Galaxy cluster cosmology in the LSST era

Constantin Payerne

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : PHYS - Physique
Spécialité : Physique Subatomique et Astroparticules
Unité de recherche : Laboratoire de Physique Subatomique et Cosmologie

Cosmologie avec les amas de galaxies à l'ère du LSST

Galaxy cluster cosmology in the LSST era

Présentée par :
Constantin PAYERNE

Direction de thèse :

Céline COMBET Directrice de recherche, LPSC Grenoble	Directrice de thèse

Thèse soutenue publiquement le **5 octobre 2023**, devant le jury composé de :

Johann COHEN-TANUGI Directeur de recherche, LUPM Montpellier	Examinateur
Raphaël GAVAZZI Chargé de recherche, LAM Marseille	Rapporteur
Sophie HENROT-VERSILLÉ Directrice de recherche, IJCLab Orsay	Rapportrice
Sophie MAUROGORDATO Directrice de recherche, OCA Nice	Examinatrice
Frédéric MAYET Professeur des Universités, LPSC Grenoble	Président
Laura SALVATI Chargée de recherche, IAS Orsay	Examinatrice

Remerciements

Je tiens d'abord à remercier l'ensemble des membres de mon jury. Je remercie très chaleureusement Frédéric Mayet, pour sa parfaite exécution de la présidence. Je tiens à remercier sincèrement Sophie Henrot-Versillé et Raphaël Gavazzi pour leur précieuse contribution à l'évaluation de mon travail de recherche, qui a permis d'enrichir la qualité de ce manuscrit. Je remercie les membres du jury Laura Salvati, Sophie Maurogordato et Johann Cohen-Tanugi pour leur intérêt porté à mon travail de thèse. Merci à toutes et tous, pour vos retours et pour nos échanges.

Je remercie Arnaud Lucotte et Laurent Derome[1] pour m'avoir permis de réaliser mon travail de recherche au LPSC, dans de parfaites conditions, malgré les nombreuses péripéties hasardeuses liées au Covid.

Merci à Céline Combet pour m'avoir encadré tout au long de cette thèse, et d'avoir su me donner la liberté d'appréhender, à l'aune sévère et juste des mathématiques et par-delà les sens communs, l'immensité absurde et magnifique des espaces cosmologiques. De surcroît, je te remercie très sincèrement pour ton aide précieuse dans la préparation de ce manuscrit.

> "*Ça parait rien de tel que consommer la page blanche, j'en jure c'est d'une affaire unique en tout genre, peut-être y trouverait-on une lumière particulière. Ou plutôt un chant clair de pater noster au large d'une île simple et lisse, tenue à la dérive dans une mer d'idées agitée par les insomnies et les arômes de caféine.*
>
> *Quel orchestre se joue-t-il qu'oreille tendue ne saurait examiner ? L'argile de l'idée ça fond le corps dans l'âme tout ça, ça agrippe tout le cœur qui l'est dans l'effort, et l'esprit ça s'éreinte comme marche au soleil. Longtemps je me suis assoupi du haut de ce le long vertige à faire se taire mon souci, c'est dire comme c'est grande force qui bat le sommeil, ça s'emmêle les démons qui dansent à parquet rompu. Mais c'est chimère à l'œil rude qui veille au chevet, qui se dîne la panse du cantique des madeleines. Ce festin me rend mi-songeur et je suis soucieux du ventre des hommes. J'abonde dans le sens des pénitents qui trouvent dans leur foi ce qui les dépasse.*"
>
> Nostra, *Le cantique des madeleines*, 1997

Thanks Calum for your friendship all along these three years. We did a lot of amazing science. I hope our likelihood paper will be cited once again before Christmas, I am almost convinced that we are already famous in one or two Pinocchio universes. I am looking forward

[1] que je remercie une seconde fois pour la confiance presque aveugle qu'il m'a accordée, en me déléguant ces nombreuses missions d'enseignement à l'Université Grenoble Alpes, qui ont abondé tout au long de ma thèse et ont parfait les conditions de travail pour la préparation de ce doctorat.

to meeting you again in Bourgoin-Jailleu to eat a soupe du jour.

Mention à l'amitié, au bon souvenir de Corentin Hanser, ce qui rend la chose juste c'est que les grands hommes font les petites histoires. J'en tiens pour témoins nos travaux de prestige "*Le cercle est-il une ellipse particulière ?*" ou encore notre recueil "*Mille fausses démonstrations de la conjecture de Riemann*", qui épinglent la marque de l'excellence au tissu noble de l'esprit.

Eu agradeço ao Eduardo, esses 10 meses foram suficientes para nos tornarmos irmãos de coração. Voltaremos a nos encontrar na Europa ou no Brasil.

À tous les moins de 30 du troisième, Manon, Marina, Miren, Corentin, Alice, Sofia, Damien et Carolina, cela va sans dire que Nietzsche répétait sans cesse dans les couloirs "Dieu est mort", mais je maintiens qu'une bande de potes comme celle-là fait un pansement redoutable. Plus de 150 memes originaux en attestent, que je consulte à la lueur d'une nostalgie frémissante.

Je remercie la cafetière percolateur du troisième étage "*Le nerf et le tendon des civilisations*" (Aër, 1987) sérum à l'anémie de sommeil, puis irradié par les accords chromatiques des collines grenobloises qui percent mes fenêtres, des idées de la veille j'en reprends les chemins sinueux, à la mémoire de la tendinite de Sisyphe.

Par ordre d'apparition dans le milieu scientifique, je remercie Céline, Laurent, Johan, David, Calum, Frédéric, Juan, Laurence, Florian, Miren, Corinne, Valérie, Fabio, Michel, Dominique, Mariana, Marina, Camille, Yuanyuan, Sam, François, Eve, Bastien, Matt, Marine, Vincent, Thibault, Fabien, Nathan, Corentin, Carolina, Robin, Gaël, Pierre, Kostia, Ben, Sebastian, Markus, Cyrille, Manon, Marina, Eduardo, Alice, Sofia, Damien.

L'esprit a ses recoins tranquilles, à l'abri des secousses de l'âme, et en cet instant j'ai une pensée apaisée pour Cécile.

Merci à mes parents, mes frères et ma sœur pour leur soutien sans conditions. Ça fait long je sais bien, mais rassurez-vous, c'est encore bien loin d'être fini, et c'est sans compter l'impact des baryons. Merci à Honorine, qui a suivi de près mon goût pour le travail de l'esprit, dès que nous aperçûmes qu'une complicité naissait entre nos doigts enlacés à la chaleur de nos cabans moleskine.

– Eugène, je sais j'ai pas les derniers yeux tout neuf, mais je dois te dire mon petit que c'est crevant que t'as l'air songeur là, je me trompe de beaucoup ?

– Barnabé camarade, on a le temps ma montre avance, je vais te parler de ma dernière fulgurance. Qu'en penses-tu de "*retarder la bêtise ne précipite pas le succès*" ?

– Mon ami, il est tôt dans le matin, t'as pas la raison à l'endroit que je vois.

Je dédie ce manuscrit à Ibrahim, à ce vieux souvenir, que ma mémoire peine à faire revivre, mais dont les échos pénètrent toujours d'un aplomb royal, l'immensité silencieuse des espaces infinis. À ces esquisses de temps sur les bancs de l'école à croiser la lance de l'esprit au fer de la science.

