider the VIMOS Public Extragalactic Redshift Survey (VIPERS; Guzzo et al. 2014) Multi-Lambda Survey (VMLS) photometric catalogue by Moutard et al. (2016), including Canada-France-Hawaii Telescope Legacy Survey (CFHTLS; Hudelot et al. 2012) *griz* Kron aperture magnitudes (Kron 1980). This catalogue covers 22 deg^2 and provides reliable photometric redshifts for more than one million galaxies with a typical accuracy of $\sigma_z \leq 0.04$, and a fraction of catastrophic failures lower than 2% down to $i \sim 23$. These statistics are based on VIPERS data, complemented with the most secure redshifts selected from other spectroscopic surveys. We remind that in VIPERS a colour-colour pre-selection was employed to enhance the effective sampling of the VIMOS spectrograph. Nevertheless, the VIPERS selection does not introduce any significant colour bias above z $\sim$ 0.6 (Guzzo et al. 2014). In addition, as we shall see

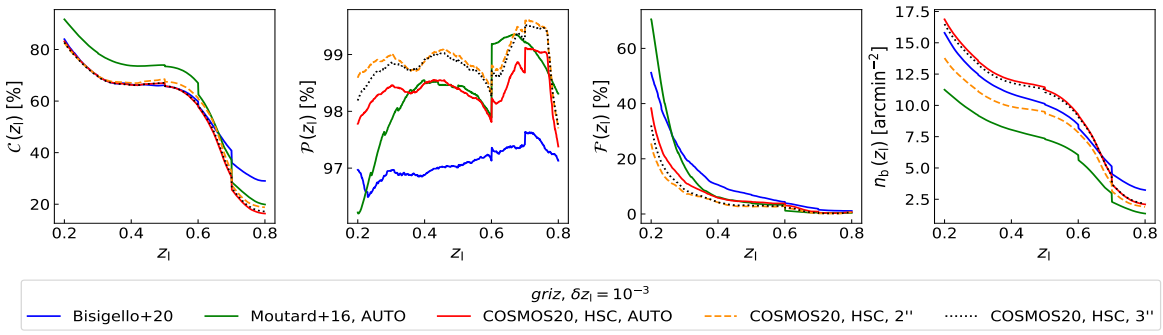

Figure 6.7: From left to right: completeness, purity, foreground failure rate, and background density as a function of z_l, from the fitted colour selection based on *griz* photometry, reported in Table A.1. The assumed z_l precision is $\delta z_\mathrm{l} = 10^{-3}$. The *griz* selection is applied to the B20 catalogue with magnitude limits corresponding to those used in the calibration process (blue solid lines), to the full depth Moutard et al. (2016) catalogue (green solid lines), and to the Weaver et al. (2022) catalogue with HSC Kron, 2″ and 3″ aperture magnitudes (solid red, dashed orange and dotted black lines, respectively), for which we imposed $i < 25$.

in the following, the selection completeness and purity obtained from the VMLS data set do not exhibit remarkable deviations from those obtained from other galaxy samples. In Fig. 6.7 we can see that, by applying the *griz* selection to the VMLS sample, we derive higher $\mathcal{P}(z_\mathrm{l})$ and lower $\mathcal{F}(z_\mathrm{l})$ compared to what we find from the B20 catalogue, on average. This agrees with what we found in Sect. 6.3.2, as the Moutard et al. (2016) catalogue is shallower than the B20 sample. For the same reason, $n_\mathrm{b}(z_\mathrm{l})$ is 3 arcmin^{-2} lower, on average. In addition, the completeness is up to 8 percent points higher for $z_\mathrm{l} < 0.6$, becoming lower for higher z_l values.

We also apply the *griz* selection to the COSMOS `CLASSIC` catalogue by Weaver et al. (2022, COSMOS20), which reaches the same photometric redshift precision as COSMOS15, namely $\sigma_z/(1+z) = 0.007$, at almost one magnitude deeper. We consider *griz* Kron, 2″ and 3″ aperture magnitudes from HSC. In addition, we select galaxies with a photometric redshift derived from at least 30 bands, and with $i < 25$, in order to consider a sample with highly reliable redshift estimates. By adopting more complex selection criteria, which may involve galaxies with photometric redshifts derived from a shared set of photometric bands, we do not expect remarkable differences in the results. Similar results for the cases with Kron, 2″ and 3″ aperture magnitudes are shown in Fig. 6.7. Compared to what we derive from the B20 sample, the completeness is similar, with the largest differences at $z_\mathrm{l} > 0.6$. In addition, $\mathcal{F}(z_\mathrm{l})$ is lower and $\mathcal{P}(z_\mathrm{l})$ is higher for any z_l. For Kron and 3″ aperture magnitudes, $n_\mathrm{b}(z_\mathrm{l})$ is slightly higher compared to that obtained from the B20 sample, on average. Lower $n_\mathrm{b}(z_\mathrm{l})$ values show up for the 2″ aperture magnitudes, which is expected as we apply the same magnitude limit for each photometric aperture definition. Indeed, for these tests we did not include aperture correction terms. Lastly, comparing the purity derived from the COSMOS20 and VMLS samples, we note that for $z_\mathrm{l} > 0.3$ the differences are below 1 percent point, on average. Thus, we conclude that the *griz* selection provides robust and reliable results on real data.

6.3.4 Validation on Flagship v2.1

We test the colour selections calibrated in Sect. 6.3.1 on the *Euclid* Flagship galaxy catalogue v2.1.10 (Euclid Collaboration in prep.), which is currently the best simulated *Euclid* galaxy catalogue available. This catalogue is based on an N-body simulation with around 4 tril-

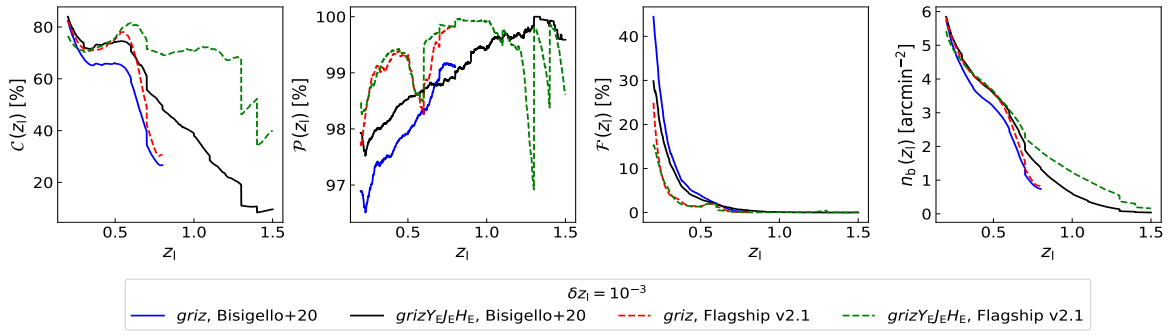

Figure 6.8: From left to right: completeness, purity, foreground failure rate, and background density as a function of z_l, from the fitted colour selections based on *griz* and *griz*$Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ bands, adopting 5σ magnitude limits. The assumed z_l precision is $\delta z_\mathrm{l} = 10^{-3}$. The solid blue and solid black lines represent the *griz* and *griz*$Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selections, respectively, applied to the B20 catalogue. The dashed red and dashed green curves represent the *griz* and *griz*$Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selections, respectively, applied to the Flagship v2.1 catalogue (Euclid Collaboration in prep.).

lion particles with mass $m_\mathrm{p} \sim 10^9\ h^{-1}\mathrm{M}_\odot$. A flat ΛCDM cosmological model was assumed, with matter density parameter $\Omega_\mathrm{m} = 0.319$, baryon density parameter $\Omega_\mathrm{b} = 0.049$, dark energy density parameter $\Omega_\Lambda = 0.681$, scalar spectral index $n_\mathrm{s} = 0.96$, Hubble parameter $h = H_0/(100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}) = 0.67$, and standard deviation of linear density fluctuations on 8 h^{-1}Mpc scales $\sigma_8 = 0.83$. The haloes were identified using `Rockstar` (Behroozi et al. 2013), and then populated with a halo occupation distribution model which was calibrated to reproduce observables such as clustering statistics as a function of galaxy luminosity. The galaxy SED templates used are the COSMOS templates from Ilbert et al. (2009), based on the models by Bruzual & Charlot (2003) and Polletta et al. (2007). In addition, galaxy photo-z probability distribution functions, namely $p(z_\mathrm{g})$, are included in Flagship, derived through a Nearest Neighbours Photometric Redshifts (NNPZ) pipeline (Euclid Collaboration: Desprez et al. 2020).

From the Flagship catalogue, we extracted a lightcone within RA $\in [158°, 160°]$ and Dec $\in [12°, 15°]$, considering the galaxies in the whole redshift range covered by the simulation, namely $z_\mathrm{g} \in [0, 3]$. Specifically, z_g is the galaxy true redshift, and we verified that the contribution of peculiar velocities does not significantly change the results. We focused on $2''$ aperture LSST *ugrizy* and *Euclid* $I_\mathrm{E}Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ photometry, as the simulated fluxes estimated for other ground-based surveys do not account for observational noise. Specifically, the photometric noise takes into account the depth expected in the southern hemisphere at the time of the third data release (DR3) for the Euclid Wide Survey. The LSST and *Euclid* 10σ magnitude limits, which are proxies for extended sources, correspond to $u < 24.4$, $g < 25.6$, $r < 25.7$, $i < 25.0$, $z < 24.3$, $y < 23.1$, $I_\mathrm{E} < 25$, $Y_\mathrm{E} < 23.5$, $J_\mathrm{E} < 23.5$, and $H_\mathrm{E} < 23.5$. The fluxes we considered are not reddened due to Milky Way extinction, consistent with the analyses performed in the previous sections.

In Fig. 6.8, we show the application of *griz* and *griz*$Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selections to Flagship. For this test, we assumed 5σ magnitude cuts for LSST *ugrizy* and *Euclid* $I_\mathrm{E}Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ bands. In addition, we show results from the B20 sample in Fig. 6.8, for which we assumed 5σ magnitude cuts rescaled from the 10σ limits listed in Euclid Collaboration: Bisigello et al. (2023, Table 1). We find that $n_\mathrm{b}(z_\mathrm{l})$ derived from Flagship agrees with that obtained from the B20 sample. The largest differences, of about 1 arcmin^{-2}, arise when the *griz*$Y_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selection is applied. We note that $n_\mathrm{b}(z_\mathrm{l}) \sim 0$ for $z_\mathrm{l} \sim 1.5$, implying that lenses at these values of z_l may not exhibit significant weak-lensing signals. Nevertheless, we verified that $n_\mathrm{b}(z_\mathrm{l})$ is enhanced

at any z_l when the selection defined for *Euclid* weak-lensing analyses (Laureijs et al. 2011; Euclid Collaboration: Scaramella et al. 2022) is assumed. This selection consists in a 10σ cut in the I_E band, corresponding to $I_\mathrm{E} < 25$ for a $2''$ aperture, yielding a galaxy density of around 39 arcmin^{-2} when applied to the Flagship data set. In fact, in this case $n_\mathrm{b}(z_\mathrm{l})$ ranges from 30 arcmin^{-2} at low z_l to 3 arcmin^{-2} at $z_\mathrm{l} = 1.5$.

As shown in Fig. 6.8, on average we obtain higher $\mathcal{P}(z_\mathrm{l})$ and lower $\mathcal{F}(z_\mathrm{l})$ for $z_\mathrm{l} < 1$ from Flagship, compared to what we derive from the B20 sample. For the *griz* selection case, $\mathcal{C}(z_\mathrm{l})$ agrees with that derived from the B20 sample, with the largest differences, of up to 16 percent points, at $z_\mathrm{l} \sim 0.5$. Larger differences in $\mathcal{C}(z_\mathrm{l})$ are obtained from the $grizY_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selection. From Flagship we obtain $\mathcal{C}(z_\mathrm{l})$ up to 10 percent points larger for $z_\mathrm{l} < 0.6$, and up to 50 percent points larger for higher z_l. We verified that this discrepancy in the completeness, in the case of the $grizY_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selection, is not significantly attenuated through the assumption of 3σ and 10σ magnitude limits on both B20 and Flagship catalogues. Analogous results were obtained by assuming limits corresponding to the magnitude distribution peaks derived from the B20 catalogue, namely $g < 24.9$, $r < 24.6$, $i < 24.3$, $z < 24.1$, $Y_\mathrm{E} < 23.8$, $J_\mathrm{E} < 23.6$, and $H_\mathrm{E} < 23.5$. Moreover, we verified that the $grizY_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selection completeness does not remarkably vary by assuming the *Euclid* weak-lensing selection defined above, namely $I_\mathrm{E} < 25$. Similar results are obtained by considering the photometric errors expected for the DR2 of the Euclid Wide Survey, assuming the corresponding 3σ, 5σ, and 10σ magnitude cuts. For each of the alternative magnitude cuts discussed in this section, we find that the $grizY_\mathrm{E}J_\mathrm{E}H_\mathrm{E}$ selection yields a purity up to 3 percent points higher at $z_\mathrm{l} > 1.2$ when it is applied to the B20 catalogue, compared to what is derived from Flagship. At $z_\mathrm{l} < 1.2$, instead, the purity obtained from B20 is 1 percent point lower, on average. Furthermore, the alternative magnitude cuts do not remarkably impact the selection purity at any z_l.

We additionally adopted SDSS fluxes, which do not include photometric noise, in place of LSST fluxes in Flagship. In this case, the completeness is up to 35 percent points larger than that derived from the B20 sample, and the purity approaches 100% for $z_\mathrm{l} > 1$, which is similar to what we derive from the B20 sample (see Fig. 6.8). Thus, the selection based on SDSS photometry is less complete and purer compared to that obtained from LSST magnitudes.

Differences in the completeness derived from the Flagship and B20 samples may originate from distinct assumptions on the physical properties of the galaxies, such as dust extinction, stellar age, nebular emission lines, or on the assumed intrinsic spectral energy distributions. This could be indicated by a different fraction of star forming galaxies in the two samples. Following B20, galaxies are classified as star forming if the following condition is satisfied,

$$\log_{10}(\mathrm{sSFR}/\mathrm{yr}^{-1}) > -10.5, \tag{6.8}$$

where sSFR is the specific star formation rate, derived from the best SED template in the catalogue by B20. We verified that, for $z_\mathrm{g} > 1$, the fraction of star forming galaxies in Flagship is consistent within 1 percent point with that derived from the catalogue by B20. Thus, we conclude that the completeness differences between the Flagship and B20 samples are not due to different star forming galaxy populations. We also verified that the $\log_{10}(\mathrm{sSFR}/\mathrm{yr}^{-1})$ distributions derived from the two data sets are compatible, having peaks at ~ -8.13 and ~ -8.35 in B20 and Flagship, respectively. The agreement of these peak values is well within 1σ of the $\log_{10}(\mathrm{sSFR}/\mathrm{yr}^{-1})$ distributions. We will be able to further investigate such completeness differences through the analysis of the first data release of the Euclid Deep Survey.

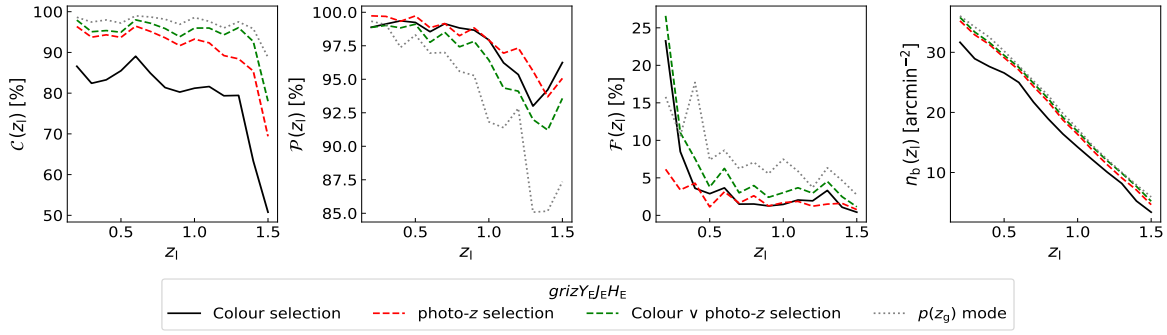

Figure 6.9: From left to right: completeness, purity, foreground failure rate, and background density as a function of z_l, obtained from Flagship v2.1. The solid black lines represent the $grizY_EJ_EH_E$ selection. The dashed green lines show the combination, through the $\vee$ logical operator, of $grizY_EJ_EH_E$ and photo-z selection (Eq. 7.2). The dashed red lines represent the photo-z selection, while the dotted black lines represent the selection based on the $p(z_g)$ mode.

6.3.5 Comparison with photo-z selections

To compare the colour selections derived in this work to selections based on the shape of the galaxy $p(z_g)$, commonly referred to as photo-z selections, we analyse the Flagship sample described in Sect. 6.3.4. We consider only the galaxies with a $p(z_g)$ estimate obtained with the NNPZ pipeline (Euclid Collaboration: Desprez et al. 2020). The NNPZ photo-zs are designed to work well for galaxies that are expected to be used in core *Euclid* weak-lensing science, namely with 5σ limits on the I_E band. Thus we impose $I_E < 25.75$, along with 5σ limits on the $Y_EJ_EH_E$ bands, namely $Y_E < 24.25$, $J_E < 24.25$, $H_E < 24.25$. Specifically, we adopt the following photo-z selection,

$$z_g^{\rm min} > z_l\,, \tag{6.9}$$

where $z_g^{\rm min}$ is the minimum of the interval containing 95% of the probability around the first mode of $p(z_g)$, namely $\overline{z}_g$. We choose $z_g^{\rm min}$ in order to derive $\mathcal{P}(z_l)$ values which are compatible with those obtained from colour selections. We verified that adding a condition on the width of $p(z_g)$ in Eq. (7.2) does not impact the results. Specifically, for the latter test, we considered the additional condition $\mathcal{A} > \mathcal{A}_{\rm min}$, where $\mathcal{A}$ is the integrated probability around $\overline{z}_g$, computed within the redshift points, which are the closest to $\overline{z}_g$, having an associated probability of $0.2p(\overline{z}_g)$. We verified that imposing $\mathcal{A}_{\rm min} = 0$ or $\mathcal{A}_{\rm min} = 0.8$ leads to compatible purity values with sub-percent differences on average. However, $\mathcal{A}_{\rm min} = 0.8$ lowers the photo-z selection completeness by around 20 percent points at all z_l. Consequently, we assume $\mathcal{A}_{\rm min} = 0$.

To perform a fair comparison of colour and photo-z selections, we consider only the $grizY_EJ_EH_E$ colour selection in this section. This is because photo-zs in Flagship were derived from the combination of ground-based and *Euclid* photometry. In Fig. 6.9, we show that the $grizY_EJ_EH_E$ selection provides, on average, a completeness 15 percent points lower than that of the photo-z selection, with similar contamination. By combining $grizY_EJ_EH_E$ and photo-z selections, through the logical operator $\vee$, the completeness increases by up to 10 percent points with respect to the case of photo-z selection alone, amounting to $\mathcal{C}(z_l) \sim 95\%$ for $z_l < 1.4$. These preliminary tests confirm the importance of the combination of colour and photo-z selections, as it leads to significantly more complete background galaxy samples. We also remark that increasing the selection completeness is key to reduce biases in the shear calibration parameters due to background selections, as we shall detail in Sect. 6.3.6. The analysis of Euclid Deep Survey data will allow for a detailed investigation of the optimal

photo-z selections for galaxy cluster weak-lensing analyses, outlining the synergies with colour selections. For example, colour selections applied to *Euclid* data could provide more robust background samples for massive or nearby galaxy clusters, as derived by Medezinski et al. (2018b). Leveraging colour selections also serves as a valuable cross-validation method for addressing the effect of unknown systematic uncertainties in photo-z estimates. Lastly, Fig. 6.9 shows the selection based only on the first mode of $p(z_{\rm g})$. Specifically, in this case we select the galaxies with $\overline{z}_{\rm g} > z_{\rm l}$. Despite $\mathcal{C}(z_{\rm l}) > 90\%$ at all $z_{\rm l}$, the purity is up to 10 percent points lower than that obtained from the $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selection.

6.3.6 Impact on shear measurements

In cluster weak-lensing analyses, the inclusion of foreground sources in the shear measurements may significantly dilute the signal (Broadhurst et al. 2005; Medezinski et al. 2007; Sifón et al. 2015; McClintock et al. 2019). As discussed in the previous sections, the calibrated colour selections provide $\mathcal{P}(z_{\rm l}) < 1$. To assess the impact of impurities on shear measurements, we express the cluster reduced tangential shear unaffected by contamination as follows (Dietrich et al. 2019):

$$g_{\rm t,true}(z_{\rm l}) = \frac{g_{\rm t}(z_{\rm l})}{\mathcal{P}(z_{\rm l})}\,, \tag{6.10}$$

where $g_{\rm t}(z_{\rm l})$ is the measured cluster reduced tangential shear at redshift $z_{\rm l}$. As the calibrated colour selections yield $\mathcal{P}(z_{\rm l}) > 0.97$, we expect at most a 3% bias on the reduced tangential shear. In addition, as discussed in Sect. 6.3.3, $\mathcal{P}(z_{\rm l})$ derived from different observed data sets with only ground-based photometry shows a scatter below 1 percent point. This scatter in $\mathcal{P}(z_{\rm l})$ is lower than the systematic uncertainty on galaxy shape measurements for stage III surveys, as we shall discuss in the following. We remark that $\mathcal{P}(z_{\rm l})$ is derived from reference fields, while galaxy clusters are overdense compared to the cosmic mean. Thus, contamination from cluster galaxies must be properly accounted for in Eq. (6.10) (see, e.g., Gruen et al. 2014; Dietrich et al. 2019). Nevertheless, such contamination is consistent with zero in the typical cluster-centric radial range adopted for mass calibration, namely at radii larger than 300 h^{-1}kpc (see, e.g., Medezinski et al. 2018b; Bellagamba et al. 2019).