Résumé

Les amas de galaxies sont des traceurs de la croissance des structures et de l'expansion de l'Univers. La mesure de leur abondance permet à la fois d'étudier la nature de l'énergie sombre et celle de la matière sombre, et figure aujourd'hui parmi les principales sondes cosmologiques exploitables à partir des futurs relevés de galaxies. La relation entre la masse des amas et le nombre de galaxies qui les composent, dite relation d'échelle, est un point clé pour l'utilisation cosmologique du comptage d'amas dans le visible. Cette relation est cependant un des facteurs limitant des analyses actuelles, car les masses des amas ne sont pas directement observables. Ces dernières peuvent néanmoins être estimées grâce à l'effet de lentille gravitationnelle. Les rayons lumineux issus de sources situées en arrière-plan des amas sont déviés par les champs gravitationnels de ces derniers; ainsi les images des galaxies que nous observons subissent une déformation cohérente permettant de reconstruire leur distribution en masse. Cette thèse se concentre sur différents aspects de l'inférence cosmologique à partir du lentillage et de l'abondance des amas, à l'ère des grands relevés cosmologiques tel que le Rubin LSST.

J'ai réalisé une partie de mon travail de thèse au sein de la Dark Energy Science Collaboration, qui prépare les analyses cosmologiques du LSST. J'ai contribué au développement de codes d'analyse pour exploiter l'effet de lentille gravitationnelle par les amas, et leur abondance, en utilisant les simulations de la collaboration. Ces simulations fournissent des catalogues de galaxies, idéaux et réalistes, tels que pourra fournir le LSST, et dont j'ai contribué à valider les propriétés. Par ailleurs, à l'aide de cette chaîne d'analyse que j'ai développée, j'ai étudié l'impact des divers effets systématiques sur la reconstruction de la masse des amas et comment ceux-ci se propagent jusqu'aux contraintes cosmologiques.

Les amas de galaxies ont une structure tri-dimensionnelle complexe et les effets de projection le long de la ligne de visée peuvent biaiser la reconstruction de la masse à partir du lentillage gravitationnel. Des informations sur la morphologie des amas peuvent être extraites en mesurant les moments multipolaires du champ de cisaillement gravitationnel. Dans le cadre du projet The Three Hundred, j'ai développé une analyse basée sur cette approche et vérifié si la reconstruction de la masse s'en trouvait améliorée. J'ai montré qu'il était possible de reconstruire ainsi l'ellipticité et l'orientation de la distribution de masse projetée. Cependant, le gain sur la reconstruction de la masse reste marginal et une forte dégénérescence existe entre la masse estimée et le paramètre d'allongement/aplatissement lorsque l'axe principal est aligné le long de la ligne de visée.

La fonction de vraisemblance (likelihood) décrit les propriétés statistiques de l'abondance des amas mesurée dans des intervalles de redshift et de masse, et est un ingrédient clé des analyses cosmologiques. J'ai développé une méthodologie novatrice pour tester la précision et la robustesse des likelihoods "binnées" utilisées dans la littérature pour le comptage d'amas (Poissonienne, Gaussienne et Gauss-Poisson (GPC)), que nous avons appliqué aux 1000 catalogues simulés de halos de matière noire PINOCCHIO. Nous avons conclu que pour des relevés de type LSST ou *Euclid*, la likelihood Poissonienne sous-estime les erreurs sur les paramètres cosmologiques de 20 à 30%, tandis que les likelihoods Gaussienne et GPC donnent des erreurs

robustes.

Par ailleurs, j'ai développé un nouveau formalisme de likelihood "non-binnée" en incluant l'effet de variance cosmique, source de bruit intrinsèque à toutes les observables liées aux fluctuations de densité de matière, tels que les amas de galaxies. Les contraintes cosmologiques obtenues sont plus fortes que dans l'approche binnée, mais ne diffèrent pas beaucoup de l'approche non-binnée standard.

Abstract

Galaxy clusters are powerful probes of the growth of structures and of the late-time expansion history of the Universe. Measuring their abundance enables us to investigate the nature of dark energy and dark matter, making them one of the primary cosmological probes to be used in future galaxy surveys.

The relation between cluster masses and the number of galaxies they contain, known as the scaling relation, is a key factor for the cosmological use of cluster counts in the optical domain. However, this relation remains one of the limiting factors in current analyses because cluster masses are not directly observable. Nevertheless, they can be estimated through the gravitational lensing effect, where light from background sources is deflected by the cluster gravitational fields. As a result, the observed galaxy images undergo a coherent distortion, allowing for the reconstruction of their mass distribution. This thesis focuses on various aspects of cosmological inference using cluster lensing and abundance in the era of wide cosmological surveys such as the Rubin LSST.

A central part of this work was performed as part of the LSST Dark Energy Science Collaboration (DESC), which focuses on the cosmological analyses using LSST data. I joined the DESC effort by contributing to the development of software tools for cluster lensing and cluster abundance analyses, using DESC simulated datasets. These datasets contain ideal and realistic LSST-like galaxy catalogs, the validation of which I contributed to. Furthermore, using the several codes I developed, I derived preliminary cosmological constraints from the cluster abundance in the DESC simulations. I also focused on the impact of the several systematic effects on the weak lensing mass reconstruction and how these propagate to the cosmological constraints.

Galaxy cluster triaxial morphology and resulting projection effects can yield biases on the mass reconstruction. Valuable information on cluster morphology can be extracted from the multipoles of the local shear and releasing the spherical symmetry assumption that is generally made. As part of the cluster simulation project The Three Hundred, I developed an analysis based on lensing shear multipoles to gain insights on the triaxial properties of the cluster mass distribution and check whether this could improve the lensing mass calibration. From this work, I found that including multipole moments of the local shear permits reconstructing the ellipticity and orientation of the projected mass distribution. However, gain on the mass reconstruction remains marginal and a strong degeneracy exists between the lensing mass estimate and elongation/flattening parameter for triaxial clusters whose major axis is aligned along the line-of-sight.

One of the key ingredients to constrain cosmology is the likelihood function, that describes the statistical properties of the observed cluster counts in bins of cluster redshift and cluster mass. I developed a methodology to test the accuracy and robustness of cluster abundance binned likelihoods that are widely used in the literature (the Poisson, the Gaussian and the Gauss-Poisson Compound likelihood (GPC)) and applied it to the 1,000 PINOCCHIO simulated dark matter halo catalogs. The main finding is that for a variety of cosmological inference setups, the Poisson likelihood underestimates errors by 20-30%, whereas the Gaussian and GPC give

similar and robust errors for LSST or *Euclid*-like surveys.

In addition, I have also developed a new unbinned cluster count formalism allowing us to account for the effect of Super-Sample Covariance, which is an inherent source of scatter for large scale structure observables such as galaxy clusters. The resulting cosmological constraints are stronger than in the binned approach but only slightly differ from the more standard unbinned method.

Introduction

" *God is nothing but the power of the universe to organize itself.*"

Lee Smolin, theoretical physicist in Ontario, Canada.

Anaximander, a pre-Socratic Greek philosopher (600 BC, painted by Raphaël, see Fig. 1) who lived in Miletus, an ancient Greek city on the modern Turkish coast, expressed profound concern for understanding the nature of the cosmos. He was the first to propose a mechanical model of the world, arranging the Sun, the Moon and stars in concentric rings around Earth, and initiated a *naturalistic* approach to understanding our Universe. He paved the way for modern science and is considered by many as one of the *Fathers of Cosmology* (Rovelli, 2023).