Furthermore, galaxy shear calibration is usually statistically derived, based on observed and simulated galaxy samples. Nevertheless, through galaxy cluster background selections, some galaxy populations may be systematically excluded. This may invalidate the statistical estimate of the shape multiplicative bias, namely m, depending on the shear measurement technique and on the actual properties of the data (Heymans et al. 2012; Miller et al. 2013; Hildebrandt et al. 2016).

The typical uncertainty on m found for stage III surveys ranges in the interval $\delta m \in [1\times10^{-2}, 3\times10^{-2}]$ (see, e.g., Jarvis et al. 2016; Melchior et al. 2017; Giblin et al. 2021). To assess the impact of colour selections on m, we consider the shape catalogues of Heymans et al. (2012), based on CFHTLS, and of Mandelbaum et al. (2018b), based on the HSC Subaru Strategic Program (HSC-SSP; Miyazaki et al. 2018a; Aihara et al. 2018). Throughout this section, we adopt a lens redshift of $z_{\rm l} = 0.5$. By applying the $griz$ selection calibrated in this work, we derive a shift in the mean shear multiplicative bias of $\Delta m = 7\times10^{-3}$ in CFHTLS and of $\Delta m = -2\times10^{-3}$ in HSC-SSP. In addition, the Oguri et al. (2012) and Medezinski et al. (2018b) colour selections provide $\Delta m = -3\times10^{-3}$ and $\Delta m = -1\times10^{-2}$ from CFHTLS, respectively, while from HSC-SSP we obtained $\Delta m = -5\times10^{-3}$ and $\Delta m = -7\times10^{-3}$, respectively. Thus, galaxy population differences due to colour cuts provide systematic effects that are within the typical m uncertainty in stage III surveys. By combining colour and photo-z selections, we expect Δm to become closer to zero. In *Euclid*-like surveys, shear has

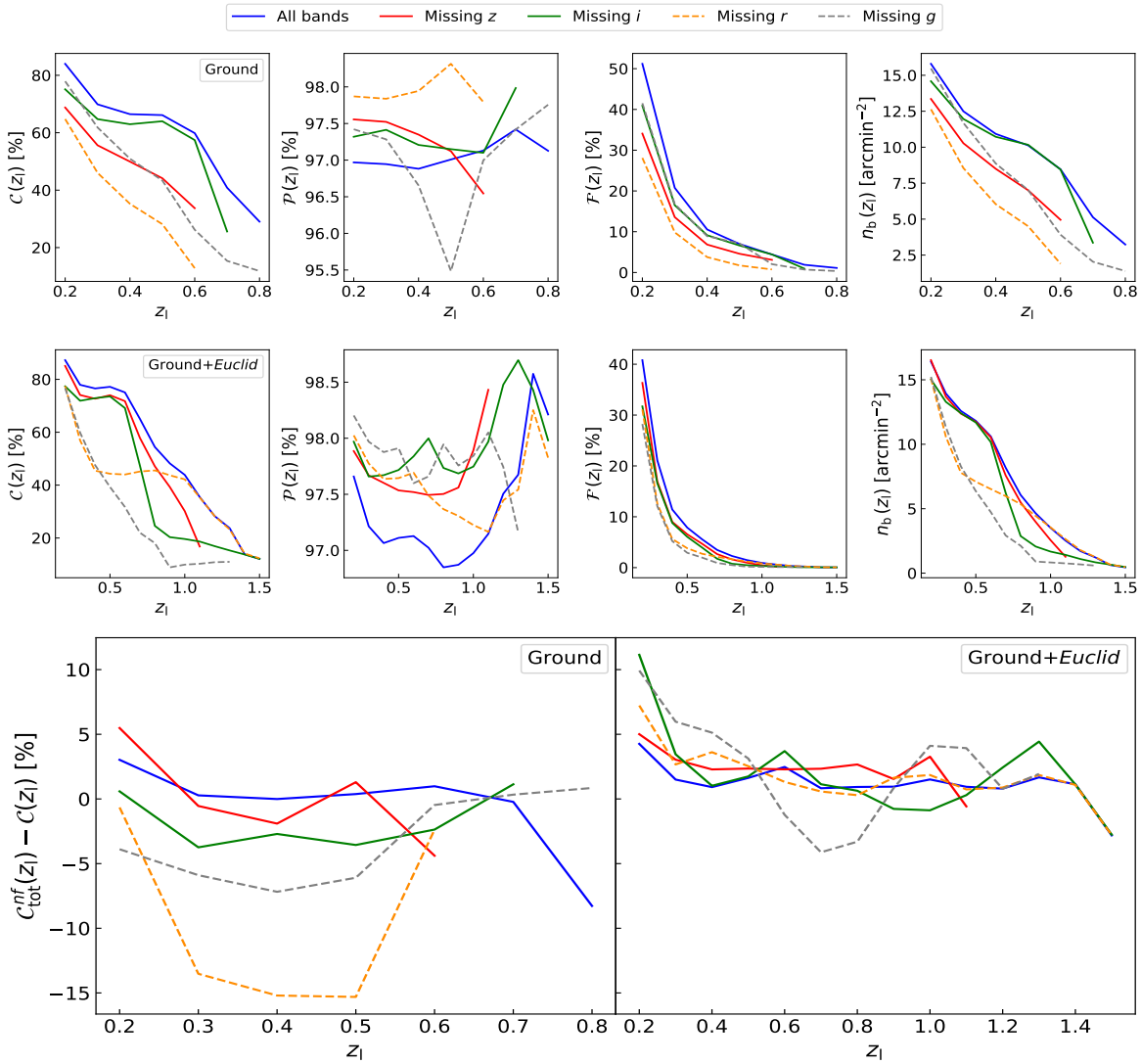

Figure 6.10: Colour selection results, obtained from the B20 catalogue, in case of missing z (solid red), i (solid green), r (dashed orange), and g (dashed grey) bands. The blue curves represent the results from the $griz$ and $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selections reported in Tables A.1 and A.2. *Top panels*, from left to right: completeness, purity, foreground failure rate, and background density are shown, in the case of ground-only photometry. *Middle panels*: colour selections from the combination of ground-based and *Euclid* photometry. The plot structure is analogous to that of top panels. *Bottom panels*: difference between $\mathcal{C}^{nf}_{\rm tot}$ and $\mathcal{C}$, for ground-only observations (left panel) and for the combination of ground-based and *Euclid* photometry (right panel).

to be calibrated within an accuracy of 2×10^{-3} (Cropper et al. 2013). As we discussed in Sect. 6.3.5, the combination of $grizY_{\rm E}J_{\rm E}H_{\rm E}$ and photo-z selections leads to 90% background completeness in the Euclid Wide Survey, on average; thus, we may expect the bias on m to be subdominant with respect to the required shear accuracy. Indeed, let us assume that 90% of galaxies, selected through the combination of $grizY_{\rm E}J_{\rm E}H_{\rm E}$ and photo-z selections, have an average m similar to that derived from stage III surveys, namely $\langle m \rangle = 0.01$. We assume that the remaining 10% of galaxies have a very biased m, namely $\langle m \rangle = 0.02$, compared to the selected population. This would imply a systematic error of $\Delta m = 10^{-3}$ in the average m of the selected population. We will delve deeper into these variations in m by examining the first data releases of the *Euclid* surveys.

6.3.7 Selection efficiency with missing bands

In this work, we derived colour selections based on *griz* and $grizY_{\rm E}J_{\rm E}H_{\rm E}$ photometry. In some cases, however, the full ground-based *griz* photometry may be not available. For example, the DES Year 3 galaxy shape catalogue was not based on g band (Gatti et al. 2021), due to issues in the point spread function estimation (Jarvis et al. 2021). Thus, we investigated the efficiency of *griz* and $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selections in the case of a missing band, based on the B20 calibration sample described in Sect. 6.1. In performing this test, we excluded the colour conditions in Tables A.1 and A.2 containing the chosen missing bands. In Fig. 6.10 we show that, in the case of ground-only observations, the absence of the r band implies the largest completeness decrease, providing $\mathcal{C}(z_{\rm l}) < 60\%$. In addition, the $z_{\rm l}$ range is substantially reduced, corresponding to $z_{\rm l} \in [0.2, 0.6]$. Also the absence of i and z bands implies a reduction of the maximum $z_{\rm l}$ for the ground-based selection, corresponding to $z_{\rm l} = 0.7$ and $z_{\rm l} = 0.6$, respectively, and a completeness decrease of up to 10 and 20 percent points, respectively. On average, a 20 percent point drop in completeness is found in absence of g band photometry. Nevertheless, in the latter case the $z_{\rm l}$ range is not reduced. We remark that the considered samples differ from case to case, as they contain only galaxies with photometry available in the required bands.

In Fig. 6.10 we show the effect of missing photometric bands on the combination of ground-based and *Euclid* observations. In this case, the lack of r band does not imply changes in $\mathcal{C}(z_{\rm l})$ for $z_{\rm l} > 1$. In the absence of i band, $\mathcal{C}(z_{\rm l})$ significantly decreases for $z_{\rm l} \gtrsim 0.7$, being below 30%, while the $z_{\rm l}$ range is not reduced. A $z_{\rm l}$ range reduction is obtained in the case of missing z or g bands, as we derived $z_{\rm l} \in [0.2, 1.1]$ and $z_{\rm l} \in [0.2, 1.3]$, respectively. On average, in the case of the combination of ground-based and *Euclid* observations, the largest completeness decrease is caused by the lack of the g band.

In this section, we defined colour selections with missing g, r, i, or z band, as subsets of the colour conditions defining the *griz* and $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selections. In order to assess the difference between the selections defined by such subsets and those that would be derived from the colour selection calibration described in Sect. 6.2, we compute $\mathcal{C}^{nf}_{\rm tot}$ for each case. In Fig. 6.10, we show the difference between $\mathcal{C}^{nf}_{\rm tot}$ and $\mathcal{C}$, the latter being derived by subsets of the colour conditions defining *griz* and $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selections. In the ground-only case, the lack of r band provides the largest $\mathcal{C}$ underestimation, as $\mathcal{C}^{nf}_{\rm tot} - \mathcal{C} \sim 15$ percent points for $z_{\rm l} \in [0.3, 0.5]$. Nevertheless, in case of other missing bands, the average $\mathcal{C}^{nf}_{\rm tot} - \mathcal{C}$ is close to 0. The same holds for the combination of ground-based and *Euclid* photometry. We conclude that *griz* and $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selections provide robust results in the case of a missing band, except for ground-only observations without the r band, for which a dedicated calibration might be needed.

6.4 Comparison with literature ground-based selections

Based on the B20 sample considered in Sect. 6.3.1, we compare our *griz* colour selection to those derived by Medezinski et al. (2010), Oguri et al. (2012), and Medezinski et al. (2018b), which are also implemented in `COMB-CL`. As detailed below, for each of these selections, we considered two versions. One includes all the colour conditions provided by the corresponding authors, while the other comprises only a subsample of such conditions, providing lower foreground contamination. `COMB-CL` includes both versions of each colour selection.

Medezinski et al. (2010) derived colour selections for three massive clusters, identified through deep Subaru imaging, by maximising their weak-lensing signal. `COMB-CL` provides the selection calibrated for the A1703 cluster at redshift $z_{\rm l} \sim 0.26$, as this is the one based

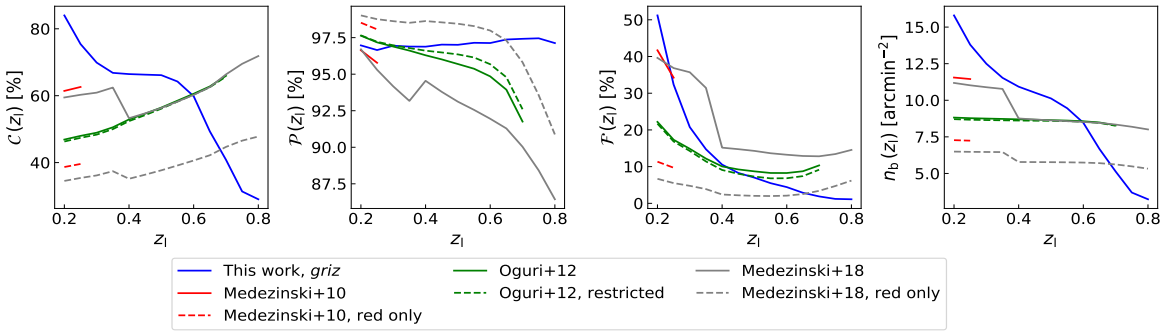

Figure 6.11: From left to right: completeness, purity, foreground failure rate, and background density, derived from the B20 catalogue. The blue solid lines represent the *griz* selection derived in this work. The red curves refer to the Medezinski et al. (2010) selection, where the solid lines are given by Eqs. (6.11) – (6.12), while the dashed lines are given by Eq. (6.11). The green curves represent the results from the Oguri et al. (2012) selection, where the solid lines are given by Eqs. (6.13) – (6.16), while the dashed lines are given by Eqs. (6.13) – (6.15). The grey curves refer to the Medezinski et al. (2018b) selection, where the solid lines are given by Eqs. (6.17) – (6.18), while the dashed lines are given by Eq. (6.17).

on *gri* photometry. This selection is expressed as follows,

$$\big[\,(g-r) < 2.17\,(r-i) - 0.37 \,\wedge \\ (g-r) < -0.6\,(r-i) + 1.85 \;\wedge\; (r-i) > 0.3\,\big] \,\vee \tag{6.11}$$

$$\big[\,(g-r) < -0.4\,(r-i) + 0.47 \;\wedge\; (r-i) < 0.3\,\big] \,\vee \\ (r-i) < -0.06\,, \tag{6.12}$$

where $\wedge$ is the logical "and" operator. Notably, Eq. (6.11) defines the red background sample, namely the galaxies redder than cluster galaxies, while Eq. (6.12) defines the blue background sample. In Fig. 6.11, we can see that this selection provides $\mathcal{C}(z_l) = 60\%$, which is 20 percent points lower than that provided by the *griz* selection calibrated in this work. In addition, from the Medezinski et al. (2010) selection we derive $\mathcal{P}(z_l) > 96\%$, which is slightly lower than that obtained from the *griz* selection discussed in Sect. 6.3.1. Foreground contamination can be attenuated by considering the red sample selection only, namely Eq. (6.11), as shown in Fig. 6.11. In this case, however, the completeness is lowered by 20 percent points. In Fig. 6.12 (upper panel), we show a comparison between the selection by Medezinski et al. (2010), namely Eqs. (6.11) and (6.12), and our *griz* selection in the $(r-i)$ - $(g-r)$ colour-colour space, by assuming $z_l = 0.26$. Within this colour-colour space, we obtain $\mathcal{C}(z_l) = 55\%$ and $\mathcal{P}(z_l) = 98\%$ from the *griz* selection, while through the Medezinski et al. (2010) selection we find $\mathcal{C}(z_l) = 62\%$ and $\mathcal{P}(z_l) = 95\%$. We remark that the full set of colour conditions defining the *griz* selection yields 80% completeness for $z_l = 0.26$, and that a calibration based on *gri* bands only would yield larger completeness values in the $(r-i)$ - $(g-r)$ space (see Sect. 6.3.7). In addition, Fig. 6.12 shows that the *griz* selection extends the selected region defined by Eq. (6.11), thus enhancing the red background sample compared to Medezinski et al. (2010). On the other hand, the *griz* selection shows a more conservative definition of the blue background sample, compared to Eq. (6.12).

Oguri et al. (2012) calibrated a selection based on *gri* photometry from the COSMOS catalogue by Ilbert et al. (2009), providing reliable results for lenses at redshift $z_l \lesssim 0.7$. This selection is expressed as

$$(g-r) < 0.3 \,\vee \tag{6.13}$$

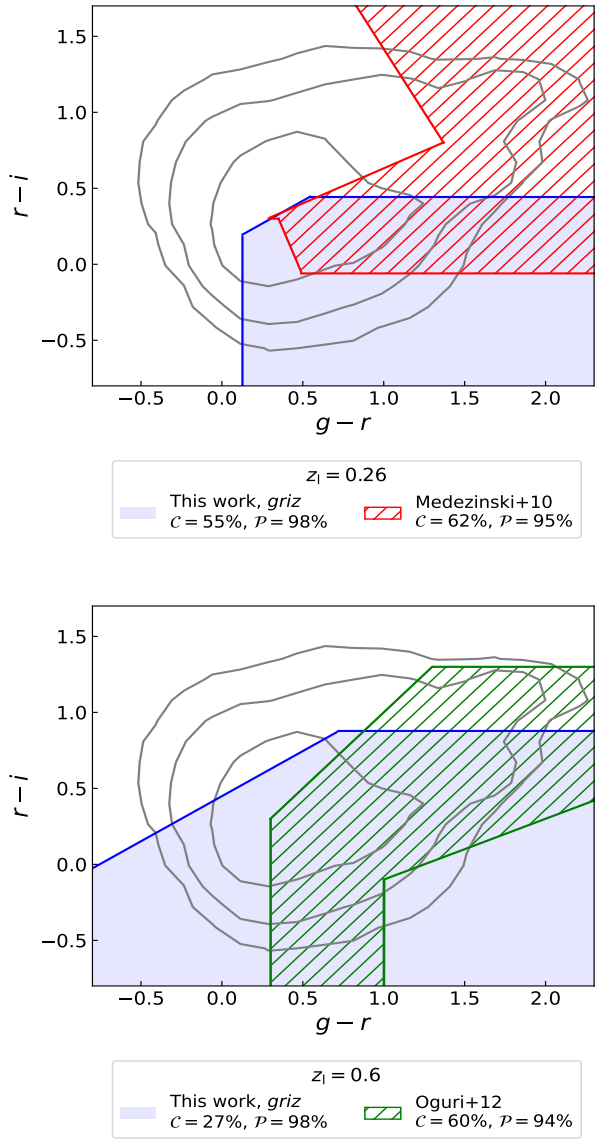

Figure 6.12: Comparison of colour selections in the $(r-i)$ - $(g-r)$ colour-colour space. The solid grey contours indicate the 68%, 95%, and 99% galaxy number density in the calibration sample, described in Sect. 6.1. The blue shaded areas represent the regions excluded by applying the *griz* selection calibrated in Sect. 6.3.1. Completeness and purity of the selections are reported in the legends. For the *griz* selection, $\mathcal{C}$ and $\mathcal{P}$ are computed by considering the colour conditions in Tab. A.1 which are defined in the $(r-i)$ - $(g-r)$ space. *Top panel*: the red hatched area shows the region excluded through the Medezinski et al. (2010) selection, and $z_l = 0.26$ is assumed. *Bottom panel*: the green hatched area shows the region excluded through the Oguri et al. (2012) selection, and $z_l = 0.6$ is assumed.

$$(r-i) > 1.3 \vee \tag{6.14}$$

$$(r-i) > (g-r) \vee \tag{6.15}$$

$$(g-r) > 1 \ \wedge \ (r-i) < 0.4\,(g-r) - 0.5. \tag{6.16}$$

The inclusion of Eq. (6.16) does not provide significant improvement in the completeness, causing a lower selection purity (Covone et al. 2014). In fact, Fig. 6.11 shows that the

selection including Eqs. (6.13) – (6.16) provides sub-percent improvements in $\mathcal{C}(z_l)$, compared to the selection including Eqs. (6.13) – (6.15) only, while $\mathcal{P}(z_l)$ and $\mathcal{F}(z_l)$ are up to 1 percent point lower and higher, respectively. With respect to the *griz* selection calibrated in this work, the Oguri et al. (2012) selection provides a purity which is up to 5 percent points lower. This explains the higher completeness values for $z_l > 0.6$. For lower z_l, the Oguri et al. (2012) selection provides a completeness up to 35 percent points lower, which is expected as the selection by Oguri et al. (2012) was calibrated for clusters at $z_l \sim 0.7$. Similarly to the comparison with the Medezinski et al. (2010) selection discussed above, in Fig. 6.12 (lower panel) we compare the *griz* and Oguri et al. (2012) selections in the $(r-i)$ - $(g-r)$ colour-colour space, assuming $z_l = 0.6$. We obtain $\mathcal{C}(z_l) = 27\%$ and $\mathcal{P}(z_l) = 98\%$ from the *griz* selection, while with the Oguri et al. (2012) selection we find $\mathcal{C}(z_l) = 60\%$ and $\mathcal{P}(z_l) = 94\%$. With respect to what we find by comparing *griz* and Medezinski et al. (2010) selections at $z_l = 0.26$, the decrease in completeness due to a purity enhancement is much larger at $z_l = 0.6$. This depends on the overlap of the foreground and background galaxy distributions in the $(r-i)$ - $(g-r)$ space. In addition, we remark that the calibration process presented in Sect. 6.2 excludes redundant colour conditions. This may partially explain the 27% completeness found in the case of the *griz* selection.

Medezinski et al. (2018b) calibrated a colour selection based on HSC observations, including *griz* bands, following an approach similar to Medezinski et al. (2010). This colour selection is expressed as

$$\begin{aligned}
&\Bigg[(g-i) < 2.276\,(r-z) - 0.152 + a(z_l) \wedge \\
&\quad (g-i) < \frac{1}{2.276}\,(r-z) + b(z_l)\left(1 + \frac{1}{2.276^2}\right) - \frac{0.152}{2.276^2} \wedge \\
&\quad (r-z) > 0.5 \wedge \\
&\quad z > 21 \Bigg] \vee
\end{aligned} \tag{6.17}$$

$$\begin{aligned}
&\Bigg\{ (r-z) < -0.0248\,z + 1.604 + c(z_l) \vee \\
&\quad \Bigg[(g-i) < \frac{1}{2.276}\,(r-z) + d(z_l)\left(1 + \frac{1}{2.276^2}\right) - \frac{0.152}{2.276^2} \wedge \\
&\quad (g-i) < 4 \Bigg] \vee \\
&\quad (r-z) < 0.5 \vee \\
&\quad z > 22 \Bigg\},
\end{aligned} \tag{6.18}$$

where

$$a(z_l) = \begin{cases} -0.7 & \text{if } z_l < 0.4 \\ -0.8 & \text{if } z_l \geq 0.4 \end{cases}, \tag{6.19}$$

$$b(z_l) = \begin{cases} 4.0 & \text{if } z_l < 0.4 \\ 1.7 & \text{if } z_l \geq 0.4 \end{cases}, \tag{6.20}$$

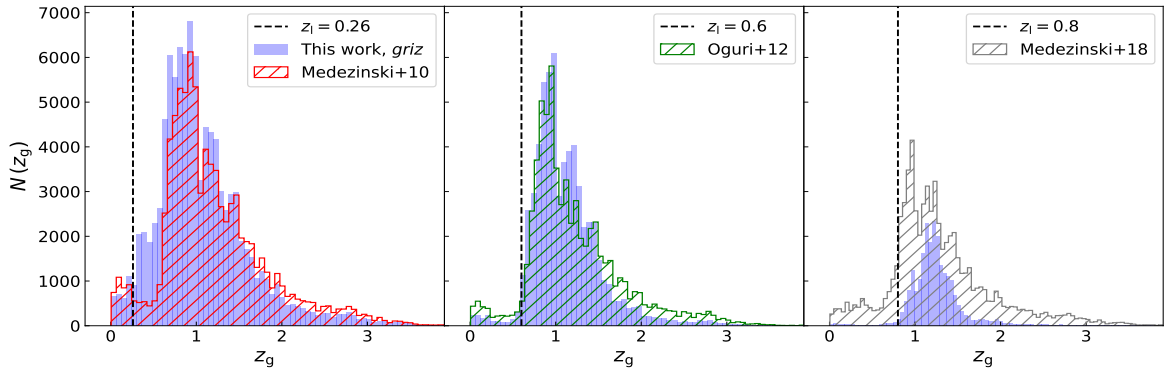

Figure 6.13: From left to right: number of selected galaxies as a function of z_g assuming $z_l = 0.26$, $z_l = 0.6$, and $z_l = 0.8$. The z_l values are represented by vertical black dashed lines. The blue histograms represent the *griz* selection calibrated in this work. The galaxy redshift distributions derived with the Medezinski et al. (2010), Oguri et al. (2012), and Medezinski et al. (2018b) selections are represented by red, green, and grey hatched histograms, respectively.

$$c(z_l) = \begin{cases} -0.8 & \text{if } z_l < 0.4 \\ -0.9 & \text{if } z_l \geq 0.4 \end{cases} , \tag{6.21}$$

$$d(z_l) = \begin{cases} 0.5 & \text{if } z_l < 0.4 \\ 0.3 & \text{if } z_l \geq 0.4 \end{cases} . \tag{6.22}$$

Similarly to Medezinski et al. (2010), Eqs. (6.17) and (6.18) define the red and blue background samples, respectively. This selection provides a much larger amount of contamination compared to that derived in this work, reaching $\mathcal{P}(z_l) < 90\%$ for $z_l > 0.7$, as shown in Fig. 6.11. However, by considering the red sample selection only, the purity improves by up to 6 percent points. In the latter case, compared to the *griz* selection detailed in Table A.1, $\mathcal{C}(z_l)$ is up to 50 percent points lower for $z_l < 0.7$. For higher z_l, the lower purity from the Medezinski et al. (2018b) selection allows for higher completeness values.

In Fig. 6.13 we show the redshift distributions of the galaxies selected through the *griz*, Medezinski et al. (2010), Oguri et al. (2012), and Medezinski et al. (2018b) selections, assuming different z_l values. At $z_l = 0.26$, the *griz* selection shows a larger number of galaxies, of the order of 10^3, which are correctly identified as background objects with $z_g < 0.6$, compared to Medezinski et al. (2010). This results in the larger completeness of the *griz* selection shown in Fig. 6.11. At $z_l = 0.6$, where the *griz* and Oguri et al. (2012) selections have the same completeness (see Fig. 6.11), the *griz* selection is less complete at $z_g > 1.5$ and more complete at lower redshifts, compared to Oguri et al. (2012). At $z_l = 0.8$, where the *griz* selection is remarkably purer than that by Medezinski et al. (2018b), a notable incompleteness of the *griz* selection is evident at any z_g, compared to Medezinski et al. (2018b). In fact, for the case of the *griz* selection, Fig. 6.13 shows that the number of rejected galaxies at high redshift increases with the number of excluded foreground galaxies. This reflects the overlapping of foreground and background galaxy distributions in the considered colour-colour spaces.

6.5 Summary

We developed a method to derive optimal galaxy colour selections for cluster weak-lensing analyses, given any set of photometric bands. To this aim, we considered all the available colour-colour combinations. Based on the galaxy catalogue by B20, we calibrated selections

based on ground-based *griz* and *Euclid* $Y_{\rm E}J_{\rm E}H_{\rm E}$ bands, with purity higher than 97%. Specifically, we showed that the *griz* selection provides a completeness between 30% and 84%, in the lens limiting redshift range $z_{\rm l} \in [0.2, 0.8]$. The inclusion of *Euclid* $Y_{\rm E}J_{\rm E}H_{\rm E}$ bands, leading to a *griz*$Y_{\rm E}J_{\rm E}H_{\rm E}$ selection, improves the completeness by up to 25 percent points in this $z_{\rm l}$ range, allowing for a galaxy selection up to $z_{\rm l} = 1.5$. In addition, for the first time in the literature, we expressed such selections as a continuous function of $z_{\rm l}$. In the following, we summarise the main results obtained from the tests presented in Sects. 6.3 and 6.4.

- The calibrated colour selections, described in Sect. 6.3.1, are stable with respect to changes in the sample limiting magnitudes and redshift.

- By applying the *griz* selection to the VMLS catalogue by Moutard et al. (2016) and to the COSMOS20 catalogue by Weaver et al. (2022), we derived completeness and purity estimates that are consistent with those obtained from the calibration sample by B20. Consequently, the calibrated selections provide stable results by assuming alternative photometric aperture definitions, obtained from different ground-based telescopes.

- The application *griz* and *griz*$Y_{\rm E}J_{\rm E}H_{\rm E}$ selections to the simulated *Euclid* Flagship galaxy catalogue v2.1.10 (Euclid Collaboration in prep.) provided a purity of around 99%, on average, which is higher than that obtained from the B20 catalogue. The completeness from the Flagship and B20 samples is compatible in the *griz* selection case, while the *griz*$Y_{\rm E}J_{\rm E}H_{\rm E}$ selection provides up to 50 percent points higher completeness from Flagship. We verified that this discrepancy does not depend on magnitude limits. In addition, we found no significant differences in the star forming galaxy fraction from the Flagship and B20 samples. A calibration of the *griz*$Y_{\rm E}J_{\rm E}H_{\rm E}$ selection based on the Euclid Deep Survey will allow for a more thorough investigation of these completeness differences.

- Based on the Flagship sample, we combined the calibrated colour selections with photo-z selections based on the $p(z_{\rm g})$ shape. We showed that in this case the completeness is up to 95%.

- We found no significant systematic effects on the multiplicative shear bias due to colour selections for stage III surveys. The first *Euclid* data releases will provide further insights into the influence of background selections on this bias.

- The calibrated colour selections provide robust results in the case of a missing band from ground-based observations, except for those without the r band, for which a dedicated calibration might be needed.

- Compared to the ground-based colour selections provided by Medezinski et al. (2010), Oguri et al. (2012), and Medezinski et al. (2018b), the *griz* selection derived in this work is purer at high redshift and more complete at low redshift.

One potential enhancement to the calibration presented in this work could entail the inclusion of a magnitude dependence in the colour cuts. This might mitigate the impact of large photometric scatter at faint magnitudes (see, e.g., Schrabback et al. 2021). In addition, enhancing the set of photometric bands in the calibration sample, for example by including the LSST y band, could remarkably improve the effectiveness of the colour selections. The performance of colour selections could further improve through dedicated calibration samples. Ongoing spectroscopic programs are specifically designed to calibrate the relationship between galaxy colours and redshifts to match the depth of the Euclid survey (Euclid Collaboration: Saglia

et al. 2022).

Furthermore, the colour selections presented in this work could improve the shear calibration in cluster weak-lensing analyses. The lensing signal from galaxy clusters differs from that of large scale structure in ways that affect both shear and photometric measurements. The dense cluster environment causes increased blending among light sources, due to both galaxy blends (Simet & Mandelbaum 2015; Everett et al. 2022) and the presence of diffuse intra-cluster light (Gruen et al. 2019; Kluge et al. 2020; Sampaio-Santos et al. 2021). In addition, cluster lines of sight exhibit characteristically stronger shear, especially at small scales (McClintock et al. 2019; Ingoglia et al. 2022). These effects can lead to distinct biases in shear measurements compared to those obtained from calibrations primarily designed for cosmic shear analyses. Through the combination of colour and photo-z selections, cluster shear calibration and mass bias can be assessed based on dedicated, multi-band cluster image simulations (see, e.g., Hernández-Martín et al. 2020).

Chapter 7

AMICO clusters in KiDS-1000: weak-lensing mass calibration

In this Chapter, we introduce the preliminary work by Lesci et al. (in prep. b) on the mass calibration of the AMICO galaxy cluster sample provided by Maturi et al. (in prep.), based on the fourth data release of KiDS (KiDS-1000, Kuijken et al. 2019). Compared to the KiDS-DR3 sample detailed in Chapt. 5, KiDS-1000 covers an area approximately 550 deg^2 larger, namely of 1000 deg^2, and additional five NIR bands from the VISTA Kilo degree INfrared Galaxy survey (VIKING, Edge et al. 2013b; Sutherland et al. 2015) are included. Hence, the number of detected clusters is remarkably larger and the redshift range used for the cluster mass calibration can be extended compared to KiDS-DR3, as we shall see in the following. Based on the methods detailed in Sect. 3.2, we measure and model the stacked weak-lensing signal of AMICO clusters in bins of redshift and mass proxy, relying on the background colour selection developed by Euclid Collaboration: Lesci et al. (2023) and detailed in Chapt. 6. We derive constraints on the mass-richness and concentration-mass scaling relations, marginalising such results over the main uncertainties affecting weak-lensing measurements and cluster detection in optical bands.

This Chapter is organised as follows. In Sect. 7.1 we present the galaxy cluster and shear samples. The measurement and stacking of the cluster weak-lensing profiles are detailed in Sect. 7.2, while Sect. 7.3 introduces the theoretical models adopted in the analysis. In Sect. 7.4 we present the assessment of the systematic uncertainties affecting the measurements, and in Sect. 7.5 we introduce the likelihood and priors assumed in the analysis. Lastly, in Sect. 7.6 we present our results and in Sect. 7.7 we draw our conclusions.

Throughout this Chapter, we assume a concordance flat ΛCDM cosmological model with $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_\mathrm{m} = 0.3$. The statistical analyses were performed using the CBL numerical libraries (Marulli et al. 2016).

7.1 The KiDS-1000 data set

The work presented in this Chapter is based on KiDS-1000 (Kuijken et al. 2019), covering 1006 tiles of 1 deg^2 each. The 2 arcsec aperture photometry in *ugri* bands is provided, with 5σ limiting magnitudes of 24.23, 25.12, 25.02 and 23.68 for the aforementioned four bands, respectively. In addition, aperture-matched $ZYJHK_\mathrm{s}$ near-infrared photometry from the fully overlapping VIKING survey (Edge et al. 2013b; Sutherland et al. 2015) is included. This is a remarkable enhancement compared to KiDS-DR3, where only *ugri* photometry was available (de Jong et al. 2017). Based on these photometric bands, the KiDS-1000 galaxy sample includes photo-z estimates for more than 100 million objects, with a typical scatter

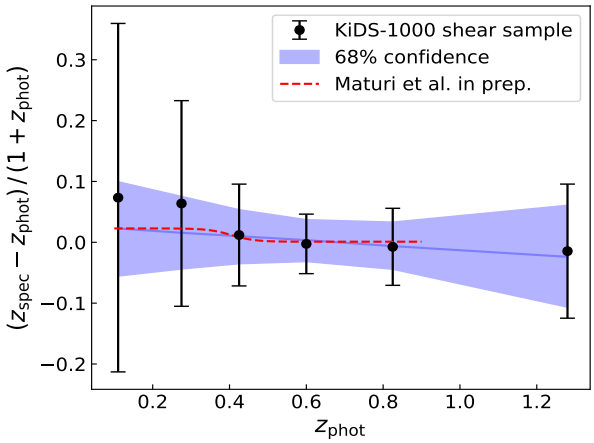

Figure 7.1: Redshift bias of KiDS-1000 galaxies and galaxy clusters. The black dots represent the measure of the galaxy redshift bias, based on the spectroscopic redshift catalogue developed by van den Busch et al. (2022) and Wright et al. (in press), with error bars corresponding to the standard deviations. The blue band represents the 68% confidence level derived from the multivariate posterior of the free parameters considered in the fit (see Sect. 7.1.2). The dashed red line shows the bias of AMICO cluster redshifts derived by Maturi et al. (in prep.).

of $\sigma_z/(1+z) = 0.072$ (Kuijken et al. 2019).

7.1.1 Galaxy cluster sample

The cluster catalogue this work is based on, named AMICO KiDS-1000 (Maturi et al. in prep.), was built up running the AMICO algorithm (Bellagamba et al. 2018) in KiDS-1000 (Kuijken et al. 2019). To perform the cluster detection, all the galaxies located in the regions affected by image artefacts have been rejected. This yields a final effective area of 840 deg^2, containing cluster detections up to $z = 1$, for a total of 52 443 objects. We consider the clusters having S/N > 3.5, resulting in a sample of 24 396 objects. We correct the cluster redshifts for the bias derived by Maturi et al. (in prep.), shown in Fig. 7.1. As for the AMICO KiDS-DR3 cluster sample (Maturi et al. 2019), the origin of this bias lies in the systematic uncertainty associated with galaxy photo-zs. By including the correction for this bias, the redshift statistical uncertainty of AMICO clusters is $\sigma_z/(1+z) = 0.0175$. In Fig. 7.2 we show the distribution of the corrected cluster redshifts and of the intrinsic richnesses. As we can see in Fig. 7.2, the number of clusters is three times larger compared to the KiDS-DR3 sample derived by Maturi et al. (2019), and also the redshift range is enhanced. Due to remarkable differences in the quality of galaxy photo-zs, thanks to the inclusion of VIKING photometry in KiDS-1000, the λ^* distribution of AMICO KiDS-DR3 clusters is not shown in Fig. 7.2 as it is not comparable to that derived by Maturi et al. (in prep.).

7.1.2 Shear sample

To estimate the weak-lensing signal produced by the AMICO KiDS-1000 galaxy clusters, we base our analysis on the KiDS-1000 gold shear catalogue (Wright et al. 2020; Hildebrandt et al. 2021; Giblin et al. 2021). This sample comprises about 21 million galaxies, covering an effective area of 777.4 deg^2, with a weighted number density of $n_{\rm eff}^{\rm gold} = 6.17$ arcmin^{-2}. Weak-lensing shear measurements from the deep KiDS r-band observations are provided, being the ones with better seeing properties and yielding the highest source density. The shear estimator used in this analysis is *lensfit* (Miller et al. 2007; Fenech Conti et al. 2017),

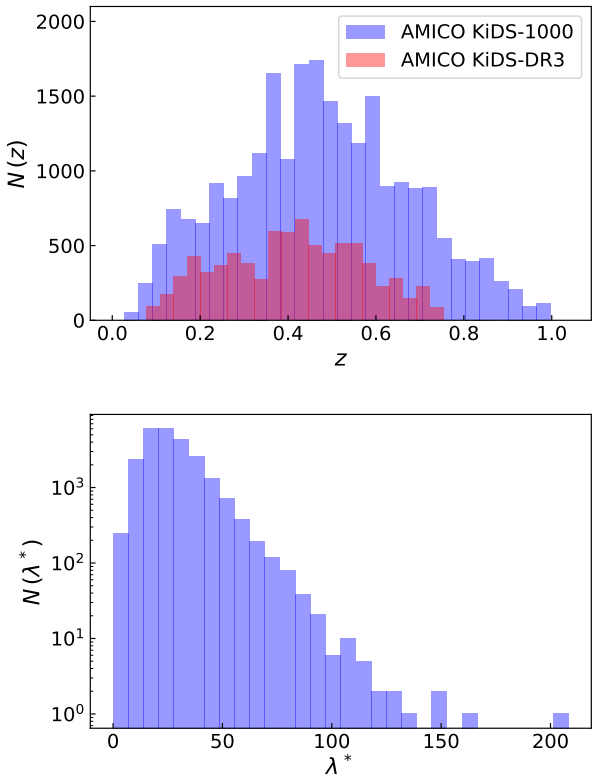

Figure 7.2: Redshift (top panel) and intrinsic richness (bottom panel) distributions of the AMICO KiDS-1000 (blue histograms) and AMICO KiDS-DR3 (red histogram) galaxy clusters with S/N > 3.5. The redshifts are corrected following Maturi et al. (in prep.) and Maturi et al. (2019).

a likelihood-based model-fitting method which has been successfully used in the analyses of other datasets, such as the Canada France Hawaii Telescope Lensing Survey (CFHTLenS, Miller et al. 2013) and the Red Cluster Survey (Hildebrandt et al. 2016). The photo-z calibration, based on deep spectroscopic reference catalogues that are reweighted with the help of a self-organising map (SOM; Wright et al. 2020), has been performed by Hildebrandt et al. (2021) in the redshift range $z \in [0.1, 1.2]$.

As in Kuijken et al. (2019), Maturi et al. (in prep.) used the Bayesian Photometric Redshift (BPZ, Benítez 2000) code to derive the galaxy photo-zs for cluster detection. However, the prior redshift probability used in BPZ by Kuijken et al. (2019) appears to generate a remarkable redshift bias for bright, low-redshift galaxies, despite it reduced uncertainties and catastrophic failures for faint galaxies at higher redshifts. For this reason, Maturi et al. (in prep.) chose the same prior adopted in the KiDS-DR3 analysis (de Jong et al. 2017). Consistently with this choice, in the analysis presented in this Chapter, we adopt the galaxy photo-zs derived by Maturi et al. (in prep.). We derive the galaxy redshift bias through the compilation of public spectroscopic data, developed by van den Busch et al. (2022) and extended by Wright et al. (in press), which includes 642 978 objects. To match the galaxies in the spectroscopic sample with those in the KiDS-1000 shear catalogue, we employ an aperture radius of 0.5 arcsec, resulting in a total of 71 494 matches. Then, we measure and fit the quantity $(z_{\rm spec} - z_{\rm phot}) / (1 + z_{\rm phot})$ as a function of $z_{\rm phot}$, where $z_{\rm spec}$ and $z_{\rm phot}$ are the spectroscopic and photometric redshifts, respectively. Specifically, we perform a Bayesian analysis via an MCMC, assuming a Gaussian likelihood and the following linear model

$$f_{\rm mod}(z_{\rm phot}) = a\, z_{\rm phot} + b\,. \tag{7.1}$$

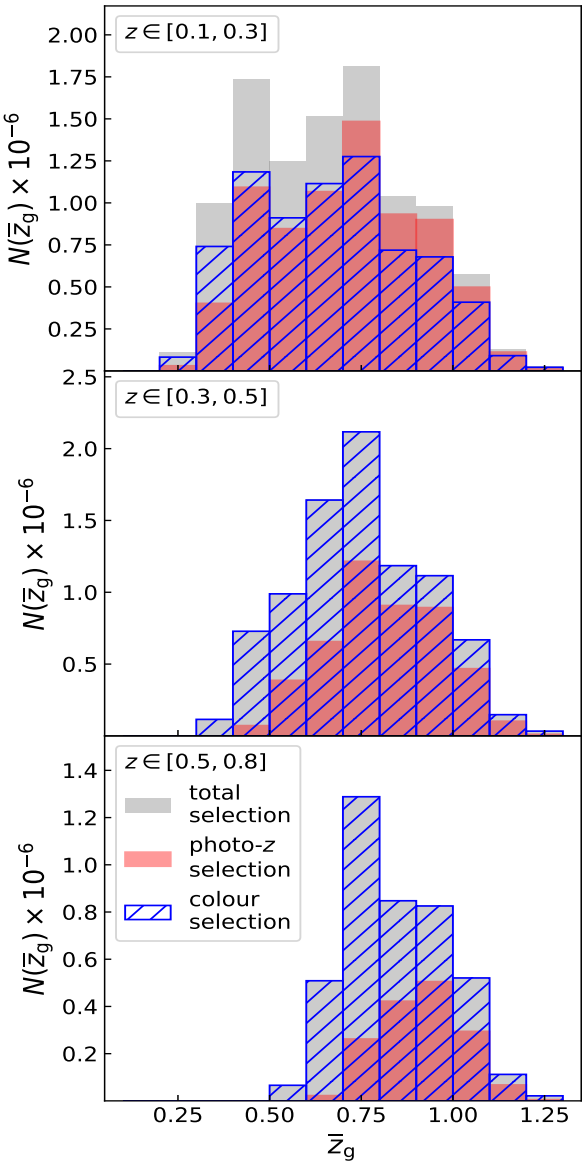

Figure 7.3: Number of galaxies satisfying the photo-z selection (red histograms) and the colour selection (blue hatched histograms), in different cluster redshift bins, namely $z \in [0.1, 0.3]$ (top panel), $z \in [0.3, 0.5]$ (middle panel), and $z \in [0.5, 0.8]$ (bottom panel). The grey histograms show the galaxies selected through the total selection criterion, defined in Eq. (7.4).

We obtain $a = -0.05 \pm 0.13$ and $b = 0.04 \pm 0.09$. In Fig. 7.1, we show that this galaxy redshift bias agrees with that derived by Maturi et al. (in prep.) for AMICO clusters. The error on the mean of the measured bias, of about 10^{-3}, is negligible compared to the typical galaxy redshift uncertainty.

7.2 Stacked weak-lensing profiles

7.2.1 Selection of background sources

If galaxies belonging to the clusters or in the foreground are mistakenly considered as background, the measured lensing signal can be significantly diluted (see, e.g., Broadhurst et al.