" *Today, the term precision cosmology is a reality,*
and cosmology is an exemplar of big-data science."

Turner (2022), *Road to Precision Cosmology.*

During the last few decades, cosmology has entered a new era of precision, propelled by the recent large sky surveys (e.g., WMAP, *Planck*, ACT, SPT, SDSS, DES, HSC) that probed high redshifts and large portions of the sky. They have provided the community with an unprecedented amount of high-quality data at different wavelengths to shed light on cosmic phenomena with great precision and map the properties and evolution of the Universe. The data allow cosmologists to test and refine the standard model of cosmology, by looking at several cosmic signatures, ranging from the relic radiation of the Big Bang to the culminant step of structure formation in the late Universe. The combination of these probes has permitted to obtain remarkable constraints that allowed us to shape with high accuracy our cosmological model and search for deviations from the standard cosmological paradigm. They also have delivered deep insights into the nature of dark energy and dark matter, two of the most important problems in modern cosmology. The most accurate cosmological constraints were obtained in 2018 from the analysis of the temperature and polarization anisotropies of the Cosmic Microwave Background (CMB), that was detected by the *Planck* satellite.

In parallel to that, the recent percent precision that was achieved has revealed some tensions between the cosmological constraints obtained from these different methods, more particularly between "early" and "late" probes, sensitive respectively to the physics occurring in the early (CMB) and late (distribution of galaxies and galaxy clusters) Universe. Through the improvement of the quality of observations and the increase in size of the cosmological data sets by the next generation surveys (e.g., Rubin LSST, *Euclid*, Simons Observatory) the current tensions might grow in significance, either revealing physics beyond our standard model of cosmology, or nuisance counterparts from uncontrolled systematic effects. The necessity of controlling the most subtle systematic effect on astronomical datasets has become one of the biggest challenges of observational cosmology, in order to maximize the scientific outcomes of future surveys and their potential synergies.

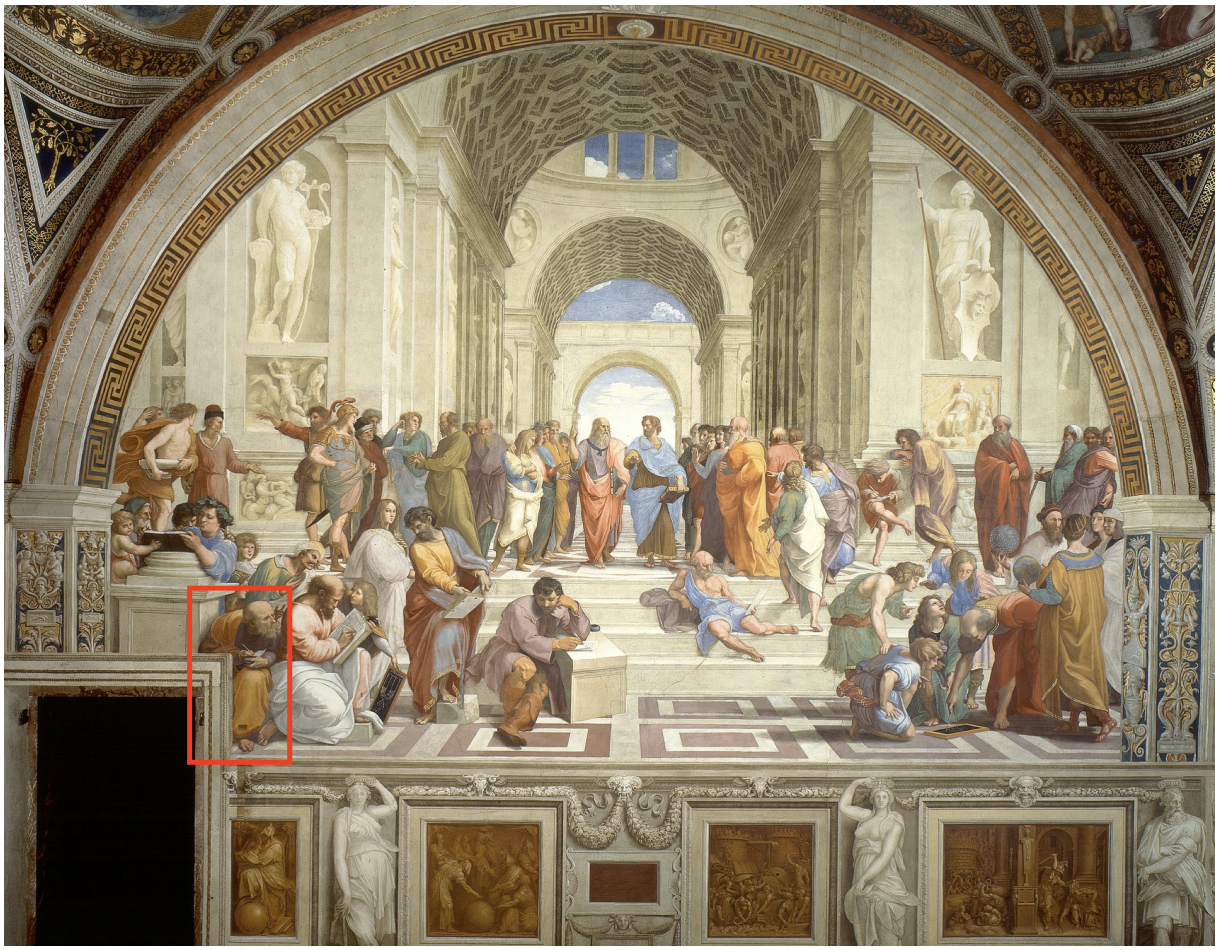

Figure 1: *Scuola di Atene* by Raphaël (1508-1512), Vatican Museum. The red square shows Anaximander, who was the first to conceive a mechanical model of the world. According to him, the Earth occupies a central position in the infinite, not supported by anything, because it has no reason to move in any direction. Earth lies at the center of a vast cosmic system, with celestial bodies arranged in concentric rings. His revolutionary ideas diverged from earlier mythological and religious interpretations, paving the way for rational and scientific investigations into understanding the fundamental nature of the cosmos.

In that context, galaxy clusters are of particular interest to provide a wealth of information on our cosmological model. Historically, they have been essential in the construction of the cosmological standard model, since they have provided some of the first evidence of this unseen matter through the motions of the intra-cluster galaxies. They are also powerful cosmological probes since their formation and abundance are sensitive to the properties of both dark energy and dark matter. Thanks to the performance of the recent and future cluster surveys, upcoming cosmological analyses will benefit from large cluster catalogs to refine cosmological models at unprecedented level.

One of the key ingredients of cluster count cosmology at optical wavelength is the cluster scaling relation that links the observed number of galaxies in clusters to their masses. The calibration of this relation is however one of the limiting factors of current analyses, since cluster masses are not directly observable and have to be inferred indirectly. As galaxy clusters correspond to the highest peaks in the matter density field, they provide us with some of the most dramatic effects of gravitational lensing, which can be used to infer cluster masses. More specifically, weak lensing in cluster fields induces a coherent distortion of background galaxy shapes that allows us to reconstruct their underlying matter distribution in a manner largely insensitive to the complex non-linear physics of galaxy clusters, which other mass reconstruction methods suffer from. This thesis focuses on different aspects of cosmological inference with galaxy cluster lensing and abundance, in the new era of wide cosmological surveys, and more particularly of the LSST, due to start in ~ 2025.