2005; Medezinski et al. 2007; Dietrich et al. 2019). To prevent such contamination, we employ stringent selections that consider both the galaxy photo-z probability distribution function, denoted as $p(z_{\rm g})$, where $z_{\rm g}$ represents the galaxy redshift, and galaxy colours. This approach is in line with previous works such as Sereno et al. (2017) and Bellagamba et al. (2019).

Specifically, to exclude galaxies with a significant probability of being at a redshift equal to or lower than that of the cluster, we adopt the following photo-z selection

$$z_{\rm g,min} > \overline{z} + 0.05\,, \tag{7.2}$$

where $z_{\rm g,min}$ is the minimum of the interval containing 95% of the probability around the first mode of $p(z_{\rm g})$, namely $\overline{z}_{\rm g}$, while $\overline{z}$ is the mean redshift of the cluster. In Eq. (7.2), the 0.05 buffer is larger than the uncertainty associated with cluster redshifts, that is $\sigma_z/(1+z) = 0.0175$.

To improve the cluster weak-lensing signal, we recover part of the population excluded through the photo-z selection by considering a colour selection (similarly to Oguri et al. 2012; Medezinski et al. 2018a; Dietrich et al. 2019; Bellagamba et al. 2019; Schrabback et al. 2021). In particular, we rely on the colour selection calibrated by Euclid Collaboration: Lesci et al. (2023) (detailed in Chapt. 6), based on *griz* photometry. This selection provides a background completeness ranging from 30% to 84% and a purity larger than 97% in the lens redshift range $z \in [0.2, 0.8]$. In addition, candidate background galaxies can pass the colour selection if

$$\overline{z}_{\rm g} > \overline{z} + 0.05\,, \tag{7.3}$$

where the 0.05 buffer is the same as in Eq. (7.2). The complete selection criterion is defined as follows

$$\text{(photo-}z\text{ selection)} \;\vee\; \text{(colour selection)}\;, \tag{7.4}$$

where $\vee$ is the "or" logical operator. To assess the performance of the two background selections in Eq. (7.4), we derive the number of galaxies that meet such selection criteria. We consider the same cluster redshift bins adopted for the weak-lensing measurements presented in the following analysis, namely $z \in [0.1, 0.3]$, $z \in [0.3, 0.5]$, and $z \in [0.5, 0.8]$. Fig. 7.3 shows that the colour selection provides the largest number of background sources at low $\overline{z}_{\rm g}$. This is expected, as the photo-z selection in Eq. (7.2) excludes a greater number of galaxies close to the clusters. In the second and third cluster redshift bins, the colour selection outperforms the photo-z selection in terms of the number of selected background galaxies at all $\overline{z}_{\rm g}$. For both $z \in [0.3, 0.5]$ and $z \in [0.5, 0.8]$, this results in about three to four million more selected galaxies compared to the photo-z selection. In addition, for $z \in [0.3, 0.5]$ and $z \in [0.5, 0.8]$, the photo-z selection does not significantly contribute to the total number of background galaxies defined via Eq. (7.4). This is in contrast to what found by Euclid Collaboration: Lesci et al. (2023) from the combination of colour and photo-z selections in *Euclid* Flagship mock data (Euclid Collaboration in prep.). Despite the photo-z selection and the number of photometric bands used for estimating photo-zs are similar to those considered in this work, Euclid Collaboration: Lesci et al. (2023) found that the photo-z selection provides larger completeness than the colour selection, with similar purity values (see Sect. 6.3.4). We expect that upcoming *Euclid* real data will clarify the reason of such differences.

7.2.2 Measure of the stacked profiles

As we discussed in Sect. 3.2, the excess surface density, $\Delta\Sigma_+$, is linked to the tangential component of the shear via Eq. (3.4). Since the weak-lensing approximation is accurate for the physical scales examined in this study, we can determine $\Delta\Sigma_+$ from observed galaxy

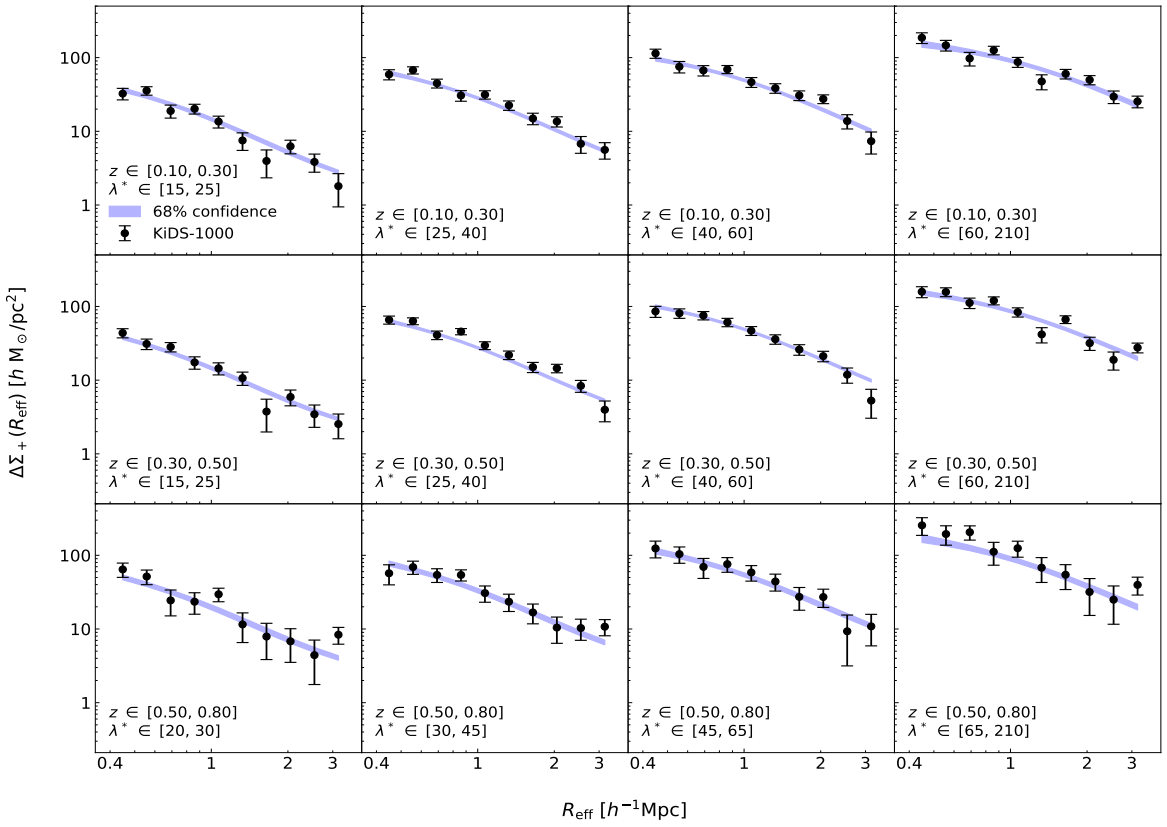

Figure 7.4: Stacked $\Delta\Sigma_+(R_{\rm eff})$ profiles of the AMICO KiDS-1000 galaxy clusters in bins of z and λ^*. The black dots show the measure, with error bars derived via bootstrap resampling. The blue bands represent the 68% confidence level derived from the multivariate posterior of all the free parameters considered in the analysis.

ellipticities (see Sect. 2.6) through the estimator defined in Eq. (3.7). For the multiplicative shear bias, m, in Eq. (3.7), we rely on the results by Giblin et al. (2021). Specifically, Giblin et al. (2021) derived m estimates in five redshift bins, which are consistent with zero and have standard deviation ranging from 10^{-2} to 2×10^{-2}. Since the galaxy photo-z calibration considered in this work differs from that adopted by Giblin et al. (2021), we extract for each galaxy a value of m from the uniform distribution $[-0.05, 0.05]$, covering the 2σ intervals of the m distributions derived by Giblin et al. (2021). We verified that extracting m values from a Gaussian distribution with mean equal to 0 and a standard deviation of 0.04 does not significantly change the results. In Eq. (3.7), for the computation of the critical density of the ith galaxy, $\Sigma_{{\rm cr},i}$, we adopt the most probable redshift of the source, namely $\overline{z}_{\rm g}$. We investigate the impact of this choice in Sect. 7.4. We do not account for the uncertainty in the lens redshift, whose impact on $\Sigma_{\rm cr}$ is negligible compared to the source one.

For the most part of the clusters in the sample, the weak-lensing signal is too low to measure the density profiles, and thus to precisely constrain masses. For this reason, we derive stacked $\Delta\Sigma_+$ estimates for ensembles of clusters, selected according to their intrinsic richness and redshift, based on Eq. (3.9). For the Kth cluster stack, we derive the effective λ^*, namely $\lambda^*_{K,{\rm eff}}$, as follows (Umetsu et al. 2014)

$$\lambda^*_{K,{\rm eff}} = \frac{\sum_{k \in K} W_k \lambda^*_k}{\sum_{k \in K} W_k} \,, \tag{7.5}$$

where k runs over the clusters in the bins of z and λ^*, and W_k is derived via Eq. (3.10) without considering any subdivisions in radial bins. For simplicity, we refer to $\lambda^*_{K,\mathrm{eff}}$ as λ^*_{eff} hereafter. We analogously derive the effective redshift, z_{eff}, and radius, R_{eff}.

We consider λ^* thresholds for which the purity of each stack is larger than 99% (Maturi et al. in prep.), namely $\lambda^* > 15$ in the cluster redshift bins $z \in [0.1, 0.3]$ and $z \in [0.3, 0.5]$, while we assume $\lambda^* > 20$ for $z \in [0.5, 0.8]$. In each of these redshift bins, we divide the clusters into 4 bins of λ^*. In addition, we consider 10 radial bins in the cluster-centric radial range $R \in [0.4, 3.5]\ h^{-1}$Mpc. By excluding the 400 h^{-1}kpc close to the centre, the contamination due to cluster members is significantly reduced (see, e.g., Medezinski et al. 2018b; Bellagamba et al. 2019). Moreover, the analysis of the shear signal in the immediate vicinity of the cluster centre is responsive to the influence of the BCG on the matter distribution and to deviations from the weak-lensing approximation. The exclusion of the scales larger than 3.5 h^{-1}Mpc reduces the cosmological dependence of the results, primarily stemming from the matter correlated with cluster haloes. Also the impact of possible anisotropic boosts, affecting correlation functions on large scales due to projection effects (Sunayama et al. 2020; Park et al. 2023; Sunayama 2023), is mitigated.

In Fig. 7.4 we show the stacked $\Delta\Sigma_+(R_{\mathrm{eff}})$ measurements in bins of redshift and mass proxy. The covariance matrix for $\Delta\Sigma_+(R_{\mathrm{eff}})$ is estimated with a bootstrap procedure with replacement, by performing 10 000 resamplings of the galaxy cluster sample. We do not consider the covariance between radial bins in different redshift and mass proxy bins, as its impact on the final results is negligible (Ingoglia et al. 2022). As we discussed in Sect. 2.6.4, if the measured signal comes from gravitational lensing only, the stacked cross-component, namely $\Delta\Sigma_\times = \Sigma_{\mathrm{cr}}\gamma_\times$, should be consistent with zero. Fig. 7.5 shows that $\Delta\Sigma_\times$ does not exhibit substantial deviation from zero across the entire radial range examined in the analysis. The associated reduced χ^2, derived by assuming a null signal and 120 degrees of freedom, corresponds to $\chi^2_{\mathrm{red}} = 0.7$.

7.3 Model

In this section, we introduce a model that enables the constraints of the scaling relation between cluster masses and mass proxies, marginalising the results over the halo miscentring parameters and the concentration-mass relation. We do not correct our model to account for the fact that the reduced shear is the weak-lensing observable. Introducing such correction would result in negligible changes in the final results, especially for the physical scales within the scope of our analysis (see, e.g., McClintock et al. 2019).

7.3.1 Lens model

In order to describe the 3D profiles of galaxy clusters, we adopt a BMO model (Eq. 3.13). As we shall detail in Sect. 7.5, we set the truncation factor of the BMO profile, namely F_{t}, as a free parameter in the analysis. In addition, we consider cluster masses within r_{200}, that is the radius enclosing a mass such that the corresponding mean density is 200 times the critical density of the Universe at that redshift. The lens model is expressed as follows

$$\Delta\Sigma_+(R) = (1 - f_{\mathrm{off}})\,\Delta\Sigma_{+,\mathrm{cen}}(R) + f_{\mathrm{off}}\,\Delta\Sigma_{+,\mathrm{off}}(R)\,, \tag{7.6}$$

where f_{off} is the fraction of haloes that belong to the miscentred population, while $\Delta\Sigma_{+,\mathrm{cen}}$ and $\Delta\Sigma_{+,\mathrm{off}}$ are the excess surface density profiles of the centred and miscentred cluster populations, respectively (see Sect. 3.2). Both $\Delta\Sigma_{+,\mathrm{cen}}$ and $\Delta\Sigma_{+,\mathrm{off}}$ include the contribution from the 2-halo term (Eq. 3.16), which is added to the centred surface density, Σ_{cen}, defined

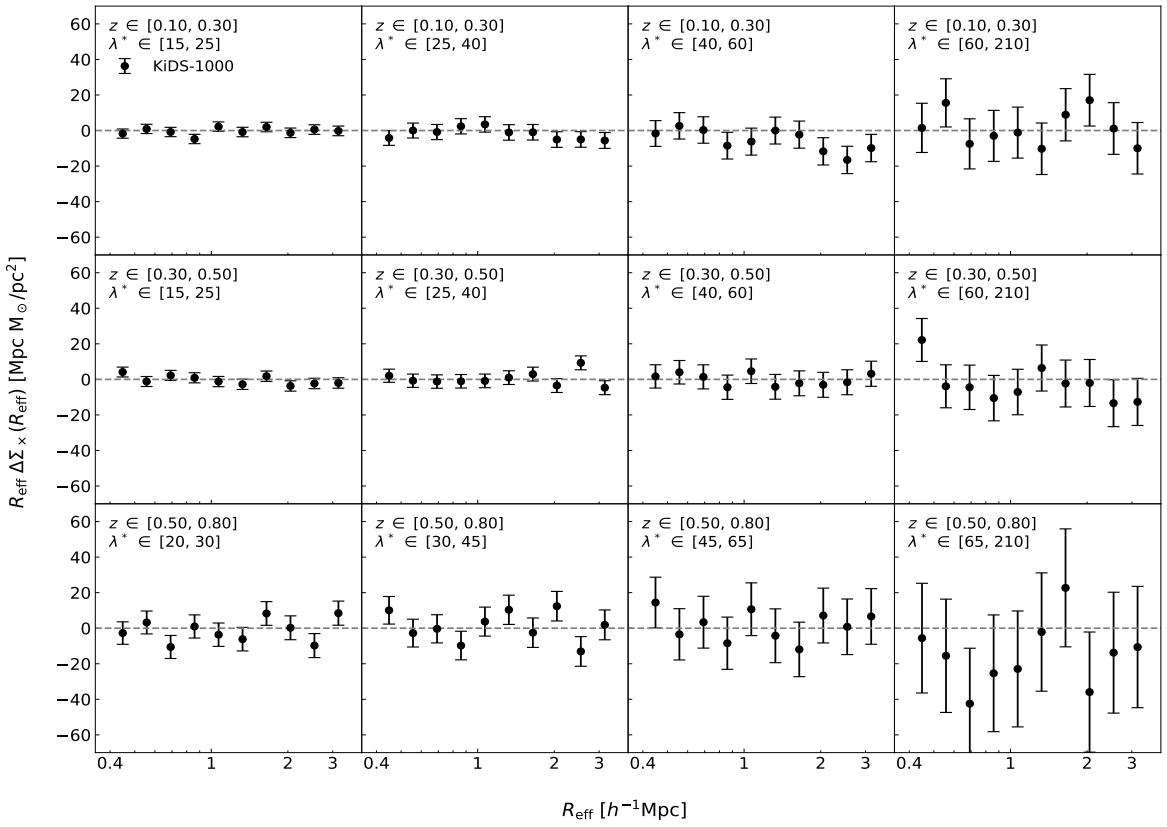

Figure 7.5: Stacked $R_{\rm eff}\Delta\Sigma_{\times}(R_{\rm eff})$ profiles of the AMICO KiDS-1000 galaxy clusters in bins of z and λ^*. The error bars are derived through bootstrap resampling.

in Eq. (3.14). To derive $\Delta\Sigma_{+,\rm off}$, we assume that the probability of a lens being at a distance $R_{\rm s}$ from the chosen centre, namely $P(R_{\rm s})$, follows a Rayleigh distribution (Eq. 3.18). In Eq. (3.18), the rms of $P(R_{\rm s})$, namely $\sigma_{\rm off}$, is expressed in units of h^{-1}Mpc and, along with $f_{\rm off}$, it is a free parameter in the analysis.

7.3.2 Mass-richness and concentration-mass relations

We simultaneously model the measured $\Delta\Sigma_+$ in all the redshift and mass proxy bins. The expected value of the excess surface density is defined as

$$\langle\Delta\Sigma_+(R_{\rm eff}|\lambda^*_{\rm eff}, z_{\rm eff})\rangle = \int_0^\infty {\rm d}M\ \Delta\Sigma_+(R_{\rm eff}|M, z_{\rm eff})\, P(M|\lambda^*_{\rm eff}, z_{\rm eff}), \tag{7.7}$$

where $\Delta\Sigma_+(R_{\rm eff}|M, z_{\rm eff})$ is given by Eq. (7.6), while $P(M|\lambda^*_{\rm eff}, z_{\rm eff})$, defined in Eq. (5.7), is a log-normal distribution with mean corresponding to the mass-richness scaling relation defined in Eq. (5.8). Specifically, in Eq. (5.8) we assume the following pivot values: $\lambda^*_{\rm piv} = 50$ and $z_{\rm piv} = 0.4$. The rms of $P(M|\lambda^*_{\rm eff}, z_{\rm eff})$ corresponds to the intrinsic scatter, namely $\sigma_{\rm intr}$, defined as follows

$$\sigma_{\rm intr} = \frac{\sigma_{\rm intr,0} + \sigma_{{\rm intr},\lambda^*}\log(\lambda^*_{\rm eff}/\lambda^*_{\rm piv})}{\sqrt{N_{\rm cl}}}, \tag{7.8}$$

where $\sigma_{\rm intr,0}$ and $\sigma_{{\rm intr},\lambda^*}$ are free to vary in the analysis, while $N_{\rm cl}$ is the number of clusters in the given redshift and mass proxy bin. By dividing $\sigma_{\rm intr}$ by $\sqrt{N_{\rm cl}}$, we neglect the cluster

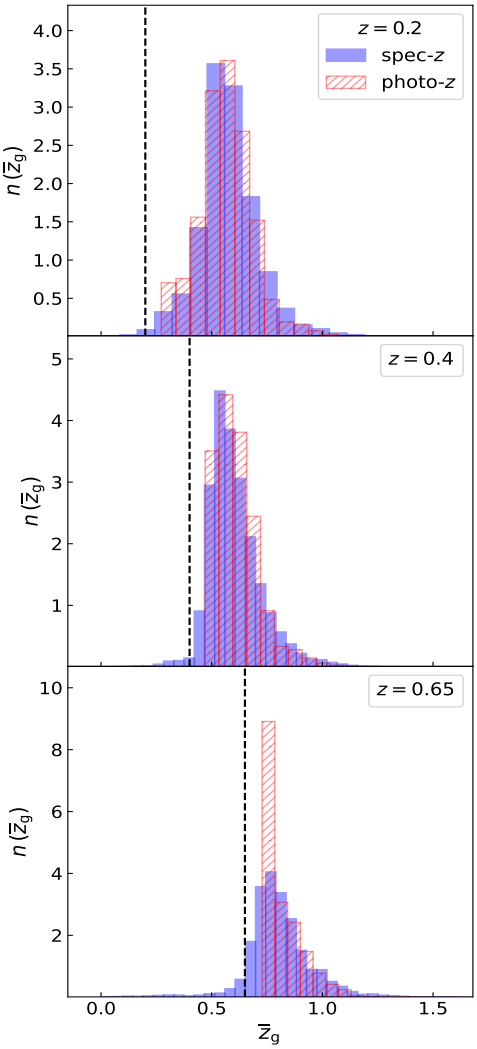

Figure 7.6: Density of galaxies satisfying the background selection as a function of $\overline{z}_{\rm g}$. The spectroscopic sample is represented by blue histograms, while the red hatched histograms show the distributions of the KiDS-1000 photometric redshifts. The vertical black dashed lines represent the lens redshifts, namely $z = 0.2$ (top panel), $z = 0.4$ (middle panel), and $z = 0.65$ (lower panel).

clustering contribution. In addition, the concentration is defined as follows

$$\log c_{200} = c_0 + c_M \log\left(\frac{M}{10^{14} M_\odot/h}\right) + c_z \log\left(\frac{1+z}{1+z_{\rm piv}}\right) , \tag{7.9}$$

where $z_{\rm piv} = 0.4$, while c_0, c_M, and c_z are free parameters in the analysis.

7.4 Systematic uncertainties

As discussed in Sect. 7.2.2, we correct the weak-lensing measurements for the multiplicative shear bias, also accounting for its uncertainty. In Sect. 7.3 we detailed a number of terms accounting for systematic uncertainties, such as the halo miscentring, enclosed in the model. We expect that the biases due to the contribution of the baryonic matter (Henson et al. 2017; Shirasaki et al. 2018; Lee et al. 2018; Beltz-Mohrmann & Berlind 2021) are absorbed by $\sigma_{\rm intr}$ and c_{200}, which are set as free parameters.

In this section we present the assessment of other relevant systematic uncertainties affecting the measured $\Delta\Sigma_+(R)$, caused by photo-z probability distributions, background selection, halo orientation, and projection effects. As we shall see in Sect. 7.5, we include such uncertainties in the covariance matrix of the likelihood function.

7.4.1 Photo-z uncertainty

In Sect. 7.2.2, we use the galaxy most probable redshift, $\overline{z}_{\rm g}$, to estimate $\Sigma_{\rm cr}$, neglecting the full shape of $p(z_{\rm g})$. In fact, inaccurate low probability tails of $p(z_{\rm g})$ may significantly bias $\Sigma_{\rm cr}$. To conservatively account for the impact of $p(z_{\rm g})$ shapes, we include the $\sigma_{\rm phot}$ term in the $\Delta\Sigma_+$ covariance matrix (see Sect. 7.5), quantifying the relative difference between $\Sigma_{\rm cr}$ derived from $\overline{z}_{\rm g}$ and that derived from the full $p(z_{\rm g})$. The latter can be estimated by inverting the following equation

$$\langle\Sigma_{\rm cr}^{-1}\rangle = \int {\rm d}z_{\rm g}\, p(z_{\rm g})\, \Sigma_{\rm cr}^{-1}(z_{\rm g}, \overline{z})\,. \tag{7.10}$$

We derive the two estimates of $\Sigma_{\rm cr}$ in the cluster redshift bins used for the stacking procedure, obtaining $\sigma_{\rm phot} = 0.040$ (i.e. 4% difference) on average. This result agrees with that found by Bellagamba et al. (2019) in KiDS-DR3, namely $\sigma_{\rm phot} = 0.042$.

7.4.2 Background selection

Foreground and cluster galaxies dilute the weak-lensing signal, as their shapes are uncorrelated with the matter distribution of galaxy clusters. To assess the fraction of such contaminants, we rely on the spectroscopic sample presented in Sect. 7.1.2. We consider three lens redshifts, corresponding to the central values of the cluster redshift bins used in the analysis, namely $z = 0.2$, $z = 0.4$, and $z = 0.65$. Then we apply the background selection, defined in Eq. (7.4), to the KiDS-1000 galaxies which have a spectroscopic counterpart. The photometric and spectroscopic redshift distributions of the selected galaxies are displayed in Fig. 7.6. Both distributions account for the lensing weights, w, discussed in Sect. 7.2.2. We find 0.5%, 1.9%, and 8.3% of contaminants for $z = 0.2$, $z = 0.4$, and $z = 0.65$, respectively. We expect such contamination to be overestimated, as the spectroscopic sample is not representative of the photometric one at high redshifts. Indeed, the r-band distribution peaks at $r = 21$ for the spectroscopic sample and at $r = 23.5$ for the photometric one.

The source galaxy redshifts influence the lensing measurements through the critical density, $\Sigma_{\rm cr}$, which depends on the ratio $D_{\rm ls}/D_{\rm s}$ (see Eq. 2.36). We find that the median of the $D_{\rm ls}/D_{\rm s}$ distribution based on photo-zs differs from the one based on spectroscopy by 0.4%, -5.4%, and 9.1% for $z = 0.2$, $z = 0.4$, and $z = 0.65$, respectively. Thus, similarly to what done in the previous section, we introduce the parameter $\sigma_{\rm bkg}$ accounting for the contamination due to the background selection, amounting to $\sigma_{\rm bkg} = 0.004$, $\sigma_{\rm bkg} = 0.054$, and $\sigma_{\rm bkg} = 0.091$ for $z \in [0.1, 0.3]$, $z \in [0.3, 0.5]$, and $z \in [0.5, 0.8]$, respectively.

7.4.3 Orientation and projections

Optical cluster finders preferentially select haloes with the major axis aligned with the line of sight, as these objects produce a larger density contrast with respect to the distribution of the field galaxies. In stacked weak-lensing analyses, this leads to overestimated cluster masses (Corless & King 2008; Dietrich et al. 2014; Osato et al. 2018; Zhang et al. 2023; Euclid Collaboration: Giocoli et al. 2023). An opposite mass bias comes from projections of secondary haloes aligned with the detected clusters, which are in fact blended in a single detection. This may cause an overestimation of the cluster optical mass proxies, and a consequent bias

Table 7.1: Parameters considered in the joint analysis of the stacked weak-lensing galaxy cluster profiles.

Parameter	Description	Prior	Posterior
α	Normalisation of the mass-richness relation	[-2, 2]	$0.18^{+0.04}_{-0.03}$
β	Slope of the mass-richness relation	[0, 5]	$1.73^{+0.10}_{-0.09}$
γ	Redshift evolution of the mass-richness relation	[-5, 5]	$-0.91^{+0.49}_{-0.51}$
$\sigma_{\rm intr,0}$	Normalisation of $\sigma_{\rm intr}$	[0.05, 1]	—
$\sigma_{\rm intr,\lambda^*}$	λ^* evolution of $\sigma_{\rm intr}$	[-5, 5]	—
c_0	Normalisation of the concentration-mass relation	[0, 1.3]	$0.42^{+0.27}_{-0.15}$
c_M	Mass evolution of the concentration-mass relation	[-1.5, 1.5]	$-0.11^{+0.18}_{-0.21}$
c_z	Redshift evolution of the concentration-mass relation	[-5, 5]	$0.29^{+1.46}_{-1.32}$
$f_{\rm off}$	Fraction of miscentred clusters	[0, 0.5]	—
$\sigma_{\rm off}$	Miscentring scale (in h^{-1}Mpc)	[0, 0.5]	—
$F_{\rm t}$	Truncation factor of the BMO density profile	$\mathcal{N}(3, 0.5)$	$2.74^{+0.45}_{-0.45}$

Notes. In the first and second columns we list the symbols and descriptions of the parameters, respectively. The third column lists the priors on the parameters, and in particular a range represents a uniform prior, while $\mathcal{N}(\mu, \sigma)$ stands for a Gaussian prior with mean μ and standard deviation σ. In the fourth column, we show the median values of the 1D marginalised posteriors, along with the 16th and 84th percentiles.

in the mass-richness relation (Myles et al. 2021; Wu et al. 2022). For the AMICO clusters in KiDS-DR3, Bellagamba et al. (2019) found that orientation effects counterbalance those caused by projections, leading to a negligible bias on the derived masses with a residual uncertainty of 3%. Also Simet et al. (2017) and Melchior et al. (2017), for redMaPPer clusters detected in DES, found that the combination of these bias sources is consistent with zero, deriving a mass corrective factor of 0.98 ± 0.03. Thus, we conservatively assume a residual uncertainty on mass due to orientation and projections of 4%.

To relate this uncertainty on mass to that on the weak-lensing profiles, we express the logarithmic dependence of $\Delta\Sigma_+$ on the mass M as follows (Melchior et al. 2017)

$$\Gamma = \frac{\mathrm{d} \ln \Delta\Sigma_+(M)}{\mathrm{d} \ln M} \,, \tag{7.11}$$

with $\Gamma \simeq 0.7$ being a good approximation for a broad range of cluster masses, redshifts, and cluster-centric distances (Melchior et al. 2017; Sereno et al. 2017; Bellagamba et al. 2019). From Eq. (7.11), we obtain

$$\frac{\delta\Delta\Sigma_+}{\Delta\Sigma_+} = \Gamma \frac{\delta M}{M} \,. \tag{7.12}$$

Based on Eq. (7.12), assuming a 4% uncertainty in mass due to orientation and projections results in a $\sigma_{\rm OP} = 0.028$ uncertainty in $\Delta\Sigma_+$.

7.5 Likelihood and priors

To derive constraints on the model parameters defined in Sect. 7.3, we perform a Bayesian analysis by means of an MCMC algorithm. For each stacked weak-lensing profile, defined in a bin of redshift and mass proxy, we assume a Gaussian likelihood function, defined as follows

$$\mathcal{L} \propto \exp(-\chi^2/2) \,, \tag{7.13}$$

with

$$\chi^2 = \sum_{i=1}^{N_\mathrm{R}} \sum_{j=1}^{N_\mathrm{R}} \left(\Delta\Sigma_{+,i}^{\mathrm{obs}} - \Delta\Sigma_{+,i}^{\mathrm{mod}}\right) C_{ij}^{-1} \left(\Delta\Sigma_{+,j}^{\mathrm{obs}} - \Delta\Sigma_{+,j}^{\mathrm{mod}}\right) , \tag{7.14}$$

where $\Delta\Sigma_{+}^{\mathrm{obs}}$ represents the observed cluster profile, given by Eq. (3.9), $\Delta\Sigma_{+}^{\mathrm{mod}}$ is the model, namely Eq. (7.7), while the indices i and j run over the number of radial bins, namely N_R, and C_{ij}^{-1} is the inverse of the covariance matrix. In particular, C_{ij} is defined as follows

$$C_{ij} = C_{ij}^{\mathrm{BS}} + C_{ij}^{\mathrm{sys}} , \tag{7.15}$$

where C_{ij}^{BS} is estimated through a bootstrap resampling, as discussed in Sect. 7.2.2, while C_{ij}^{sys} accounts for the systematic uncertainties defined in Sect. 7.4, and it is written as follows

$$C_{ij}^{\mathrm{sys}} = (\sigma_{\mathrm{phot}}^2 + \sigma_{\mathrm{bkg}}^2 + \sigma_{\mathrm{OP}}^2)\, \Delta\Sigma_{+,i}^{\mathrm{obs}} \Delta\Sigma_{+,j}^{\mathrm{obs}} . \tag{7.16}$$

Systematic errors are considered as correlated since they impact all cluster bins in a similar fashion. We jointly model the weak-lensing signal in all redshift and mass proxy bins, neglecting the correlation between radial bins in different stacks as discussed in Sect. 7.2.2.

In Table 7.1 we list the parameter priors adopted in the analysis. We assume uniform priors on the mass-richness scaling relation parameters in Eq. (5.7), namely α, β, and γ, and on the intrinsic scatter parameters in Eq. (7.8), that is $\sigma_{\mathrm{intr},0}$ and $\sigma_{\mathrm{intr},\lambda^*}$. Uniform priors are also assumed for the concentration-mass relation parameters in Eq. (7.9), namely c_0, c_M, and c_z, as well as for the miscentring parameters f_{off} and σ_{off}, defined in Sect. 7.3. The priors on f_{off} and σ_{off}, namely $f_{\mathrm{off}} \in [0, 0.5]$ and $\sigma_{\mathrm{off}} \in [0, 0.5]$ h^{-1}Mpc, agree with the comparisons of cluster centres derived from simulated (Yan et al. 2020; Sommer et al. 2023) and observed (Saro et al. 2015; Zhang et al. 2019; Seppi et al. 2023) galaxy, ICM, and dark matter distributions. For the truncation factor of the BMO profile, F_t, we assume a Gaussian prior with mean equal to 3 and an rms of 0.5, in agreement with what derived by Oguri & Hamana (2011).

7.6 Results

In the last column of Table 7.1 we list the constraints obtained for the free model parameters considered in the analysis. Regarding the parameters of the mass-richness relation, we derive $\alpha = 0.18^{+0.04}_{-0.03}$, $\beta = 1.73^{+0.10}_{-0.09}$, $\gamma = -0.91^{+0.49}_{-0.51}$. The statistical uncertainty on α is similar to that derived by Bellagamba et al. (2019), while the one on γ shows a 20% improvement, and the uncertainty on β is 20% larger. We remark that the analysis carried out in this work is not fully comparable with that performed by Bellagamba et al. (2019). The modelling procedure is different, as we derive the mass-richness relation directly from the stacked profiles. Also the set of free parameters differs from that assumed by Bellagamba et al. (2019), since we marginalise the mass-richness relation over σ_{intr} and consider a mass and redshift evolution of c_{200}. Furthermore, the galaxy weighted density in the KiDS-1000 shear sample, amounting to 6.17 arcmin^{-2} (Giblin et al. 2021), is lower than that associated with the KiDS-DR3 sample, namely 8.53 arcmin^{-2} (Hildebrandt et al. 2017), as a consequence of redshift and SOM-gold selections (Wright et al. 2020; Giblin et al. 2021).

The parameters $\sigma_{\mathrm{intr},0}$ and $\sigma_{\mathrm{intr},\lambda^*}$, along with the miscentring parameters f_{off} and σ_{off}, are not constrained. We expect that the joint analysis of cluster weak-lensing and counts will yield competitive results on $\sigma_{\mathrm{intr},0}$ and $\sigma_{\mathrm{intr},\lambda^*}$ (see, e.g., Murata et al. 2019; Lesci et al. 2022a). Regarding the concentration-mass relation in Eq. (7.9), we derive $c_0 = 0.42^{+0.27}_{-0.15}$,

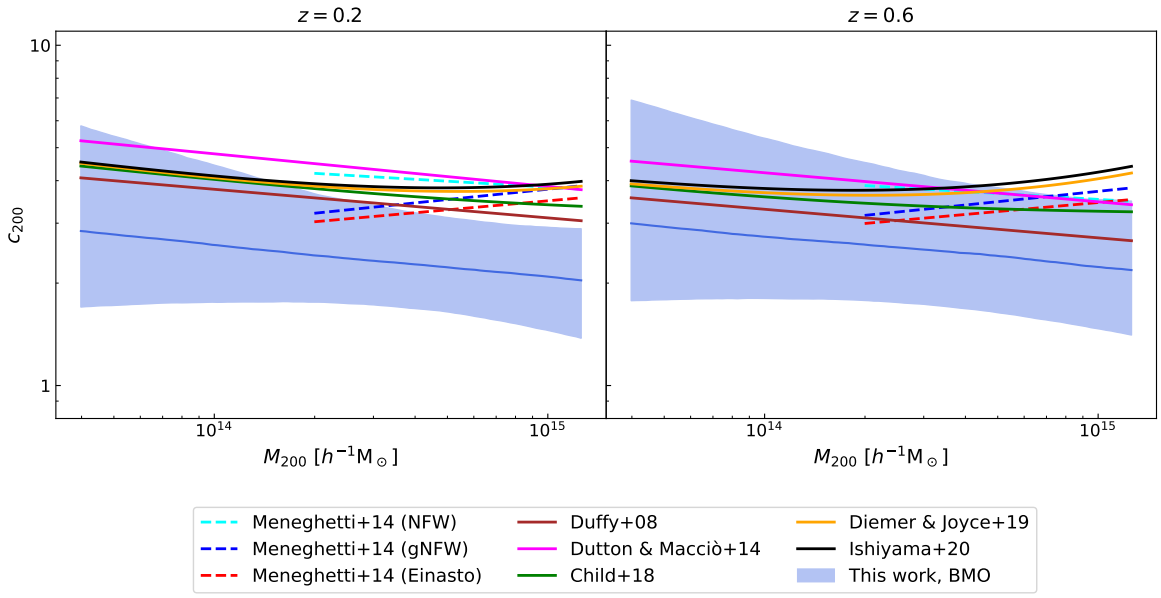

Figure 7.7: Concentration as a function of M_{200} assuming $z = 0.2$ (left panel) and $z = 0.6$ (right panel). The results by Duffy et al. (2008) (solid brown), Dutton & Macciò (2014) (solid magenta), Child et al. (2018) (solid green), Diemer & Joyce (2019) (solid orange), and Ishiyama et al. (2021) (solid black), obtained assuming an NFW profile, are shown. The dashed lines display the results by Meneghetti et al. (2014) assuming NFW (cyan), generalised NFW (gNFW, Zhao 1996) (blue), and Einasto (Einasto 1965) (red) profiles. The blue band shows the 68% confidence level of the model in Eq. (7.9). The covered mass range is based on the mass-richness relation constraints listed in Table 7.1.

$c_M = -0.11^{+0.18}_{-0.21}$, and $c_z = 0.29^{+1.46}_{-1.32}$. As we show in Fig. 7.7, these constraints agree within 1σ with the typical predictions obtained from simulations (Duffy et al. 2008; Dutton & Macciò 2014; Meneghetti et al. 2014; Child et al. 2018; Diemer & Joyce 2019; Ishiyama et al. 2021) at high redshift, while we find a 2σ agreement at low redshift and high mass. The mass range covered in Fig. 7.7 is based on the mass-richness relation constraints. In addition, we find $F_\mathrm{t} = 2.74^{+0.45}_{-0.45}$, with the posterior median showing a 0.5σ shift with respect to the median of the prior distribution. This agrees with the results by Oguri & Hamana (2011), who found $F_\mathrm{t} = 3$ by fixing the concentration parameter, and $F_\mathrm{t} = 2.6$ when the concentration is fitted to the data. Finally, in Fig. 7.8 we show the 1D and 2D posterior distributions of the main parameters constrained in the analysis.

We verified that by replacing the BMO density profile (Eq. 3.13) with a simple NFW model (Eq. 3.11), the median values of α, β, and γ are shifted within the 1σ interval of the posteriors reported in Table 7.1. In addition, the intrinsic scatter parameters are not constrained, in agreement with the results based on the BMO profile. On the other hand, we find that by assuming an NFW profile the miscentring parameters are constrained. In fact, we find $f_\mathrm{off} = 0.43^{+0.05}_{-0.09}$ and $\sigma_\mathrm{off} = 0.28^{+0.04}_{-0.04}$ h^{-1}Mpc. Despite the result on σ_off is in agreement with observations and simulations, f_off is larger than expected (Yan et al. 2020; Sommer et al. 2023; Saro et al. 2015; Zhang et al. 2019; Seppi et al. 2023). In addition, the constraint on f_off suggests that a wider prior is required for this parameter, covering larger values of f_off. As a result of this excess of miscentred clusters, the assumption of an NFW profile yields a very steep concentration-mass relation with a large normalisation, namely $c_0 = 1.09^{+0.15}_{-0.19}$, $c_M = -0.66^{+0.24}_{-0.24}$, and $c_z = 0.80^{+1.73}_{-1.67}$. Such relation disagrees with simulations for $M_{200} < 2 \times 10^{14}$ $h^{-1}\mathrm{M}_\odot$.

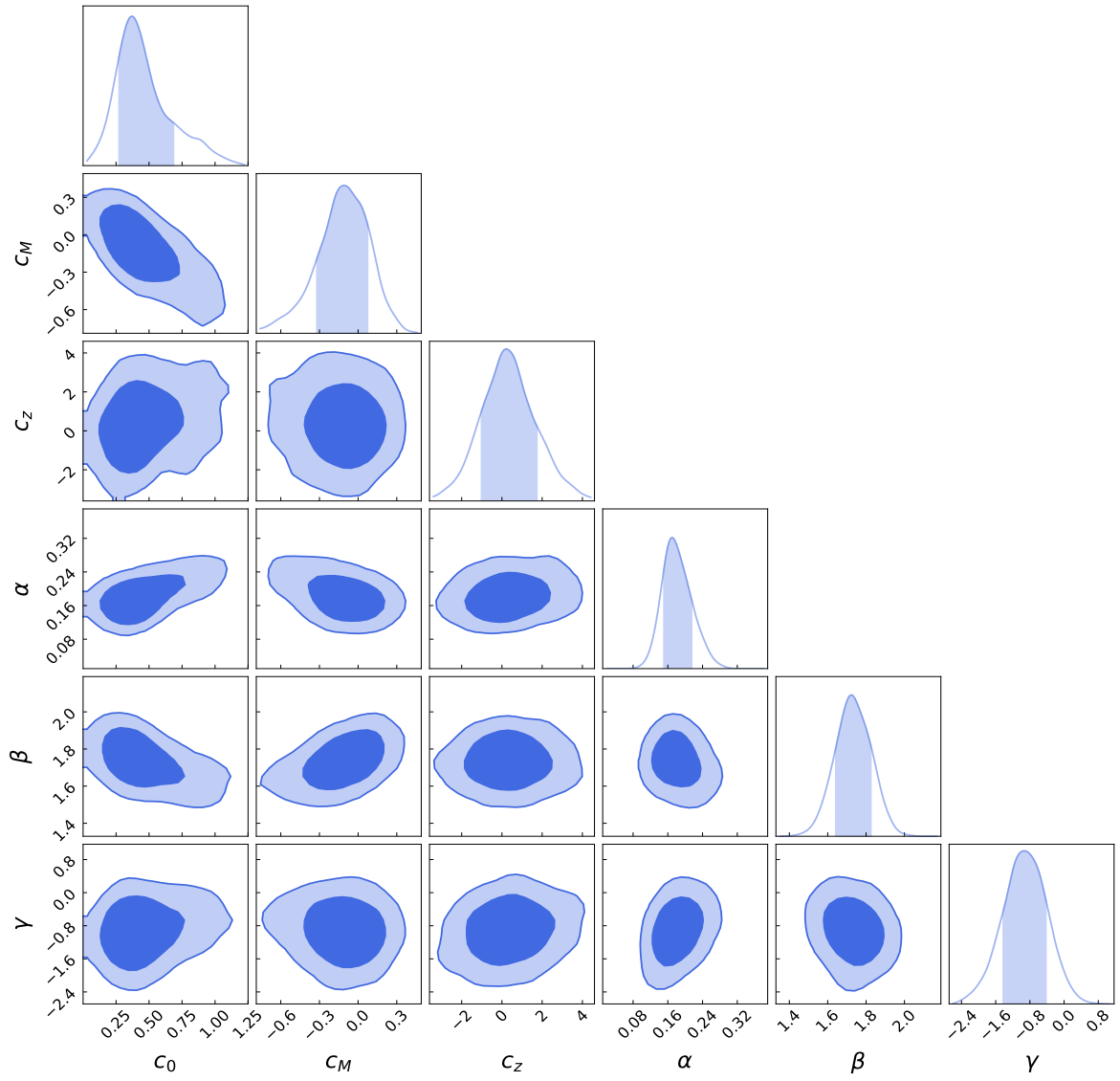

Figure 7.8: Constraints on c_0, c_M, c_z, α, β, and γ. The shown results are marginalised over the posteriors of the other parameters listed in Table 7.1. The confidence ellipses correspond to 68% and 95%, while the bands over the 1D marginalised posteriors represent the intervals between 16th and 84th percentiles.

7.7 Summary and perspectives

In this Chapter, we presented the work by Lesci et al. (in prep. b) on the weak-lensing mass calibration of the AMICO galaxy cluster sample provided by Maturi et al. (in prep.), based on KiDS-1000 Kuijken et al. (2019). This cluster sample extends over an effective area of 840 deg^2, which is more than double the effective area obtained in KiDS-DR3 (Maturi et al. 2019). Also the redshift range covered by the KiDS-1000 cluster sample, reaching $z = 1$, is much larger than that obtained by Maturi et al. (2019), reaching up to $z = 0.8$. As a result, the number of clusters is three times larger than that derived in KiDS-DR3, corresponding to 24 396 clusters with S/N > 3.5.