I started this work in 2020, at the peak of the pandemic. While my Ph.D. was planned for the first LSST data, the delay caused by the pandemic required us to redefine part of the work to use mainly simulated datasets. My manuscript is split into 6 chapters.

Chapter 1 presents the landscape of modern cosmology. We introduce the basic concepts of Einstein's general relativity and we present the standard model of cosmology. We also review the main cosmological probes and leading experiments, along with the tensions that have appeared in the recent analyses.

Chapter 2 focuses on the formalism of how galaxy clusters can be used to provide insight into cosmology through their abundance and their weak lensing signal. We also introduce the concepts of Bayesian inference in which likelihoods play a central role.

Chapter 3 introduces the Vera C. Rubin Observatory, the LSST and the Dark Energy Science Collaboration (DESC) that prepares the cosmological analysis of the LSST data. We also introduce the simulated data from DESC, the Data Challenge 2, that we will extensively use in the work presented in Chapter 4.

Chapter 4 presents the studies I undertook as part of the LSST DESC. I detail my contribution to the development of software tools for cluster lensing and cluster abundance, using DESC simulated data sets. Using the several codes that I developed, I present the preliminary cosmological constraints from the cluster abundance in the DESC simulations. I also present my work on the impact of several systematic effects on the weak lensing mass reconstruction and how these propagate to the cluster scaling relation.

Chapter 5 focuses on many aspects of cluster abundance likelihoods, that are key statistical ingredients to constrain cosmology. I present a new methodology to test the accuracy and robustness of cluster abundance likelihoods. In addition, I present a new unbinned cluster count formalism allowing us to account for the effect of Super-Sample Covariance, which is an inherent source of scatter for large-scale structure probes such as galaxy clusters.

Chapter 6 presents my work on weak lensing mass inference using multipolar moments of the cluster shear field. Galaxy cluster triaxial morphology and resulting projection effects can yield biases on the mass reconstruction. Valuable information on cluster morphology can be extracted by measuring the multipole coefficients of the cluster weak lensing shear and releasing

the spherical symmetry assumption that is generally made.

I present my conclusions and my perspectives in Chapter 7.

Contents

Chapter 1

Introduction: Theoretical and observational landscape of cosmology

Contents

The standard model of cosmology allows us to understand the evolution of the Universe, from the Big Bang to the present days, by using a restricted number of parameters and relying on simple assumptions about homogeneity and isotropy of the Universe. In particular, it provides a framework to accurately describe the evolution of matter perturbations leading to the emergence of large scale structures, including galaxy clusters, that are the focus of this thesis. Today, this model is favored by a variety of astrophysical and cosmological observations, and the key parameters are being continually refined from new observations. In particular, the next decade will permit to enter the high-precision cosmology era, that will benefit from large cosmological surveys from a variety of ground-based and space telescopes.

This introductory chapter aims to present both the key theoretical aspects and observational landscape of modern cosmology. In Section 1.1 we present the formalism of Einstein's general relativity as the foundation of the cosmological model. The solutions of Einstein's field equations for a homogeneous and isotropic Universe are shown in Section 1.2, which enable us to derive the evolution of the expansion as a function of the energy density content in the Universe. In

Section 1.3, we present the standard model of cosmology, and we introduce the main cosmological probes and the observational landscape in Section 1.4. Then we make a brief review of the challenges of modern cosmology in Section 1.5 and conclude in Section 1.6.

1.1 General relativity

The current models for cosmology are based on Einstein's theory of general relativity. General relativity is a successful geometric theory of gravitation that provides the natural framework for writing the laws of physics in curved space-time. This framework is based on the non-Euclidean Riemann geometry, where the infinitesimal physical distance ds between two neighboring locations with respective 4-coordinates x and x' and separated by $dx = (dx^0, dx^1, dx^2, dx^3)$ follows

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu, \tag{1.1}$$

where $g_{\mu\nu}$ is the space-time *metric*. From $g_{\mu\nu}$ one can derive the acceleration of a test particle, which is given by

$$p^\mu = \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho} \frac{dx^\nu}{d\tau} \frac{d^2x^\rho}{d\tau}, \tag{1.2}$$

where τ labels the worldline (proper time) of the particle, $\{x^\mu(\tau)\}$ are the four-dimensional coordinates, and $\Gamma^\mu_{\nu\rho}$ are the Christoffel symbols that are given by

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\alpha} \left(\frac{\partial g_{\alpha\nu}}{\partial x^\mu} + \frac{\partial g_{\alpha\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\alpha} \right). \tag{1.3}$$

In curved space-time, the dynamics of a free particle (that is not subject to external non-gravitational forces) is given a null acceleration $p^\mu = 0$, giving the *geodesic* equation. Once the local geometry is no longer flat, i.e., the $g_{\mu\nu}$ terms are non-trivial functions of the 4-coordinates x, then $\Gamma^\lambda_{\mu\nu} \neq 0$, and the *linear* solution $\frac{d^2x^\mu}{d\tau^2} = 0$ (in Cartesian coordinates) is no longer valid. Then, the $g_{\mu\nu}$ terms act as a source for the non-linear dynamics of a test particle in a given system of coordinates. Einstein (1916) proposed that the evolution of the space-time geometry $g_{\mu\nu}$ is coupled to the local energy momentum and stress. This is encoded in a set of non-linear partial differential equations called the Einstein's field equations (EFE) that are given by

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.4}$$

where $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + g_{\mu\nu}\Lambda$. $R_{\mu\nu}$ is the Ricci tensor, derived from the Christoffel symbols $\Gamma^\lambda_{\mu\nu}$ thus relying on the partial derivatives of the metric. $T_{\mu\nu}$ is the energy-momentum tensor, that accounts for the various contributions to the total energy content, momentum and stress in the Universe, and Λ is a cosmological constant.

As a summary, Einstein's theory of gravity tells that in the absence of any external non-gravitational force, gravity appears as an intrinsic geometrical property of space and time, since any perturbation in the mass/energy content in $T_{\mu\nu}$ sources the evolution of the metric in $R_{\mu\nu}$, causing modification of the local space-time curvature and then affects particle dynamics. Reversely, a highly perturbed metric dictates the energy density how to behave. General relativity then refines the Newton's law of universal gravitation.