We carried out the analysis within the redshift range $z \in [0.1, 0.8]$, considering richness cuts ensuring a purity larger than 99% for each cluster bin. Compared to the weak-lensing

analysis performed by Bellagamba et al. (2019) in KiDS-DR3, we extended the cluster redshift range and improved the modelling of the stacked cluster profiles. Indeed, we jointly modelled all the stacked profiles in bins of redshift and richness. This allowed us to simultaneously constrain the mass-richness and concentration-mass relations, marginalising over the miscentring parameters. Thus, the mass calibration presented in this work accounts for the main biases affecting weak-lensing analyses, also including the treatment of systematic uncertainties due to photo-z errors, background selection, halo orientation, and projection effects.

Regarding the mass-richness scaling relation parameters, we found $\alpha = 0.18^{+0.04}_{-0.03}$, $\beta = 1.73^{+0.10}_{-0.09}$, $\gamma = -0.91^{+0.49}_{-0.51}$. The statistical uncertainty on α is similar to that derived by Bellagamba et al. (2019), while the one on γ shows a 20% improvement, and the uncertainty on β is 20% larger. Furthermore, we found $c_0 = 0.42^{+0.27}_{-0.15}$, $c_M = -0.11^{+0.18}_{-0.21}$, and $c_z = 0.29^{+1.46}_{-1.32}$, implying that the concentration parameter c_{200}, defined in Eq. (7.9), is in agreement with simulations (Duffy et al. 2008; Dutton & Macciò 2014; Meneghetti et al. 2014; Child et al. 2018; Diemer & Joyce 2019; Ishiyama et al. 2021). We verified that BMO and NFW profiles provide statistically consistent results on the mass-richness relation. However, an excess of miscentred clusters is required to fit an NFW profile to the data, with this yielding a very steep concentration-mass relation which disagrees with simulations for $M_{200} < 2 \times 10^{14}$ $h^{-1}\mathrm{M}_\odot$. We will thoroughly investigate these results in the near future.

A series of additional tests shall be performed to further improve the analysis presented in this Chapter. We list them in the following:

- The redshift distribution of the galaxy spectroscopic sample provided by van den Busch et al. (2022) and Wright et al. (in press), introduced in Sect. 7.1, can be properly reweighted through a SOM algorithm (following Wright et al. 2020). This would improve the assessment of contamination by foreground and cluster galaxies detailed in Sect. 7.4.2.

- Based on the reference spectroscopic sample detailed in Sect. 7.1, the contamination can be estimated as a function of cluster-centric distances. Any excesses of contaminants close to the cluster centres would be due to cluster galaxies. In addition, if the contamination showed a trend with cluster radii, this should be accounted for in the likelihood function.

- In Sect. 7.2.2, we showed that $\Delta\Sigma_\times$ is consistent with zero. This is expected if the measured signal comes from gravitational lensing only. Another null test involves measuring $\Delta\Sigma_+$ around random points. In the absence of residual systematic uncertainties, such signal should be consistent with zero.

- In the model in Eq. (7.7) we neglected the impact of the uncertainties on the lensing-weighted quantities, namely $\lambda^*_{\rm eff}$, $z_{\rm eff}$, and $R_{\rm eff}$. To account for such uncertainties, amounting to 2% on average, the integrand in Eq. (7.7) must be convolved with the probability to have a true lensing-weighted quantity given an observed one.

- The AMICO algorithm provides two cluster mass proxies, namely λ^*, considered in this Chapter and in Chapt. 5, and the signal amplitude, A (Maturi et al. 2019, in prep.). We will also carry out a stacked weak-lensing analysis based on A. This will allow the assessment of the mass proxy impact on the cosmological analyses, based on counts and clustering, performed by Lesci et al. (in prep. c).

Chapter 8

Summary and conclusions

In this Thesis, we discussed the main achievements of modern cosmology, which led to the definition of the concordance ΛCDM model. We detailed the central role of galaxy clusters in cosmology, presenting a series of studies covering the main steps of state-of-the-art cosmological analyses based on these cosmic tracers, spanning from mass calibration to the cosmological results based on their statistics. In the following, we summarise the main original results of these studies.

- In Chapter 4, we focused on the analysis carried out by Lesci et al. (2023) on the 3D 2PCF of the galaxy cluster sample presented in Planck Collaboration XXVII (2016), based on observations of the SZ effect. It is known that the masses of these clusters are likely underestimated, due to the use of a scaling relation calibrated through X-ray observations. Recent studies, such as the ones by Planck Collaboration XXIV (2016) and Planck Collaboration VI (2020), have highlighted the need for high values of the Planck mass bias, denoted as $b_{\rm SZ}$, to reconcile the constraints on σ_8 obtained from Planck cluster counts and CMB observations. Simulations and weak-lensing observations suggest that $(1 - b_{\rm SZ}) \sim 0.8$, whereas reconciling cluster counts with CMB cosmological results would require a value of $(1 - b_{\rm SZ}) = 0.62 \pm 0.03$. Through the analysis of the Planck cluster 2PCF in the redshift bins $z < 0.2$ and $z \in [0.2, 0.8]$, and within the cluster-centric radial range $r \in [10, 150]$ h^{-1}Mpc, Lesci et al. (2023) obtained $(1 - b_{\rm SZ}) = 0.62 \pm 0.12$. This constraint agrees with the one obtained from the combination of CMB and cluster count observations, emphasizing the necessity for very low values of $(1 - b_{\rm SZ})$ to reconcile cluster statistics with CMB results. Finally, by employing the same $b_{\rm SZ}$ priors as Planck Collaboration XXIV (2016), which are derived from weak-lensing measurements, we obtained constraints on $\Omega_{\rm m}$ from the 2PCF. These constraints not only agree with those obtained from cluster counts, but are also competitive in terms of uncertainty. Furthermore, we derived a reduction of approximately 30% in the uncertainty of $\Omega_{\rm m}$ when cluster counts and 2PCF are combined. However, we note that σ_8 remains unconstrained in our analysis. Future stage IV CMB experiments (Abazajian et al. 2016) are expected to detect about 10^5 galaxy clusters based on the SZ effect. This will significantly improve the calibration of the hydrostatic mass bias from cluster clustering, and might shed light on the degeneracy between σ_8 and mass bias.

- In Chapter 5, we presented the analyses based on cluster counts and 2PCF performed by Lesci et al. (2022a) and Lesci et al. (2022b), respectively. These studies are based on the cluster catalogue provided by Maturi et al. (2019), built up running the AMICO algorithm (Bellagamba et al. 2018) on KiDS-DR3 (de Jong et al. 2017). The sample

contains 7988 galaxy clusters over an effective area of 377 deg^2 and in the redshift range $z \in [0.1, 0.8]$. In Lesci et al. (2022a), we considered a subsample containing 3652 galaxy clusters having an intrinsic richness of $\lambda^* \geq 20$ and redshift $z \in [0.1, 0.6]$. We also corrected the models accounting for the incompleteness and the impurity of the sample. From the joint analysis of cluster counts and the weak-lensing measurements provided by Bellagamba et al. (2019), we found $\Omega_{\rm m} = 0.24^{+0.03}_{-0.04}$, $\sigma_8 = 0.86^{+0.07}_{-0.07}$, and $S_8 \equiv \sigma_8(\Omega_{\rm m}/0.3)^{0.5} = 0.78^{+0.04}_{-0.04}$. These results are competitive with those derived from recent cosmic shear (Troxel et al. 2018; Hikage et al. 2019; Asgari et al. 2021), cluster counts (Bocquet et al. 2019; Costanzi et al. 2019), and CMB (Hinshaw et al. 2013; Planck Collaboration VI 2020) analyses. We also showed that the inclusion of cluster counts significantly improves the constraints on the mass-richness scaling relation derived from weak lensing only. In the 2PCF analysis performed by Lesci et al. (2022b), we adopted the same redshift selection used for the counts' analysis, namely $z \in [0.1, 0.6]$. Since our clustering model does not rely on sample completeness estimates, we chose a less restrictive richness threshold, namely $\lambda^* > 15$, which ensures a sample purity close to 100%. This yields a subsample of 4934 galaxy clusters. Assuming the scaling relation constraints derived by Lesci et al. (2022a) as priors, we obtained $\Omega_{\rm m} = 0.28^{+0.05}_{-0.04}$, $\sigma_8 = 0.82^{+0.14}_{-0.12}$, and $S_8 = 0.80^{+0.08}_{-0.08}$, which are consistent within 1σ with the results obtained from CMB experiments and from the most recent analyses of the late Universe. Furthermore, by assuming the cosmological results derived by Planck Collaboration VI (2020, Table 2, TT, TE, and EE+lowE) as priors, we derived a robust constraint on the normalisation of the mass-richness relation. This reaffirms the key importance of clustering in constraining cluster masses, as also shown in Chapter 4.

- In Chapter 6, we detailed the work by Euclid Collaboration: Lesci et al. (2023). In this study, we developed a method to derive galaxy colour selections for cluster weak-lensing analyses, given any set of photometric bands and considering all the available colour-colour combinations. In fact, an accurate selection of lensed background galaxies is crucial to derive a reliable cluster weak-lensing signal and, in turn, to obtain accurate cluster mass measurements to be used in cosmological analyses. Colour selections are among the most reliable tools to accomplish this task. Based on the galaxy catalogue by Bisigello et al. (2020), we calibrated selections relying on ground-based *griz* and *Euclid* $Y_{\rm E}J_{\rm E}H_{\rm E}$ bands, with purity larger than 97%. The *griz* selection provides completeness values ranging between 30% and 84%, in the lens limiting redshift range $z_{\rm l} \in [0.2, 0.8]$. Including the *Euclid* $Y_{\rm E}J_{\rm E}H_{\rm E}$ bands, leading to the calibration of a $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selection, enhances the completeness by up to 25 percent points within this $z_{\rm l}$ range, and extends the galaxy selection up to $z_{\rm l} = 1.5$. In addition, for the first time in the literature, we expressed such selections as a continuous function of $z_{\rm l}$. We showed that both *griz* and $grizY_{\rm E}J_{\rm E}H_{\rm E}$ selections provide stable results when applied to external real and simulated data, regardless the photometric aperture definition and the differences in the transmission curves of the photometric filters. We demonstrated that the combination of colour and photo-z selections leads to 95% completeness in *Euclid* galaxy mock data, and that colour selections do not disrupt the calibration of cosmic shear samples. In addition, we showed that the *griz* selection is purer at high redshift and more complete at low redshift, compared to ground-based colour selections found in the literature. One possible enhancement to the calibration method presented by Euclid Collaboration: Lesci et al. (2023) is the inclusion of a magnitude dependence in the colour cuts, which might mitigate the impact of the uncertainties affecting faint magnitudes (Schrabback et al. 2021). Also the inclusion of additional ground-based bands might remarkably improve the results. Furthermore, the colour selection calibrated by

Euclid Collaboration: Lesci et al. (2023) could improve the shear calibration in cluster weak-lensing analyses, based on dedicated multi-band cluster image simulations.

- In Chapter 7, we presented the galaxy cluster weak-lensing mass calibration carried out by Lesci et al. (in prep. b). This analysis is based on the AMICO cluster catalogue provided by Maturi et al. (in prep.), derived from KiDS-1000 (Kuijken et al. 2019). The cluster sample contains 24 396 objects with S/N $>$ 3.5, corresponding to three times the number derived in KiDS-DR3 (Maturi et al. 2019), over an effective area of 840 deg^2 and up to redshift $z = 1$. We measured the stacked cluster weak-lensing signal in bins of richness and redshift. Specifically, we considered the clusters in the redshift range $z \in [0.1, 0.8]$, assuming richness cuts that ensure a purity larger than 99% for each cluster bin. Through the joint modelling of all the stacked profiles, we simultaneously constrained the mass-richness and concentration-mass relations, marginalising over the halo miscentring parameters and accounting for the main systematic uncertainties affecting cluster weak-lensing measurements. The constraints on the mass-richness relation have a precision which is compatible to that obtained by Bellagamba et al. (2019) in KiDS-DR3. Indeed, despite the larger number of galaxy clusters, the galaxy weighted density in the KiDS-1000 shear sample, amounting to 6.17 arcmin^{-2} (Giblin et al. 2021), is lower than that associated with the KiDS-DR3 sample, namely 8.53 arcmin^{-2} (Hildebrandt et al. 2017), as a consequence of redshift and SOM-gold selections (Wright et al. 2020; Giblin et al. 2021). Regarding the parameters of the concentration-mass relation in Eq. (7.9), we found $c_0 = 0.42^{+0.27}_{-0.15}$, $c_M = -0.11^{+0.18}_{-0.21}$, and $c_z = 0.29^{+1.46}_{-1.32}$. These results agree with simulations (Duffy et al. 2008; Dutton & Macciò 2014; Meneghetti et al. 2014; Child et al. 2018; Diemer & Joyce 2019; Ishiyama et al. 2021). We verified that BMO and NFW profiles provide statistically consistent results on the mass-richness relation. However, an excess of miscentred clusters is required to fit an NFW profile to the data, and this implies a very steep concentration-mass relation which disagrees with simulations for $M_{200} < 2 \times 10^{14}\ h^{-1}\mathrm{M}_\odot$. We plan to conduct a comprehensive investigation into these findings in upcoming studies.

The methodologies presented in this thesis have been contributing to the enhancement of the cosmological pipelines developed for the *Euclid* mission (Laureijs et al. 2011). Particularly, the techniques outlined in Chapter 7, which pertain to the stacking of cluster weak-lensing signals and the associated mass calibration process, are poised to become foundational components of analogous pipelines within the *Euclid* framework. We remark that mass calibration plays a pivotal role in cosmological analyses based on cluster statistics. Indeed, systematic uncertainties affecting weak-lensing cluster masses can lead to significant biases in the cosmological parameters derived from current surveys (Abbott et al. 2020).

Notably, the systematic exploration of both statistical and local properties of galaxy clusters is paving the way for novel inquiries on fundamental physics, which extend beyond the primary aim of the *Euclid* mission, that is constraining the dark energy equation of state. These methodologies are poised to deliver groundbreaking insights into diverse cosmological paradigms, ranging from the investigation of modified gravity theories (Dima & Vernizzi 2018; Langlois 2019; Cardone et al. 2021; Vogt et al. 2024) and the impact of massive neutrinos (Costanzi et al. 2013; Hagstotz et al. 2019), to elucidating the dynamics of interacting dark matter models (Peter et al. 2013; Robertson et al. 2017; Eckert et al. 2022) and non-minimally coupled dark matter scenarios (Bertolami et al. 2022; Gandolfi et al. 2023; Zamani et al. 2024). Thus, by leveraging the rich statistical and morphological information encoded within galaxy clusters, ongoing and future cosmological surveys have the potential to deepen our understanding of fundamental physical phenomena and unveil novel avenues for probing

the underlying fabric of the Universe.

Bibliography

Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, arXiv e-prints, arXiv:1610.02743

Abbott, T. M. C., Adamów, M., Aguena, M., et al. 2021, ApJ Suppl., 255, 20

Abbott, T. M. C., Aguena, M., Alarcon, A., et al. 2020, Phys. Rev. D, 102, 023509

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

Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, J. Cosm. Astro-Particle Phys., 2019, 056

Aguado-Barahona, A., Rubiño-Martín, J. A., Ferragamo, A., et al. 2022, A&A, 659, A126

Aguena, M., Avestruz, C., Combet, C., et al. 2021, MNRAS, 508, 6092

Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, Publ. Astron. Soc. Japan, 70, S8

Alcock, C. & Paczynski, B. 1979, Nature, 281, 358

Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, Annu. Rev. Astron. Astrophys., 49, 409

Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, Living Reviews in Relativity, 21, 2

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

Angulo, R. E. & Hahn, O. 2022, Living Reviews in Computational Astrophysics, 8, 1

Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, MNRAS, 426, 2046

Applegate, D. E., von der Linden, A., Kelly, P. L., et al. 2014, MNRAS, 439, 48

Arnaud, M. 2005, in Background Microwave Radiation and Intracluster Cosmology, ed. F. Melchiorri & Y. Rephaeli, 77

Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, A&A, 517, A92

Asgari, M., Lin, C.-A., Joachimi, B., et al. 2021, A&A, 645, A104

Balmès, I., Rasera, Y., Corasaniti, P. S., & Alimi, J. M. 2014, MNRAS, 437, 2328

Baltz, E. A., Marshall, P., & Oguri, M. 2009, J. Cosm. Astro-Particle Phys., 2009, 015

Bardeau, S., Soucail, G., Kneib, J. P., et al. 2007, A&A, 470, 449

Bartelmann, M. & Schneider, P. 2001, Phys. Rept., 340, 291

Bassett, B. & Hlozek, R. 2010, in Dark Energy: Observational and Theoretical Approaches, ed. P. Ruiz-Lapuente, 246

Bayliss, M. B., Ruel, J., Stubbs, C. W., et al. 2016, ApJ Suppl., 227, 3

Becker, M. R. & Kravtsov, A. V. 2011, ApJ, 740, 25

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

Bellagamba, F., Roncarelli, M., Maturi, M., & Moscardini, L. 2018, MNRAS, 473, 5221

Bellagamba, F., Sereno, M., Roncarelli, M., et al. 2019, MNRAS, 484, 1598

Beltz-Mohrmann, G. D. & Berlind, A. A. 2021, ApJ, 921, 112

Benítez, N. 2000, ApJ, 536, 571

Bertolami, O., Lima, M. M., & Mena, F. C. 2022, General Relativity and Gravitation, 54, 82

Beutler, F., Saito, S., Seo, H.-J., et al. 2014, MNRAS, 443, 1065

Bhattacharya, S., Heitmann, K., White, M., et al. 2011, ApJ, 732, 122

Bianchi, E. & Rovelli, C. 2010, arXiv e-prints, arXiv:1002.3966

Bisigello, L., Kuchner, U., Conselice, C. J., et al. 2020, MNRAS, 494, 2337

Biviano, A. & Mamon, G. A. 2023, A&A, 670, A17

Blandford, R. D. & Narayan, R. 1992, Annu. Rev. Astron. Astrophys., 30, 311

Bocquet, S., Dietrich, J. P., Schrabback, T., et al. 2019, ApJ, 878, 55

Bocquet, S., Saro, A., Dolag, K., & Mohr, J. J. 2016, MNRAS, 456, 2361

Böhringer, H., Schuecker, P., Guzzo, L., et al. 2004, A&A, 425, 367

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440

Borgani, S. 2008, in A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure, ed. M. Plionis, O. López-Cruz, & D. Hughes, Vol. 740, 24

Borgani, S. & Kravtsov, A. 2011, Advanced Science Letters, 4, 204

Borm, K., Reiprich, T. H., Mohammed, I., & Lovisari, L. 2014, A&A, 567, A65

Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601

Bridle, S. L., Kneib, J. P., Bardeau, S., & Gull, S. F. 2002, in The Shapes of Galaxies and their Dark Halos, ed. P. Natarajan, 38–46

Broadhurst, T., Takada, M., Umetsu, K., et al. 2005, ApJ Lett., 619, L143

Broadhurst, T. J., Ellis, R. S., Koo, D. C., & Szalay, A. S. 1990, Nature, 343, 726

Brunner, H., Liu, T., Lamer, G., et al. 2022, A&A, 661, A1

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

Cardone, V. F., Karmakar, P., De Petris, M., & Maoli, R. 2021, Phys. Rev. D, 103, 064065

Carlstrom, J. E., Holder, G. P., & Reese, E. D. 2002, Annu. Rev. Astron. Astrophys., 40, 643

Castro, T., Borgani, S., Dolag, K., et al. 2021, MNRAS, 500, 2316

Cavaliere, A. & Fusco-Femiano, R. 1976, A&A, 49, 137

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560

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

CHEX-MATE Collaboration. 2021, A&A, 650, A104

Child, H. L., Habib, S., Heitmann, K., et al. 2018, ApJ, 859, 55

Chiu, I. N., Okumura, T., Oguri, M., et al. 2020, MNRAS, 498, 2030

Chown, R., Omori, Y., Aylor, K., et al. 2018, ApJ Suppl., 239, 10

Clerc, N., Adami, C., Lieu, M., et al. 2014, MNRAS, 444, 2723

Clerc, N. & Finoguenov, A. 2023, in Handbook of X-ray and Gamma-ray Astrophysics. Edited by Cosimo Bambi and Andrea Santangelo, 123

Coles, P. & Lucchin, F. 2002, Cosmology: The Origin and Evolution of Cosmic Structure, Second Edition

Contreras, S. & Zehavi, I. 2023, arXiv e-prints, arXiv:2305.19628

Corasaniti, P.-S., Sereno, M., & Ettori, S. 2021, ApJ, 911, 82

Corless, V. L. & King, L. J. 2008, MNRAS, 390, 997

Costanzi, M., Rozo, E., Simet, M., et al. 2019, MNRAS, 488, 4779

Costanzi, M., Villaescusa-Navarro, F., Viel, M., et al. 2013, J. Cosm. Astro-Particle Phys., 2013, 012

Covone, G., Sereno, M., Kilbinger, M., & Cardone, V. F. 2014, ApJ Lett., 784, L25

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937

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

Cropper, M., Hoekstra, H., Kitching, T., et al. 2013, MNRAS, 431, 3103

Cropper, M., Pottinger, S., Niemi, S., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave, ed. H. A. MacEwen, G. G. Fazio, M. Lystrup, N. Batalha, N. Siegler, & E. C. Tong, 99040Q

Cuceu, A., Farr, J., Lemos, P., & Font-Ribera, A. 2019, J. Cosm. Astro-Particle Phys., 2019, 044

Cui, W., Borgani, S., Dolag, K., Murante, G., & Tornatore, L. 2012, MNRAS, 423, 2279

da Ângela, J., Outram, P. J., & Shanks, T. 2005, MNRAS, 361, 879

Dalal, R., Strauss, M. A., Sunayama, T., et al. 2021, MNRAS, 507, 4016

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371

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

de Bernardis, P., Ade, P. A. R., Bock, J. J., et al. 2000, Nature, 404, 955

de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, A&A, 582, A62

de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., et al. 2017, A&A, 604, A134

de la Torre, S. & Guzzo, L. 2012, MNRAS, 427, 327

Deng, Z., Xia, X., & Fang, L.-Z. 1994, ApJ, 431, 506

Despali, G., Giocoli, C., Angulo, R. E., et al. 2016, MNRAS, 456, 2486

Dhawan, S., Jha, S. W., & Leibundgut, B. 2018, A&A, 609, A72

Di Valentino, E., Mena, O., Pan, S., et al. 2021, arXiv e-prints, arXiv:2103.01183

Diemer, B. & Joyce, M. 2019, ApJ, 871, 168

Diemer, B. & Kravtsov, A. V. 2014, ApJ, 789, 1

Dietrich, J. P., Bocquet, S., Schrabback, T., et al. 2019, MNRAS, 483, 2871

Dietrich, J. P., Zhang, Y., Song, J., et al. 2014, MNRAS, 443, 1713

Dima, A. & Vernizzi, F. 2018, Phys. Rev. D, 97, 101302

Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, MNRAS, 390, L64

Dutton, A. A. & Macciò, A. V. 2014, MNRAS, 441, 3359

Eckert, D., Ettori, S., Robertson, A., et al. 2022, A&A, 666, A41

Eckert, D., Finoguenov, A., Ghirardini, V., et al. 2020, The Open Journal of Astrophysics, 3, 12

Edge, A., Sutherland, W., Kuijken, K., et al. 2013a, The Messenger, 154, 32

Edge, A., Sutherland, W., Kuijken, K., et al. 2013b, The Messenger, 154, 32

Einasto, J. 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

Einstein, A. 1915, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 778

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

Estrada, J., Sefusatti, E., & Frieman, J. A. 2009, ApJ, 692, 265

Ettori, S., Donnarumma, A., Pointecouteau, E., et al. 2013, Space Science Reviews, 177, 119

Ettori, S. & Molendi, S. 2011, Memorie della Societa Astronomica Italiana Supplementi, 17, 47

Euclid Collaboration, Adam, R., Vannier, M., et al. 2019, A&A, 627, A23

Euclid Collaboration, Blanchard, A., Camera, S., et al. 2020, A&A, 642, A191

Euclid Collaboration: Bisigello, L., Conselice, C. J., Baes, M., et al. 2023, MNRAS, 520, 3529

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

Euclid Collaboration: Fumagalli, A., Saro, A., Borgani, S., et al. 2022, arXiv e-prints, arXiv:2211.12965

Euclid Collaboration: Giocoli, C., Meneghetti, M., Rasia, E., et al. 2023, arXiv e-prints, arXiv:2302.00687

Euclid Collaboration: Lesci, G. F., Sereno, M., Radovich, M., et al. 2023, arXiv e-prints, arXiv:2311.16239

Euclid Collaboration: Pocino, A., Tutusaus, I., Castander, F. J., et al. 2021, A&A, 655, A44

Euclid Collaboration: Saglia, R., De Nicola, S., Fabricius, M., et al. 2022, A&A, 664, A196

Euclid Collaboration: Scaramella, R., Amiaux, J., Mellier, Y., et al. 2022, A&A, 662, A112

Euclid Collaboration: Schirmer, M., Jahnke, K., Seidel, G., et al. 2022, A&A, 662, A92

Euclid Collaboration: Sereno, M. et al. in prep.

Everett, S., Yanny, B., Kuropatkin, N., et al. 2022, ApJ Suppl., 258, 15

Fenech Conti, I., Herbonnet, R., Hoekstra, H., et al. 2017, MNRAS, 467, 1627

Ferragamo, A., Barrena, R., Rubiño-Martín, J. A., et al. 2021, A&A, 655, A115

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

Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, ApJ, 473, 576

Fox, C., Mahler, G., Sharon, K., & Remolina González, J. D. 2022, ApJ, 928, 87

Frenk, C. S., White, S. D. M., Davis, M., & Efstathiou, G. 1988, ApJ, 327, 507

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

Fumagalli, A., Costanzi, M., Saro, A., Castro, T., & Borgani, S. 2023

Gandolfi, G., Haridasu, B. S., Liberati, S., & Lapi, A. 2023, ApJ, 952, 105

García-Farieta, J. E., Marulli, F., Moscardini, L., Veropalumbo, A., & Casas-Miranda, R. A. 2020, MNRAS, 494, 1658

Garrel, C., Pierre, M., Valageas, P., et al. 2022, A&A, 663, A3

Gatti, M., Sheldon, E., Amon, A., et al. 2021, MNRAS, 504, 4312

Giacconi, R., Murray, S., Gursky, H., et al. 1974, ApJ Suppl., 27, 37

Gianfagna, G., Rasia, E., Cui, W., et al. 2023, MNRAS, 518, 4238

Giblin, B., Heymans, C., Asgari, M., et al. 2021, A&A, 645, A105

Gil-Marín, H., Wagner, C., Verde, L., Porciani, C., & Jimenez, R. 2012, J. Cosm. Astro-Particle Phys., 2012, 029

Giocoli, C., Marulli, F., Moscardini, L., et al. 2021, A&A, 653, A19

Giocoli, C., Moscardini, L., Baldi, M., Meneghetti, M., & Metcalf, R. B. 2018, MNRAS, 478, 5436

Giocoli, C., Tormen, G., & Sheth, R. K. 2012, MNRAS, 422, 185

Giri, S. K. & Schneider, A. 2021, J. Cosm. Astro-Particle Phys., 2021, 046

Gorenstein, P., Fabricant, D., Topka, K., Harnden, F. R., J., & Tucker, W. H. 1978, ApJ, 224, 718

Grego, L., Carlstrom, J. E., Reese, E. D., et al. 2001, ApJ, 552, 2

Gruen, D., Seitz, S., Brimioulle, F., et al. 2014, MNRAS, 442, 1507

Gruen, D., Zhang, Y., Palmese, A., et al. 2019, MNRAS, 488, 4389

Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, Astron. J., 116, 3040

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108

Gwyn, S. D. J. 2012, Astron. J., 143, 38

Hagstotz, S., Costanzi, M., Baldi, M., & Weller, J. 2019, MNRAS, 486, 3927

Hamilton, A. J. S. 1992, ApJ Lett., 385, L5

Hamilton, A. J. S. 1998, in Astrophysics and Space Science Library, Vol. 231, The Evolving Universe, ed. D. Hamilton, 185

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

Hattori, M., Kneib, J., & Makino, N. 1999, Progress of Theoretical Physics Supplement, 133, 1

Henson, M. A., Barnes, D. J., Kay, S. T., McCarthy, I. G., & Schaye, J. 2017, MNRAS, 465, 3361

Herbonnet, R., Crawford, A., Avestruz, C., et al. 2022, MNRAS, 513, 2178

Herbonnet, R., Sifón, C., Hoekstra, H., et al. 2020, MNRAS, 497, 4684

Hernández-Martín, B., Schrabback, T., Hoekstra, H., et al. 2020, A&A, 640, A117

Heymans, C. & Heavens, A. 2003, MNRAS, 339, 711

Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323

Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, MNRAS, 427, 146

Hikage, C., Oguri, M., Hamana, T., et al. 2019, Publ. Astron. Soc. Japan, 71, 43

Hildebrandt, H., Choi, A., Heymans, C., et al. 2016, MNRAS, 463, 635

Hildebrandt, H., van den Busch, J. L., Wright, A. H., et al. 2021, A&A, 647, A124

Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454

Hilton, M., Hasselfield, M., Sifón, C., et al. 2018, ApJ Suppl., 235, 20

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJ Suppl., 208, 19

Ho, M., Ntampaka, M., Rau, M. M., et al. 2022, Nature Astronomy, 6, 936

Hoekstra, H., Herbonnet, R., Muzzin, A., et al. 2015, MNRAS, 449, 685

Hoekstra, H., Mahdavi, A., Babul, A., & Bildfell, C. 2012, MNRAS, 427, 1298

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Hudelot, P., Cuillandre, J. C., Withington, K., et al. 2012, VizieR Online Data Catalog, II/317

Hurier, G. & Lacasa, F. 2017, A&A, 604, A71

Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017, ApJ, 848, 128

Ibitoye, A., Tramonte, D., Ma, Y.-Z., & Dai, W.-M. 2022, ApJ, 935, 18

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Ingoglia, L., Covone, G., Sereno, M., et al. 2022, MNRAS, 511, 1484

Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, MNRAS, 506, 4210

Ivezic, Z., Axelrod, T., Brandt, W. N., et al. 2008, Serbian Astronomical Journal, 176, 1

Iye, M., Karoji, H., Ando, H., et al. 2004, Publ. Astron. Soc. Japan, 56, 381

Jackson, J. C. 1972, MNRAS, 156, 1P

Jarvis, M., Bernstein, G. M., Amon, A., et al. 2021, MNRAS, 501, 1282

Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, MNRAS, 460, 2245

Jeans, J. H. 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Johnston, D. E., Sheldon, E. S., Wechsler, R. H., et al. 2007, arXiv e-prints, arXiv:0709.1159

Jones, C. & Forman, W. 1984, ApJ, 276, 38

Käfer, F., Finoguenov, A., Eckert, D., et al. 2019, A&A, 628, A43

Kaiser, N. 1984, ApJ Lett., 284, L9

Kaiser, N. 1987, MNRAS, 227, 1

Kaiser, N. 1995, ApJ Lett., 439, L1

Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460

Keihänen, E., Kurki-Suonio, H., Lindholm, V., et al. 2019, A&A, 631, A73

Kerscher, M., Szapudi, I., & Szalay, A. S. 2000, ApJ Lett., 535, L13

Kilbinger, M. 2015, Reports on Progress in Physics, 78, 086901

Kilbinger, M., Fu, L., Heymans, C., et al. 2013, MNRAS, 430, 2200

Klein, M., Mohr, J. J., Desai, S., et al. 2018, MNRAS, 474, 3324

Kluge, M., Neureiter, B., Riffeser, A., et al. 2020, ApJ Suppl., 247, 43

Kneib, J.-P. & Natarajan, P. 2011, A&A Rev., 19, 47

Kodi Ramanah, D., Wojtak, R., Ansari, Z., Gall, C., & Hjorth, J. 2020, MNRAS, 499, 1985

Korytov, D., Rangel, E., Bleem, L., et al. 2023, The Open Journal of Astrophysics, 6, 24

Kravtsov, A. V., Vikhlinin, A., & Nagai, D. 2006, ApJ, 650, 128

Kron, R. G. 1980, ApJ Suppl., 43, 305

Kuijken, K. 1999, A&A, 352, 355

Kuijken, K. 2011, The Messenger, 146, 8

Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, A&A, 625, A2

Labatie, A., Starck, J.-L., Lachièze-Rey, M., & Arnalte-Mur, P. 2010, arXiv e-prints, arXiv:1009.1232

Lacasa, F. & Grain, J. 2019, A&A, 624, A61

Lacey, C. & Cole, S. 1994, MNRAS, 271, 676

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJ Suppl., 224, 24

Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64

Langlois, D. 2019, International Journal of Modern Physics D, 28, 1942006

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193

Le Brun, A. M. C., McCarthy, I. G., Schaye, J., & Ponman, T. J. 2014, MNRAS, 441, 1270

Le Brun, A. M. C., McCarthy, I. G., Schaye, J., & Ponman, T. J. 2017, MNRAS, 466, 4442

Lee, B. E., Le Brun, A. M. C., Haq, M. E., et al. 2018, MNRAS, 479, 890

Lemaître, G. 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

Lesci, G. F., Marulli, F., Moscardini, L., et al. 2022a, A&A, 659, A88

Lesci, G. F., Nanni, L., Marulli, F., et al. 2022b, A&A, 665, A100

Lesci, G. F., Veropalumbo, A., Sereno, M., et al. 2023, A&A, 674, A80

Lesci, G. F. et al. in prep. b

Lesci, G. F. et al. in prep. c

Lewis, A. & Challinor, A. 2006, Phys. Rept., 429, 1

Lewis, A. & Challinor, A. 2011, CAMB: Code for Anisotropies in the Microwave Background, Astrophysics Source Code Library, record ascl:1102.026

Li, P. 2023, ApJ Lett., 950, L14

Linde, A. 2007, in Lecture Notes in Physics, Berlin Springer Verlag, ed. M. Lemoine, J. Martin, & P. Peter, Vol. 738, 1

Linde, A. D. 1983, Physics Letters B, 129, 177

Linder, E. V. 2003, Phys. Rev. Lett., 90, 091301

Lindholm, V., Finoguenov, A., Comparat, J., et al. 2021, A&A, 646, A8

Liu, A., Bulbul, E., Ghirardini, V., et al. 2022, A&A, 661, A2

LSST Dark Energy Science Collaboration. 2012, arXiv e-prints, arXiv:1211.0310

LSST Dark Energy Science Collaboration. 2021, arXiv e-prints, arXiv:2101.04855

LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv e-prints, arXiv:0912.0201

Makiya, R., Ando, S., & Komatsu, E. 2018, MNRAS, 480, 3928

Mana, A., Giannantonio, T., Weller, J., et al. 2013, MNRAS, 434, 684

Mandelbaum, R., Lanusse, F., Leauthaud, A., et al. 2018a, MNRAS, 481, 3170

Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018b, Publ. Astron. Soc. Japan, 70, S25

Mandelbaum, R., Rowe, B., Bosch, J., et al. 2014, ApJ Suppl., 212, 5

Martinet, N., Schneider, P., Hildebrandt, H., et al. 2018, MNRAS, 474, 712

Marulli, F., Bianchi, D., Branchini, E., et al. 2012, MNRAS, 426, 2566

Marulli, F., Veropalumbo, A., García-Farieta, J. E., et al. 2021, ApJ, 920, 13

Marulli, F., Veropalumbo, A., & Moresco, M. 2016, Astronomy and Computing, 14, 35

Marulli, F., Veropalumbo, A., Sereno, M., et al. 2018, A&A, 620, A1

Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13

Maturi, M., Bellagamba, F., Radovich, M., et al. 2019, MNRAS, 485, 498

Maturi, M., Fedeli, C., & Moscardini, L. 2011, MNRAS, 416, 2527

Maturi, M., Meneghetti, M., Bartelmann, M., Dolag, K., & Moscardini, L. 2005, A&A, 442, 851

Maturi, M. et al. in prep.

McCarthy, I. G., Schaye, J., Bird, S., & Le Brun, A. M. C. 2017, MNRAS, 465, 2936

McClintock, T., Varga, T. N., Gruen, D., et al. 2019, MNRAS, 482, 1352

Medezinski, E., Battaglia, N., Umetsu, K., et al. 2018a, Publ. Astron. Soc. Japan, 70, S28

Medezinski, E., Broadhurst, T., Umetsu, K., et al. 2007, ApJ, 663, 717

Medezinski, E., Broadhurst, T., Umetsu, K., et al. 2010, MNRAS, 405, 257

Medezinski, E., Oguri, M., Nishizawa, A. J., et al. 2018b, Publ. Astron. Soc. Japan, 70, 30

Melchior, P., Gruen, D., McClintock, T., et al. 2017, MNRAS, 469, 4899

Melchior, P., Suchyta, E., Huff, E., et al. 2015, MNRAS, 449, 2219

Melin, J. B., Aghanim, N., Bartelmann, M., et al. 2012, A&A, 548, A51

Melin, J. B., Bartlett, J. G., & Delabrouille, J. 2006, A&A, 459, 341

Meneghetti, M., Rasia, E., Merten, J., et al. 2010, A&A, 514, A93

Meneghetti, M., Rasia, E., Vega, J., et al. 2014, ApJ, 797, 34

Meszaros, P. 1974, A&A, 37, 225

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., & van Waerbeke, L. 2007, MNRAS, 382, 315

Miralda-Escude, J. 1991, ApJ, 370, 1

Miyazaki, S., Komiyama, Y., Kawanomoto, S., et al. 2018a, Publ. Astron. Soc. Japan, 70, S1

Miyazaki, S., Oguri, M., Hamana, T., et al. 2018b, Publ. Astron. Soc. Japan, 70, S27

Mo, H. J. & White, S. D. M. 1996, MNRAS, 282, 347

Monna, A., Seitz, S., Balestra, I., et al. 2017, MNRAS, 466, 4094

Moresco, M., Veropalumbo, A., Marulli, F., Moscardini, L., & Cimatti, A. 2021, ApJ, 919, 144

Moscardini, L., Matarrese, S., & Mo, H. J. 2001, MNRAS, 327, 422

Moutard, T., Arnouts, S., Ilbert, O., et al. 2016, A&A, 590, A102

Murata, R., Oguri, M., Nishimichi, T., et al. 2019, Publ. Astron. Soc. Japan, 71, 107

Myles, J., Gruen, D., Mantz, A. B., et al. 2021, MNRAS, 505, 33

Naess, S., Aiola, S., Austermann, J. E., et al. 2020, J. Cosm. Astro-Particle Phys., 2020, 046

Nagai, D., Vikhlinin, A., & Kravtsov, A. V. 2007, ApJ, 655, 98

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Norberg, P., Baugh, C. M., Gaztañaga, E., & Croton, D. J. 2009, MNRAS, 396, 19

Oguri, M., Bayliss, M. B., Dahle, H., et al. 2012, MNRAS, 420, 3213

Oguri, M. & Hamana, T. 2011, MNRAS, 414, 1851

Okabe, N., Takada, M., Umetsu, K., Futamase, T., & Smith, G. P. 2010a, Publ. Astron. Soc. Japan, 62, 811

Okabe, N., Zhang, Y. Y., Finoguenov, A., et al. 2010b, ApJ, 721, 875

Orlowski-Scherer, J., Di Mascolo, L., Bhandarkar, T., et al. 2021, A&A, 653, A135

Osato, K., Nishimichi, T., Oguri, M., Takada, M., & Okumura, T. 2018, MNRAS, 477, 2141

Park, Y., Sunayama, T., Takada, M., et al. 2023, MNRAS, 518, 5171

Peacock, J. A., Cole, S., Norberg, P., et al. 2001, Nature, 410, 169

Peacock, J. A. & Dodds, S. J. 1996, MNRAS, 280, L19

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

Peebles, P. J. E. & Hauser, M. G. 1974, ApJ Suppl., 28, 19

Penna-Lima, M., Bartlett, J. G., Rozo, E., et al. 2017, A&A, 604, A89

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Peter, A. H. G., Rocha, M., Bullock, J. S., & Kaplinghat, M. 2013, MNRAS, 430, 105

Pierre, M. 2022, arXiv e-prints, arXiv:2205.05499

Pierre, M., Pacaud, F., Adami, C., et al. 2016, A&A, 592, A1

Piffaretti, R. & Valdarnini, R. 2008, A&A, 491, 71

Planck Collaboration VI. 2020, A&A, 641, A6

Planck Collaboration XX. 2014, A&A, 571, A20

Planck Collaboration XXIV. 2016, A&A, 594, A24

Planck Collaboration XXIX. 2014, A&A, 571, A29

Planck Collaboration XXVII. 2016, A&A, 594, A27

Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81

Pratt, G. W., Arnaud, M., Biviano, A., et al. 2019, Space Science Reviews, 215, 25

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

Qin, F., Howlett, C., Stevens, A. R. H., & Parkinson, D. 2022, ApJ, 937, 113

Qin, J., Melia, F., & Zhang, T.-J. 2021, MNRAS, 502, 3500

Rana, D., More, S., Miyatake, H., et al. 2023, MNRAS, 522, 4181

Rasia, E., Meneghetti, M., Martino, R., et al. 2012, New Journal of Physics, 14, 055018

Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. 2012, MNRAS, 425, 1951

Reischke, R., Maturi, M., & Bartelmann, M. 2016, MNRAS, 456, 641

Remolina González, J. D., Sharon, K., Mahler, G., et al. 2021, ApJ, 920, 98

Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, ApJ, 876, 85

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, Astron. J., 116, 1009

Robertson, A., Massey, R., & Eke, V. 2017, MNRAS, 465, 569

Roos, M. & Harun-or-Rashid, S. M. 2000, arXiv e-prints, astro

Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, ApJ, 785, 104

Rykoff, E. S., Rozo, E., Hollowood, D., et al. 2016, ApJ Suppl., 224, 1

Salvati, L., Douspis, M., & Aghanim, N. 2020, A&A, 643, A20

Salvati, L., Douspis, M., Ritz, A., Aghanim, N., & Babul, A. 2019, A&A, 626, A27

Salvati, L., Saro, A., Bocquet, S., et al. 2022, ApJ, 934, 129

Sampaio-Santos, H., Zhang, Y., Ogando, R. L. C., et al. 2021, MNRAS, 501, 1300

Saro, A., Bocquet, S., Rozo, E., et al. 2015, MNRAS, 454, 2305

Sartoris, B., Biviano, A., Fedeli, C., et al. 2016, MNRAS, 459, 1764

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses

Schneider, P. & Seitz, C. 1995, A&A, 294, 411

Schrabback, T., Applegate, D., Dietrich, J. P., et al. 2018, MNRAS, 474, 2635

Schrabback, T., Bocquet, S., Sommer, M., et al. 2021, MNRAS, 505, 3923

Scoccimarro, R. 2004, Phys. Rev. D, 70, 083007

Secco, L. F., Samuroff, S., Krause, E., et al. 2022, Phys. Rev. D, 105, 023515

Seppi, R., Comparat, J., Nandra, K., et al. 2023, A&A, 671, A57

Sereno, M. 2015, MNRAS, 450, 3665

Sereno, M., Covone, G., Izzo, L., et al. 2017, MNRAS, 472, 1946

Sereno, M. & Ettori, S. 2015, MNRAS, 450, 3675

Sereno, M., Ettori, S., Lesci, G. F., et al. 2020, MNRAS, 497, 894

Sereno, M., Veropalumbo, A., Marulli, F., et al. 2015, MNRAS, 449, 4147

Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. 2021, ApJ Suppl., 254, 24

Shan, H., Liu, X., Hildebrandt, H., et al. 2018, MNRAS, 474, 1116

Sheldon, E. S., Johnston, D. E., Frieman, J. A., et al. 2004, Astron. J., 127, 2544

Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1

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

Shi, D., Wang, X., Zheng, X., et al. 2023a, arXiv e-prints, arXiv:2303.09726

Shi, J., Sunayama, T., Kurita, T., et al. 2023b, arXiv e-prints, arXiv:2306.09661

Shirasaki, M., Lau, E. T., & Nagai, D. 2018, MNRAS, 477, 2804

Sifón, C., Hoekstra, H., Cacciato, M., et al. 2015, A&A, 575, A48

Simet, M. & Mandelbaum, R. 2015, MNRAS, 449, 1259

Simet, M., McClintock, T., Mandelbaum, R., et al. 2017, MNRAS, 466, 3103

Smith, G. P., Mazzotta, P., Okabe, N., et al. 2016, MNRAS, 456, L74

Sommer, M. W., Schrabback, T., Ragagnin, A., & Rockenfeller, R. 2023, arXiv e-prints, arXiv:2306.13187

Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676

Stern, C., Dietrich, J. P., Bocquet, S., et al. 2019, MNRAS, 485, 69

Sunayama, T. 2023, MNRAS, 521, 5064

Sunayama, T., Park, Y., Takada, M., et al. 2020, MNRAS, 496, 4468

Sunyaev, R. A. & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173

Sutherland, W., Emerson, J., Dalton, G., et al. 2015, A&A, 575, A25

Tang, L., Lin, H.-N., & Liu, L. 2023, Chinese Physics C, 47, 015101

Taruya, A. & Hiramatsu, T. 2008, ApJ, 674, 617

Taruya, A., Nishimichi, T., & Saito, S. 2010, Phys. Rev. D, 82, 063522

The Dark Energy Survey Collaboration. 2005, arXiv e-prints, astro

Thuan, T. X. & Izotov, Y. I. 2000, arXiv e-prints, astro

Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878

To, C., Krause, E., Rozo, E., et al. 2021a, Phys. Rev. Lett., 126, 141301

To, C.-H., Krause, E., Rozo, E., et al. 2021b, MNRAS, 502, 4093

Tonegawa, M., Park, C., Zheng, Y., et al. 2020, ApJ, 897, 17

Tozzi, P. 2007, in The Invisible Universe: Dark Matter and Dark Energy, ed. L. Papantonopoulos, Vol. 720, 125

Troxel, M. A., MacCrann, N., Zuntz, J., et al. 2018, Phys. Rev. D, 98, 043528

Tsujikawa, S. 2013, Classical and Quantum Gravity, 30, 214003

Umetsu, K. 2020, A&A Rev., 28, 7

Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, ApJ, 795, 163

van den Busch, J. L., Wright, A. H., Hildebrandt, H., et al. 2022, A&A, 664, A170

Vazza, F., Jones, T. W., Brüggen, M., et al. 2017, MNRAS, 464, 210

Velliscig, M., van Daalen, M. P., Schaye, J., et al. 2014, MNRAS, 442, 2641

Veropalumbo, A., Marulli, F., Moscardini, L., Moresco, M., & Cimatti, A. 2014, MNRAS, 442, 3275

Veropalumbo, A., Marulli, F., Moscardini, L., Moresco, M., & Cimatti, A. 2016, MNRAS, 458, 1909

Viola, M., Cacciato, M., Brouwer, M., et al. 2015, MNRAS, 452, 3529

Vogt, S. M. L., Bocquet, S., Davies, C. T., Mohr, J. J., & Schmidt, F. 2024, arXiv e-prints, arXiv:2401.09959

Voit, G. M. 2005, Reviews of Modern Physics, 77, 207

von der Linden, A., Mantz, A., Allen, S. W., et al. 2014, MNRAS, 443, 1973

Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, ApJ, 646, 881

Watson, W. A., Iliev, I. T., D'Aloisio, A., et al. 2013, MNRAS, 433, 1230

Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJ Suppl., 258, 11

Wicker, R., Douspis, M., Salvati, L., & Aghanim, N. 2022, in European Physical Journal Web of Conferences, Vol. 257, mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera, 00046

Wright, A. et al. in press

Wright, A. H., Hildebrandt, H., Kuijken, K., et al. 2019, A&A, 632, A34

Wright, A. H., Hildebrandt, H., van den Busch, J. L., & Heymans, C. 2020, A&A, 637, A100

Wu, H.-Y., Costanzi, M., To, C.-H., et al. 2022, MNRAS, 515, 4471

Xu, W., Ramos-Ceja, M. E., Pacaud, F., Reiprich, T. H., & Erben, T. 2022, A&A, 658, A59

Xu, Z., Adachi, S., Ade, P., et al. 2021, Research Notes of the American Astronomical Society, 5, 100

Xue, Y.-J. & Wu, X.-P. 2000, ApJ, 538, 65

Yan, Z., Raza, N., Van Waerbeke, L., et al. 2020, MNRAS, 493, 1120

Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2006, MNRAS, 373, 1159

Yu, H., Tozzi, P., Borgani, S., Rosati, P., & Zhu, Z. H. 2011, A&A, 529, A65

Zaldarriaga, M. & Seljak, U. 1999, Phys. Rev. D, 59, 123507

Zamani, S., Salzano, V., & Bettoni, D. 2024, arXiv e-prints, arXiv:2401.02221

Zhang, Y., Jeltema, T., Hollowood, D. L., et al. 2019, MNRAS, 487, 2578

Zhang, Z., Wu, H.-Y., Zhang, Y., et al. 2023, MNRAS, 523, 1994

Zhao, H. 1996, MNRAS, 278, 488

Zhu, Z., Xu, H., Hu, D., et al. 2021, ApJ, 908, 17

Zohren, H., Schrabback, T., Bocquet, S., et al. 2022, A&A, 668, A18

Zubeldia, Í. & Challinor, A. 2019, MNRAS, 489, 401

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

Appendix A

Colour selection parameterisation

In Tables A.1 and A.2 we report the parameterisation of *griz* and $griz Y_{\rm E} J_{\rm E} H_{\rm E}$ colour selections, respectively, described in Sect. 6.3.1.

Table A.1: Calibrated colour selection based on *griz* photometry. The listed colour conditions are combined through the $\vee$ logical operator.

Colour condition	Parameters	$z_{\rm l}$ range
$(g-r) < s\,(r-z) + c$	$s = 9.39\,z_{\rm l}^3 - 7.93\,z_{\rm l}^2 + 0.84\,z_{\rm l} + 1.69$ $c = -31.8\,z_{\rm l}^3 + 35.71\,z_{\rm l}^2 - 12.78\,z_{\rm l} + 1.04$	$[0.2, 0.7]$
$(g-r) < s\,(i-z) + c$	$s = 7.64\,z_{\rm l}^2 - 7.49\,z_{\rm l} + 2.75$ $c = -2.53\,z_{\rm l}^2 + 1.75\,z_{\rm l} - 0.32$	$[0.2, 0.8]$
$(r-i) > c$	$c = 1.28\,z_{\rm l} + 0.11$	$[0.2, 0.6]$
$(g-r) < s\,(r-i) + c$	$s = 2.07\,z_{\rm l}^2 - 1.75\,z_{\rm l} + 1.99$ $c = -25.25\,z_{\rm l}^3 + 28.86\,z_{\rm l}^2 - 11.72\,z_{\rm l} + 1.34$	$[0.2, 0.6]$
$(r-i) < s\,(r-z) + c$	$s = 8.05\,z_{\rm l}^3 - 14.37\,z_{\rm l}^2 + 6.87\,z_{\rm l} + 0.73$ $c = -8.42\,z_{\rm l}^3 + 16.11\,z_{\rm l}^2 - 9.54\,z_{\rm l} + 0.68$	$[0.2, 0.8]$
$(g-i) > s\,(i-z) + c$	$s = -9.17\,z_{\rm l} + 3.24$ $c = 7.07\,z_{\rm l} - 0.45$	$[0.2, 0.3]$
$(g-i) < s\,(r-z) + c$	$s = -47.54\,z_{\rm l}^4 + 84.36\,z_{\rm l}^3 - 53.03\,z_{\rm l}^2 + 13.64\,z_{\rm l} + 0.48$ $c = 56.05\,z_{\rm l}^4 - 107.76\,z_{\rm l}^3 + 72.88\,z_{\rm l}^2 - 21.09\,z_{\rm l} + 1.93$	$[0.2, 0.8]$
$(g-z) < s\,(r-z) + c$	$s = 1.70$ $c = -21.04\,z_{\rm l}^4 + 16.37\,z_{\rm l}^3 + 1.47\,z_{\rm l}^2 - 3.46\,z_{\rm l} + 0.59$	$[0.2, 0.7]$
$(g-r) < s\,(g-z) + c$	$s = -8.94\,z_{\rm l}^3 + 13.28\,z_{\rm l}^2 - 6.55\,z_{\rm l} + 1.73$ $c = -2.53\,z_{\rm l}^2 + 2.24\,z_{\rm l} - 0.80$	$[0.2, 0.7]$
$(g-r) < c$	$c = 1.46\,z_{\rm l}^2 - 1.43\,z_{\rm l} + 0.40$	$[0.2, 0.5]$
$(r-i) > s\,(i-z) + c$	$s = 8.45\,z_{\rm l}^2 - 6.93\,z_{\rm l} + 1.67$ $c = -2.53\,z_{\rm l}^2 + 3.48\,z_{\rm l} - 0.35$	$[0.2, 0.5]$
$(i-z) > c$	$c = -1.53\,z_{\rm l}^2 + 2.15\,z_{\rm l} - 0.01$	$[0.2, 0.6]$
$(g-z) > s\,(r-z) + c$	$s = -0.58\,z_{\rm l} - 1.42$ $c = -10.10\,z_{\rm l}^2 + 15.76\,z_{\rm l} - 0.52$	$[0.2, 0.5]$
$(g-i) < s\,(r-i) + c$	$s = 0.24\,z_{\rm l} + 1.53$ $c = -0.91\,z_{\rm l} + 0.33$	$[0.2, 0.6]$

Notes. In the first column, the colour conditions are listed. The parameters of such conditions are shown in the second column, while in the last column the $z_{\rm l}$ ranges are listed. From top to bottom, the ith row corresponds to the ith iteration of the iterative process described in Sect. 6.2.

Table A.2: Calibrated colour selection based on $grizY_{\rm E}J_{\rm E}H_{\rm E}$ photometry. The listed colour conditions are combined through the $\vee$ logical operator.

Colour condition	Parameters	$z_{\rm l}$ range
$(g-i) < s\,(r-H_{\rm E}) + c$	$s = 30.81\,z_{\rm l}^4 - 79.74\,z_{\rm l}^3 + 73.28\,z_{\rm l}^2 - 28.85\,z_{\rm l} + 5.07;\;\; c = -31.2\,z_{\rm l}^4 + 78.62\,z_{\rm l}^3 - 71.16\,z_{\rm l}^2 + 27.57\,z_{\rm l} - 4.41$	$[0.2, 1.0]$
$(g-i) < s\,(z-H_{\rm E}) + c$	$s = 1.41\,z_{\rm l}^2 - 2.42\,z_{\rm l} + 2.1;\;\; c = -2.42\,z_{\rm l}^3 + 3.67\,z_{\rm l}^2 - 1.57\,z_{\rm l} - 0.28$	$[0.2, 1.3]$
$(g-r) < s\,(r-z) + c$	$s = 0.39\,z_{\rm l} + 1.33;\;\; c = -5.23\,z_{\rm l}^2 + 2.6\,z_{\rm l} - 0.83$	$[0.2, 0.7]$
$(g-Y_{\rm E}) < s\,(z-H_{\rm E}) + c$	$s = -0.01\,z_{\rm l} + 1.65;\;\; c = 0.37\,z_{\rm l}^2 - 1.45\,z_{\rm l} + 0.08$	$[0.2, 1.5]$
$(r-i) > s\,(J_{\rm E}-H_{\rm E}) + c$	$s = 1.64\,z_{\rm l} - 1.33;\;\; c = -0.72\,z_{\rm l}^2 + 1.49\,z_{\rm l} + 0.41$	$[0.2, 0.6]$
$(g-i) < s\,(z-J_{\rm E}) + c$	$s = -0.05\,z_{\rm l}^4 - 2.44\,z_{\rm l}^3 + 6.19\,z_{\rm l}^2 - 4.35\,z_{\rm l} + 2.31;\;\; c = -0.29\,z_{\rm l}^2 - 0.35\,z_{\rm l} - 0.21$	$[0.2, 1.3]$
$(g-i) < s\,(i-J_{\rm E}) + c$	$s = -2.17\,z_{\rm l}^3 + 5.01\,z_{\rm l}^2 - 3.25\,z_{\rm l} + 2.11;\;\; c = -2.79\,z_{\rm l}^3 + 4.06\,z_{\rm l}^2 - 2.7\,z_{\rm l} - 0.18$	$[0.2, 1.1]$
$(g-r) < s\,(r-H_{\rm E}) + c$	$s = -25.93\,z_{\rm l}^3 + 42.07\,z_{\rm l}^2 - 22.02\,z_{\rm l} + 4.2;\;\; c = 14.96\,z_{\rm l}^3 - 30.3\,z_{\rm l}^2 + 18.03\,z_{\rm l} - 3.52$	$[0.2, 0.7]$
$(g-i) > s\,(i-Y_{\rm E}) + c$	$s = -5.31\,z_{\rm l} + 1.99;\;\; c = 7.07\,z_{\rm l} - 0.45$	$[0.2, 0.3]$
$(g-r) < s\,(i-Y_{\rm E}) + c$	$s = -26.87\,z_{\rm l}^4 + 63.55\,z_{\rm l}^3 - 51.34\,z_{\rm l}^2 + 17.08\,z_{\rm l} - 0.8;\;\; c = 18.25\,z_{\rm l}^4 - 46.35\,z_{\rm l}^3 + 38.77\,z_{\rm l}^2 - 13.53\,z_{\rm l} + 1.31$	$[0.2, 1.0]$
$(g-r) < s\,(Y_{\rm E}-H_{\rm E}) + c$	$s = 2.07\,z_{\rm l}^2 - 2.39\,z_{\rm l} + 1.76;\;\; c = -10.29\,z_{\rm l}^3 + 11.18\,z_{\rm l}^2 - 4.19\,z_{\rm l} + 0.33$	$[0.2, 0.7]$
$(g-r) < s\,(i-H_{\rm E}) + c$	$s = -4.02\,z_{\rm l}^3 + 8.91\,z_{\rm l}^2 - 5.87\,z_{\rm l} + 1.75;\;\; c = -2.55\,z_{\rm l}^3 + 0.24\,z_{\rm l}^2 + 1.44\,z_{\rm l} - 0.61$	$[0.2, 0.9]$
$(g-z) < s\,(z-H_{\rm E}) + c$	$s = 1.43\,z_{\rm l}^4 - 5.51\,z_{\rm l}^3 + 7.29\,z_{\rm l}^2 - 3.79\,z_{\rm l} + 2.23;\;\; c = 0.32\,z_{\rm l}^2 - 1.55\,z_{\rm l} - 0.17$	$[0.2, 1.4]$
$(g-Y_{\rm E}) < s\,(i-H_{\rm E}) + c$	$s = -2.25\,z_{\rm l}^4 + 6.28\,z_{\rm l}^3 - 6.63\,z_{\rm l}^2 + 3.09\,z_{\rm l} + 1.16;\;\; c = -1.86\,z_{\rm l}^3 + 3.75\,z_{\rm l}^2 - 3.43\,z_{\rm l} + 0.26$	$[0.2, 1.1]$
$(g-z) < s\,(i-J_{\rm E}) + c$	$s = -6.33\,z_{\rm l}^4 + 15.05\,z_{\rm l}^3 - 11.67\,z_{\rm l}^2 + 3.25\,z_{\rm l} + 1.41;\;\; c = 8.54\,z_{\rm l}^4 - 23.53\,z_{\rm l}^3 + 22.33\,z_{\rm l}^2 - 9.11\,z_{\rm l} + 0.74$	$[0.2, 1.2]$
$(r-z) > s\,(i-z) + c$	$s = -5.60\,z_{\rm l}^2 + 2.31\,z_{\rm l} - 0.61;\;\; c = 4.69\,z_{\rm l}^2 - 0.67\,z_{\rm l} + 0.80$	$[0.2, 0.7]$
$(g-J_{\rm E}) < s\,(r-H_{\rm E}) + c$	$s = 2.76\,z_{\rm l}^2 - 2.91\,z_{\rm l} + 2.14;\;\; c = -21.51\,z_{\rm l}^3 + 23.09\,z_{\rm l}^2 - 7.14\,z_{\rm l} + 0.12$	$[0.2, 0.7]$
$(r-i) > s\,(z-J_{\rm E}) + c$	$s = -1.21\,z_{\rm l}^2 + 0.8\,z_{\rm l} + 0.2;\;\; c = 1.01\,z_{\rm l} + 0.15$	$[0.2, 0.5]$
$(g-z) < s\,(i-H_{\rm E}) + c$	$s = -3.66\,z_{\rm l}^4 + 12.87\,z_{\rm l}^3 - 16.26\,z_{\rm l}^2 + 7.83\,z_{\rm l} + 0.36;\;\; c = 6.29\,z_{\rm l}^4 - 21.87\,z_{\rm l}^3 + 26.63\,z_{\rm l}^2 - 13.19\,z_{\rm l} + 1.43$	$[0.2, 1.3]$
$(g-i) < s\,(r-J_{\rm E}) + c$	$s = -2.48\,z_{\rm l}^3 + 6.11\,z_{\rm l}^2 - 5.46\,z_{\rm l} + 2.59;\;\; c = 11.18\,z_{\rm l}^4 - 27.18\,z_{\rm l}^3 + 20.47\,z_{\rm l}^2 - 4.48\,z_{\rm l} - 0.72$	$[0.2, 1.0]$
$(g-z) < s\,(g-J_{\rm E}) + c$	$s = -2.16\,z_{\rm l}^3 + 5.88\,z_{\rm l}^2 - 5.06\,z_{\rm l} + 2.12;\;\; c = 2.33\,z_{\rm l}^3 - 6.82\,z_{\rm l}^2 + 6.18\,z_{\rm l} - 2.32$	$[0.2, 1.3]$
$(i-Y_{\rm E}) > s\,(z-Y_{\rm E}) + c$	$s = -0.02\,z_{\rm l} + 0.42;\;\; c = 2.81\,z_{\rm l}^3 - 5.05\,z_{\rm l}^2 + 3.36\,z_{\rm l} + 0.08$	$[0.2, 0.8]$
$(g-r) < s\,(z-J_{\rm E}) + c$	$s = 13.22\,z_{\rm l}^4 + 31.92\,z_{\rm l}^3 - 25.03\,z_{\rm l}^2 + 7.6\,z_{\rm l} + 0.42;\;\; c = 14.13\,z_{\rm l}^4 - 34.81\,z_{\rm l}^3 + 27.79\,z_{\rm l}^2 - 9.36\,z_{\rm l} + 0.83$	$[0.2, 1.1]$
$(g-i) < s\,(i-H_{\rm E}) + c$	$s = -11.75\,z_{\rm l}^4 + 33.51\,z_{\rm l}^3 - 33.07\,z_{\rm l}^2 + 12.73\,z_{\rm l} - 0.33;\;\; c = 10.74\,z_{\rm l}^4 - 35.01\,z_{\rm l}^3 + 37.81\,z_{\rm l}^2 - 16.31\,z_{\rm l} + 1.62$	$[0.2, 1.1]$
$(g-J_{\rm E}) < s\,(i-H_{\rm E}) + c$	$s = -0.03\,z_{\rm l} + 1.7;\;\; c = 21.04\,z_{\rm l}^4 - 48.34\,z_{\rm l}^3 + 38.51\,z_{\rm l}^2 - 13.25\,z_{\rm l} + 1.21$	$[0.2, 0.9]$
$(r-i) < s\,(z-J_{\rm E}) + c$	$s = -0.66\,z_{\rm l}^3 + 1.27\,z_{\rm l}^2 - 0.3\,z_{\rm l} + 1.37;\;\; c = 0.58\,z_{\rm l}^2 - 1.84\,z_{\rm l} - 0.38$	$[0.2, 1.3]$
$(i-Y_{\rm E}) < s\,(i-H_{\rm E}) + c$	$s = 2.90\,z_{\rm l} + 0.49;\;\; c = -6.06\,z_{\rm l} + 0.66$	$[0.2, 0.3]$
$(r-z) > s\,(Y_{\rm E}-J_{\rm E}) + c$	$s = 0.58\,z_{\rm l} + 0.29;\;\; c = 1.01\,z_{\rm l} + 0.45$	$[0.2, 0.6]$
$(g-r) < s\,(z-H_{\rm E}) + c$	$s = 1.1\,z_{\rm l}^3 + 0.69\,z_{\rm l}^2 - 1.14\,z_{\rm l} + 1.09;\;\; c = -8.67\,z_{\rm l}^3 + 9.14\,z_{\rm l}^2 - 3.35\,z_{\rm l} + 0.24$	$[0.2, 0.9]$
$(g-H_{\rm E}) < s\,(r-H_{\rm E}) + c$	$s = 0.78\,z_{\rm l}^2 - 1.03\,z_{\rm l} + 1.89;\;\; c = -13.09\,z_{\rm l}^3 + 15.15\,z_{\rm l}^2 - 5.42\,z_{\rm l} + 0.22$	$[0.2, 0.7]$
$(g-Y_{\rm E}) < s\,(r-J_{\rm E}) + c$	$s = 12.07\,z_{\rm l}^3 - 15.52\,z_{\rm l}^2 + 5.71\,z_{\rm l} + 1.08;\;\; c = -33.67\,z_{\rm l}^3 + 42.93\,z_{\rm l}^2 - 17.52\,z_{\rm l} + 1.61$	$[0.2, 0.7]$

Notes. The table structure is analogous to that of Table A.1.