1.2 Homogeneous FLRW universe

For an arbitrary form of the metric $g_{\mu\nu}$ and energy-momentum tensor $T_{\mu\nu}$, EFE are difficult to solve. Applied to our Universe, we can build a simplified version of the metric based on

the geometrical properties of *homogeneity* and *isotropy*, which are both major statements (also hypothesis) of the cosmological principle, the milestone of most cosmological models. The cosmological principle tells that the matter distribution is roughly statistically homogeneous and isotropic when viewed at large enough scales and for any observer in the Universe, as motivated by observations. Using these assumptions, the metric reaches a very simple form called the Friedmann (1924); Lemaître (1931); Robertson (1935); Walker (1937) (FLRW) metric. The FLRW metric can be represented by re-writing the interval ds with respect to time t and spherical spatial coordinates (r, θ, ϕ) such that

$$ds^2 = -c^2dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right), \tag{1.5}$$

where $a(t)$ is the scale factor, and k is the *curvature* of the metric associated to spatial coordinates (r, θ, ϕ). Both scale factor and curvature are degrees of freedom that are not suppressed by prescriptions of the cosmological principle and tell that the geometry is possibly not *static* (is $da/dt \neq 0$?) and not flat (is $k \neq 0$?). Since r corresponds to comoving coordinates that are not "affected" by the scale factor, the physical distance between two objects which are respectively at $r = 0$ and at the radial (or line-of-sight) comoving distance $r = r_0$ is given by

$$D = a(t)\int_0^{r_0} \frac{dr'}{1-kr'} = a(t)\chi(r_0), \tag{1.6}$$

where χ is the transverse comoving distance and is given by (Hogg, 1999)

$$\chi(r_0) = \begin{cases} \frac{1}{\sqrt{|k|}}\sin\left(\sqrt{|k|}r_0\right) & \text{if } k > 0, \\ r_0 & \text{if } k = 0, \\ \frac{1}{\sqrt{|k|}}\sinh\left(\sqrt{|k|}r_0\right) & \text{if } k < 0. \end{cases} \tag{1.7}$$

Both possible time dependent *dilatation/contraction* of space through $a(t)$ and curvature k affect the physical distance between objects.

Cosmological redshift and distances

All of our observations of the Universe come to us through photons, that travel at the speed of light from every distant source to our telescopes. In a curved spacetime, the motion of a photon follows $ds^2 = 0$. Consider a point source at $r = r_1$, that emits light at $t = t_1$. From Eq. (1.5), the photon radial coordinate $r(t)$ follows

$$\frac{dr(t)}{dt} = -\frac{a(t)}{\sqrt{1+kr(t)^2}}. \tag{1.8}$$

Here, $dr/dt < 0$ denotes that the light rays move towards us. We can re-write the above equation as

$$\int_{t_1}^{t_0} \frac{cdt'}{a(t')} = \int_0^{r_1} \frac{dr}{\sqrt{1+kr^2}}. \tag{1.9}$$

The right side of the equation is constant with time and denotes the comoving transverse distance to the source. Considering that the source emits a second signal at time $t_1 + \delta t_1$, we can re-write the same equation. It will be received on earth ($r = 0$) at time $t_0 + \delta t_0$. Then we have

$$0 = \int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{cdt'}{a(t')} - \int_{t_1}^{t_0} \frac{cdt'}{a(t')} = \frac{c\delta t_1}{a(t_1)} - \frac{c\delta t_0}{a(t_0)}. \tag{1.10}$$

This is a fundamental result for observational cosmology applications. The time interval between the two signals δt_0 measured by an observer on earth is not equal to the time delay measured by the emitter. From the above, we have

$$\frac{\delta t_0}{\delta t_1} = \frac{a(t_1)}{a(t_0)}. \tag{1.11}$$

Since the speed of light $c = \lambda/\delta t$ is constant for any observer, where λ is the measured wavelength of a photon, we get

$$\frac{\lambda_0}{\lambda_1} = \frac{a(t_0)}{a(t_1)}. \tag{1.12}$$

Setting $a(t_0) = 1$ for our present time $t = t_0$, the observed wavelength of a photon emitted at time $t_1 < t_0$ is given by

$$\lambda_0 = \frac{\lambda_1}{a(t_1)}. \tag{1.13}$$

The *redshift* of a photon emitted, with wavelength λ_1 at time t_1 and observed with wavelength λ_0 at time t_0,is then defined as

$$z = \frac{\lambda_0 - \lambda_1}{\lambda_1}. \tag{1.14}$$

From the previous equations, the observed cosmological redshift of a photon is linked to the scale factor at the time of emission via

$$a(t_1) = \frac{1}{1+z}, \tag{1.15}$$

where $z = 0$ corresponds to present time[1]. All distances to a given source in the Universe can now be related to its cosmological redshift z (Hogg, 1999; Bartelmann & Schneider, 2001). The radial comoving distance is given by

$$r(z) = d_H \int_0^z \frac{dz'}{H(z')/H_0}, \tag{1.16}$$

where $H(z) = \dot{a}/a$ is the Hubble time-dependent parameter describing the time evolution of the scale factor $a(z)$ ($\dot{a}$ is the time derivative of a), $d_H = c/H_0$ is the Hubble distance where H_0 is the Hubble parameter at present time $H(z = 0)$. The radial comoving distance is the fundamental distance quantity since all other distance definitions rely on it. In particular, the angular diameter distance (that will be central to the weak lensing work presented in Chapter 2) is defined as the ratio of an object's physical transverse size to its angular size. It is given by

$$D_A(z) = \frac{\chi(r[z])}{(1+z)}. \tag{1.17}$$

where $\chi(r)$ is the transverse comoving distance.

Evolution of the scale factor

Solving the EFE requires making assumptions on the energy content in the Universe. In the homogeneous universe, each source of Einstein's gravity is treated as a fluid. Considering the simplest case of a perfect fluid, i.e. with no viscosity, the stress-energy tensor can be written in

[1]This shift in wavelength is usually referred to as the *cosmological* redshift, to be differentiated to other mechanisms, e.g., the Doppler effect due to the relative peculiar motion of sources, see Section 1.4.3.

terms of the 4-velocity field $u^\mu = (1, 0, 0, 0)$ of the fluid at rest, its energy density ρ and pressure p, such as (Schutz, 1985; Piattella, 2018)

$$T_{\mu\nu} = (p + \rho)u^\mu u^\nu - p g_{\mu\nu}. \tag{1.18}$$

Combining the EFE with the simplified form of the metric and the stress-energy tensor of a perfect fluid, we get the two Friedmann equations (Friedmann, 1924), namely

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2} \tag{1.19}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{c^2}. \tag{1.20}$$

From Eq. (1.19), the scale factor depends on the energy content denoted by ρ but is also impacted by the Λ-term and k-term. They both can be written in terms of their respective volumetric *densities* $\rho_\Lambda = \Lambda c^2/8\pi G$ and $\rho_k = -3kc^2/8\pi G a^2$. Defining $H_0 = H(t = t_0)$ to be the Hubble parameter at present time, with $H(t) = \dot{a}/a$ and the critical density[2]

$$\rho_c = \frac{3H^2}{8\pi G}, \tag{1.21}$$

the more convenient form of the first Friedmann equation in Eq. (1.19) is given by

$$H(t)^2 = H_0^2 \sum_i \Omega_i(t), \tag{1.22}$$

where the sum runs over all possible sources of the dynamics of $a(t)$, i.e., the ρ contribution in Eq. (1.19), plus ρ_Λ and ρ_k. To each of these contributions is associated a dimensionless *cosmological parameter*

$$\Omega_i(t) = \frac{\rho_i(t)}{\rho_c(t)}. \tag{1.23}$$

As a general property of the FLRW metric and Einstein's field equations, the curvature and cosmological constant terms are given in Eq. (1.22) respectively by $\Omega_k = -kc^2/(aH(t))^2$ and $\Omega_\Lambda = \Lambda c^2/3H(t)^2$ directly from Eq. (1.19). The time dependence of all the other different densities can be found using once again the EFE. From Eq. (1.4) the local conservation of the local energy-momentum tensor is given by $\nabla_\mu T^{\mu\nu} = 0$. Using the expression of $T_{\mu\nu}$ in Eq. (1.18) and the FLRW metric, we get the conservation equation

$$\dot{\rho} + 3H(\rho + p) = 0. \tag{1.24}$$

The Friedmann equation (Eq. (1.19)) and the conservation equation (Eq. (1.24)) can be used to describe the joint evolution of the scale factor a, the pressure p and density ρ. To remove one degree of freedom, it is common to use an equation of state for each fluid with index i where the pressure is linked to the density by

$$p_i = w_i \rho_i. \tag{1.25}$$

Eq. (1.24) then becomes

$$\sum_i \dot{\rho}_i + \rho_i 3H(1 + w_i) = 0. \tag{1.26}$$

[2]In the Big Bang model, the critical density is defined as the average density in the Universe that is just sufficient to halt the expansion at some time infinitely far in the future.

The dynamics of each component can be considered independently, if there is no energy transfer between them, so the above sum is null if each term is null too. This implies that the individual background density ρ_i evolves as

$$\rho_i \propto \exp\left(-3\int_{a_0}^{a}(1+w_i)d\ln a\right) \quad \rightarrow \quad \rho_i \propto a^{-3(1+w_i)}, \tag{1.27}$$

where we obtain the last equation for the specific case of constant equation of state i.e. $w(a) =$ cste. Similarly, it is common to compare the cosmological constant contribution to a fluid with an equation of state $w = -1$, and the curvature with $w = -2$, that can be directly deduced from Eq. (1.19). Using these results, we have the revised version of the Friedmann equation

$$H(t)^2 = H_0^2 \sum_i \Omega_{i,0} a^{-3(1+w_i)}, \tag{1.28}$$

where $\Omega_{i,0}$ are today's values of the cosmological parameters, that follow the closure relation

$$1 = \sum_i \Omega_{i,0}. \tag{1.29}$$

Solving completely Eq. (1.19), we get that the scale factor at any time $t > t_0$ is given by

$$a(t) = a(t_0)\exp\left(H_0 \int_{t_0}^{t} dt' \sqrt{\sum_i \Omega_{i,0}\ a(t')^{-3(1+w_i)}}\right). \tag{1.30}$$

This equation gives the time evolution of physical distances in the Universe in Eq. (1.5) that may be *contracting* or *expanding*, depending on the energy content (from both "particles", Λ and k contributions) through cosmic time.

1.3 Standard model of cosmology

In this section, we present the main features of the standard ΛCDM cosmological model, that is widely used to describe the background evolution of the Universe and the evolution of the cosmological perturbations, and is corroborated by a plethora of observations through the past decades. In the last section, we have seen how the FLRW formalism and the hypothesis of a perfect fluid allow us to understand the dynamics of the scale factor. We refer now to a cosmological model to be a combination of knowledge describing the Universe's expansion, its geometrical properties, its composition, and how it started and evolved through time.

1.3.1 The ΛCDM paradigm

The *standard* ΛCDM cosmological model (ΛCDM for Λ-Cold Dark Matter, also called LCDM) is supported by a plethora of observational signatures. It is usually called the "concordance" cosmological model in that sense (see Section 1.4). It is a parametrization of the Big Bang cosmological model, in which the universe expanded from a very dense and hot initial state about 13.8×10^9 yr ago (we detail the ΛCDM thermal history of the Universe in Section 1.3.4). According to the ΛCDM paradigm, the Universe is filled with relativistic species such as photons and neutrinos (respectively denoted by Ω_γ and Ω_ν, where the full "radiation" parameter writes $\Omega_r = \Omega_\gamma + \Omega_\nu$), non-relativistic species such as baryons and a hypothetical *dark matter* (Ω_b and Ω_c, where the matter parameter is $\Omega_m = \Omega_b + \Omega_c$), a *dark energy* contribution (denoted by Ω_Λ), and curvature (Ω_k, which is assumed to be zero in most of used versions of ΛCDM). The

Total energy density	Ω					
Contributions	Ω_m		Ω_r		Ω_Λ	Ω_k
Equation of state w	0		1/3		-1	-1/3
sub-species	Ω_b	Ω_c	Ω_γ	Ω_ν	$\times$	$\times$

Table 1.1: Cosmological parameters and equation of state for ΛCDM paradigm

cosmological parameters $\{\Omega_i\}$, their equation of state and sub-species are listed in Table 1.1. Whereas photons, neutrinos and baryons are *standard* particles, the nature of dark energy and dark matter in ΛCDM is still unknown. Dark energy and dark matter represent respectively 70% and 25% of the energy content in the Universe and explain a variety of astrophysical and cosmological observations where standard physics remains powerless.

- **Dark matter:** The first surprising key feature of ΛCDM is dark matter. In the 1930s F. Zwicky inferred the mass of the Coma galaxy cluster from the motion of its member galaxies, and then compared it to an estimate based on galaxy counts. In his article (Zwicky, 1937) he concluded that the dynamical mass of the cluster was 400 times larger, unveiling the presence of an unseen massive substance (what he called *dunkle Materie*, german for dark matter). Later, studying the galaxy rotation curves, Rubin et al. (1980) emitted the hypothesis of halos of non-luminous matter, since observed galaxies should have had around six times as much dark matter as visible mass (i.e., the baryon mass) to fit their observations. Recently, even more precise inference of the dark matter content in the Universe was achieved by studying the Cosmic Microwave Background (CMB) anisotropies (that will be presented in Section 1.3.4) revealed that $\Omega_m = 0.315$ (Planck Collaboration et al., 2020a), where the baryon fraction is very low ($\Omega_c/\Omega_b \sim 7$). Similar results were obtained browsing a broad range of wavelengths and tracers, that we will discuss later, and have corroborated the presence of non-baryon matter filling the Universe in-homogeneously.

- **Dark energy:** First measurements of today's expansion rate H_0 were made by Lemaître (1931) (his article was first published in french in 1927) and more accurately by Hubble (1929), who discovered that nearby galaxies appeared to be moving away from us at speeds proportional to their distances. Later, Riess et al. (1998); Perlmutter et al. (1999) used high-redshift Type1a supernovae to constrain the expansion rate H_0 and showed that the expansion is accelerating ($\ddot{a} > 0$) in opposition to the gravity that would tend to slow it. By considering a homogeneous dark energy field that acts as a fluid with negative pressure against gravity (represented by Λ), they have shown that the Universe accounts for a significant contribution of dark energy compared to other species, such as $\Omega_\Lambda = 0.68 \pm 0.10$ (Riess et al., 1998). Since then, this value was confirmed and even more strongly constrained for example by the *Planck* data.

ΛCDM provides a suitable parametrization of the background dynamics of the Universe, describing how the expansion changes over time. In Table 1.2 we present some of the cosmological parameters of ΛCDM we discussed in the earlier sections, obtained recently by the *Planck* mission (Planck Collaboration et al., 2020a). The first block corresponds to the background parameters necessary to derive the expansion rate at any cosmological time. We have that $\Omega_b \sim 15\%$ of Ω_m, and that $\Omega_c \sim 85\%$ of Ω_m. We also see that the curvature parameter is compatible with 0. The parameters that are listed in the second block correspond to the theory of cosmological perturbations.

Cosmological parameter	Best fit and 1σ error
Hubble constant H_0	67.36 ± 0.54
Dark energy Ω_Λ	0.6847 ± 0.0073
Matter Ω_m	0.3153 ± 0.0073
Baryons $\Omega_b h^2$	0.02237 ± 0.00015
Cold dark matter $\Omega_c h^2$	0.1200 ± 0.0012
Curvature Ω_k	-0.011 ± 0.013
Amplitude of matter fluctuations σ_8	0.8111 ± 0.0060
Spectral index n_s	0.9649 ± 0.0042
Normalization of the power spectrum $\ln(10^9 A_s)$	3.044 ± 0.014

Table 1.2: Cosmological parameters and uncertainties from Planck Collaboration et al. (2020a). The first block contains *background* parameters, where $h = H_0/100$. The second block gives the key parameters to describe the level of matter fluctuations, and will be discussed in the next sections.

1.3.2 Cosmological perturbations

From the very complex Einstein's field equations, the FLRW formalism is well adapted to understand with little effort the expansion history of an *idealized* homogeneous and isotropic Universe. However, where we live is full of inhomogeneities (for instance galaxies, clusters of galaxies, filaments, etc.). In an expanding Universe, structures continue to grow and merge under the influence of gravity, forming a vast inhomogeneous cosmic web. We can easily look at the traces of inhomogeneities in the cosmological fluids through the large-scale structures filling our observable Universe. In the standard cosmological perturbation theory, these fluctuations are usually considered to be very small, describing the Universe to be close to FLRW, with some local perturbations.

In a perturbative approach, the pressure and the velocity fields of the cosmological fluid are slightly perturbed around their background values and are given by

$$\rho(t) \to \rho(t) + \delta\rho(\vec{x}, t), \tag{1.31}$$

$$p(t) \to p(t) + \delta p(\vec{x}, t), \tag{1.32}$$

$$\vec{v}(t) = 0 \to \delta\vec{v}(\vec{x}, t). \tag{1.33}$$

In classical Newtonian perturbation theory, we can apply the classical hydrodynamics equations to the cosmological fluid, which are

$$\text{Continuity equation} : \frac{\partial \rho}{\partial t} = -\nabla(\rho\vec{v}), \tag{1.34}$$

$$\text{Euler equation} : \frac{\partial \vec{v}}{\partial t} + \vec{v} \times \nabla \cdot \vec{v} = -\frac{1}{\rho}\nabla P - \nabla\Phi, \tag{1.35}$$

$$\text{Poisson equation} : \nabla^2\Phi = 4\pi G\delta\rho. \tag{1.36}$$

We can linearize the above equations, by expanding them to first order in small perturbations $\delta\rho$, $\delta\vec{v}$ and δp. The zeroth order term gives the steady unperturbed solution, depending on the background $\bar{\rho}$ and $\bar{p}$. For the first perturbative order, we can combine the derivative of Eq. (1.34) with the divergence of Eq. (1.35), then using Eq. (1.36) to eliminate the gravitational potential Φ. We get the equation of motion of the density perturbation $\delta\rho$ such as

$$\left(\frac{\partial^2}{\partial t^2} - c_s^2\nabla^2\right)\delta\rho = 4\pi G\bar{\rho}\delta\rho, \tag{1.37}$$

where $c_s^2 = (\partial P/\partial\rho)$ is the square of the speed of sound of the fluid. Making the plane wave decomposition of the above equation, i.e. $\delta\rho \propto \exp i(wt - \vec{k}\cdot\vec{x})$ we find

$$w^2 = \frac{4\pi^2 c_s^2}{\lambda_J^2}\left(\frac{\lambda_J^2}{\lambda^2} - 1\right) \quad \text{where} \quad \lambda_J = c_s\sqrt{\frac{\pi}{G\bar{\rho}}} \tag{1.38}$$

is the *Jeans* length. Let us define two regimes. For perturbations with wavelength $\lambda \geq \lambda_J$, $w^2 < 0$; the perturbations are unstable under gravity and start increasing exponentially. The Jeans length tells when gravity dominates over acoustic wave propagation that travels at the speed of sound. When $\lambda \leq \lambda_J$, the matter perturbations will oscillate back and forth without collapsing.

So far, we have considered a static Universe, i.e., $a(t) = 1$. This formalism has some limitations when considering an expanding Universe. To account for expansion, we need to re-scale all the lengths by a factor of $a(t)$, then we modify the time and spatial derivatives in the equations of motion. By defining the local density $\rho(\vec{x},t) = \bar{\rho}(t)[1+\delta(\vec{x},t)]$ where $\delta(\vec{x},t)$ is the *overdensity* field, we find that δ follows

$$\left(\frac{\partial^2}{\partial t^2} + 2H\frac{\partial}{\partial t} - \frac{c_s^2}{a^2}\nabla^2\right)\delta = 4\pi G\bar{\rho}\delta. \tag{1.39}$$

Eq. (1.39) is equivalent to Eq. (1.37) but when considering an expanding Universe. The growth of linear perturbations in an expanding universe is then given by solving this differential equation. Considering the Fourier transform of the overdensity field

$$\delta(\vec{x},t) = \int d\vec{k}\delta_k(t)\exp\{-i\vec{k}\cdot\vec{x}\}, \tag{1.40}$$

we have

$$\ddot{\delta}_k + 2H\dot{\delta}_k = \frac{c_s^2}{a}(k_J^2 - k^2)\delta_k, \tag{1.41}$$

where $k_J = 2\pi a/\lambda_J$ is the Jeans wave-number. We see that the perturbations are now influenced by the *expansion* terms H and a. But including H and a also makes the search for a solution very difficult. As for the static approach, one can identify the unstable regime, i.e., when perturbations $\lambda \gg \lambda_J$ are much larger than the Jeans length, i.e., $k^2 \ll k_J^2$, (when gravitation dominates against pressure). Neglecting the pressure contribution, by considering the $c_s^2\nabla^2$ term to be negligible in Eq. (1.39), only the time derivatives remain. We can search for a general time-dependent solution. We find that the perturbation writes

$$\delta(\vec{x},t) = D_+(t)\epsilon_+(\vec{x}) + D_-(t)\epsilon_-(\vec{x}), \tag{1.42}$$

where $\epsilon_+(\vec{x})$ and $\epsilon_-(\vec{x})$ are initial perturbed fields that are constant with respect to time, and D_+ and D_- are decaying and growing modes. We find

$$D_- = H \quad \text{and} \quad D_+ = \frac{5}{2}\frac{H(a)}{H_0}\Omega_{m,0}\int_0^a \frac{da'}{[a'E(a')]^3}, \tag{1.43}$$

where $\Omega_{m,0}$ is the today's cosmological parameter for matter fluid and $E(a) = H(a)/H_0$. From these equations, it is simple to deduce that whilst physics in the early universe sets up the value of the density perturbations $\epsilon_{+/-}$ their subsequent growth through cosmic time is dependent on the expansion history of the universe. For example for a matter-dominated Universe, the scale factor evolves as $a(t) \propto t^{2/3}$ (from Eq. (1.22)), and the growing modes increase as $D_+(t) \propto t^{2/3}$. Reversely, in a universe dominated by dark energy, $a(t) \propto \exp(Ht)$, the growth of fluctuations with $D_+(t) \propto \exp(-Ht)$ vanishes rapidly.

The *classical* theory of cosmological perturbations is *valid* when the typical extent of density variations is much smaller than the Hubble length $1/H$ as well as for weak gravitational fields. For such scales, the relativistic effects of the metric fluctuation can be ignored and most of the essential physics can be extracted from a Newtonian approach. Moreover, the *Newtonian* description of the gravitational field Φ acting as an external "force" is still inherent to the classical theory of perturbations, whereas general relativity tells us that gravitation is due to metric perturbation.

For more accurate results, the theory of perturbations needs to be treated from Einstein's general relativity directly (Mukhanov et al., 1992). According to EFE in Eq. (1.4), any density perturbation now perturbs the geometry of space-time itself. We can now write the perturbed energy-momentum tensor $T_{\mu\nu}$ and metric tensor $g_{\mu\nu}$ as

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \quad \text{and} \quad T_{\mu\nu} = \bar{T}_{\mu\nu} + \delta T_{\mu\nu}, \tag{1.44}$$

where $\bar{g}_{\mu\nu}$ is the background space-time metric in the perturbed Universe. Both $T_{\mu\nu}$ and $G_{\mu\nu}$ in the EFE in Eq. (1.4) can be expanded the same way order-by-order. The first and second-order equations of motion,

$$\bar{G}_{\mu\nu} = \frac{8\pi G}{c^4}\bar{T}_{\mu\nu} \quad \text{and} \quad \delta G_{\mu\nu} = \frac{8\pi G}{c^4}\delta T_{\mu\nu}, \tag{1.45}$$

represent respectively the motion of the background and perturbed Universe. The first equation leads to the Friedmann equations in Eqs. (1.19) and (1.20). The second equation now describes how both metric and energy perturbations are coupled and evolve. In the Newtonian gauge (see Appendix A), the "perturbed" interval for a flat Universe (i.e. $k = 0$) is given by (Peter & Uzan, 2013)

$$ds^2 = a(\eta)^2[-(1+2\Phi)d\eta^2 + (1+2\Phi)d\vec{x}^2], \tag{1.46}$$

where η is the conformal time defined with respect to time t by $a(\eta)d\eta = dt$ and Φ is the scalar metric perturbation, i.e., the gravitational potential. The energy-momentum tensor perturbation depends on the pressure, density and velocity perturbations around their mean values (as mentioned in the "classical" perturbation theory), that now evolves jointly with the metric perturbation Φ. Taking a matter-dominated universe, and after a significant amount of calculations (see full derivation in Appendix A) we get

$$\delta = -2\Phi + \frac{2}{3\mathcal{H}}\nabla^2\Phi \quad \text{giving} \quad \delta_k = -2\Phi_k - \frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2\Phi_k, \tag{1.47}$$

where δ_k is the Fourier transform of δ. When $k \gg \mathcal{H}$ (sub−Hubble scales), we find that

$$\delta_k \approx -\frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2\Phi_k \propto \eta^2 \propto t^{2/3}. \tag{1.48}$$

So, "sub−Hubble" fluctuations grow by $\propto t^{2/3}$ in the matter-dominated era, and we find the same result as obtained applying the classical perturbation theory in an expanding Universe in Eq. (1.43). However, the "super-Hubble" scale corresponds to $k \ll \mathcal{H}$. Converting to physical scales $k_{\text{phys}} = k/a$, we have $k_{\text{phys}} \ll H$. The perturbation follows $\delta_k = -2\Phi_k$, so the perturbations are fixed or *frozen*, there is no growth. In the linearized GR formalism, we see that density perturbations begin to grow when they enter the horizon, which was not considered in the classical theory of perturbations. At lower scales, we find the same results obtained from classical perturbation theory in an expanding universe. Thus, for sub-horizon scales, the magnitude of perturbation at time t should be

$$\delta_k(t) = \begin{cases} (a/a_k)\delta_{k,\text{prim}} & \text{if } a > a_k, \\ \delta_{k,\text{prim}} & \text{if } a < a_k. \end{cases}, \tag{1.49}$$

where a_k is the scale factor at the time when the scale k enters the horizon, and $\delta_{k,\mathrm{prim}} = -2\Phi_k$ is the "primordial" density perturbation, that will be discussed in Section 1.3.4 (for this example we do not consider the Λ-dominated era occurring after the matter domination, that will tend to slow down the growth of perturbation at all scales). As the Universe is filled with multiple fluids, all the perturbed components will evolve on a common background, whereas the growth of all possible perturbations will be affected by the coupling between species, which one dominates the background expansion, etc. (We discuss this aspect in Appendix A).

1.3.3 Matter power spectrum

From the equations above that govern the evolution of cosmological perturbations, we can derive very relevant quantities for studying the growth of fluctuations and the emergence of large-scale structures that will be of interest in this thesis. Considering a density field with overdensity $\delta \equiv \delta(\vec{x}, t)$, we can investigate its clustering by measuring the 2-point correlation function

$$\xi_\delta(\vec{x}, \vec{x}', t) = \langle \delta(\vec{x}, t)\delta(\vec{x}', t)\rangle. \tag{1.50}$$

The brackets $\langle\cdot\rangle$ denote an averaging over many realizations of the Universe, so it is not *feasible* in reality. Since we have access to one realization, statistical cosmology relies on the ergodic theorem first proposed by Boltzmann (1896) that states that the ensemble average (over multiples versions of the Universe) and the spatial average over one realization of a random field at a fixed time are equivalent, if spatial correlations (the strength of ξ_δ) decay sufficiently rapidly over the survey volume. The cosmological principle (homogeneity and isotropy), the weak cosmological principle (field fluctuations are *small* over Hubble scales) and the ergodicity theorem represent the "Fair Sample Hypothesis" proposed by Peebles (1980), that ensure that the 2-point correlation function is a robust quantity to describe the matter clustering when estimated from data.

For a Gaussian random field δ, that will be considered in this section, the statistical properties are fully captured by the 2-point correlation function, and all odd-order correlation functions vanish. Moreover, thanks to the cosmological principle, the 2-point correlation function of δ only depends on time and $r = |\vec{x} - \vec{x}'|$. In Fourier space, we have

$$\delta_D^{(3)}(\vec{k} - \vec{k}')P_\delta(\vec{k}, t) = \langle \delta(\vec{k}, t)\delta^*(\vec{k}', t)\rangle \tag{1.51}$$

where $\delta_D^{(3)}$ is the three dimensional Dirac function and P_δ is the *power spectrum* of the field δ (also noted P_{ab} when considering two fields δ_a and δ_b in Eq. (1.50), we will use it later in Chapter 2). The auto-correlation function now verifies

$$\xi_\delta(r, t) = \int_0^{+\infty} \frac{dk}{2\pi^2} k^2 P_\delta(k, t) \frac{\sin(kr)}{kr}. \tag{1.52}$$

At this point, as we mentioned before, the power spectrum of a species represented by the random field δ is sourced by the primordial power spectrum that we will note P_0, that is assumed to be generated during the inflation era (see Section 1.3.4). For the inflaton scalar field, as well as motivated by the analysis of CMB anisotropies by *Planck*, the primordial power spectrum $P_0(k)k^3$ is almost scale-invariant, thus follows a power-law form

$$P_0(k)k^3 \propto A_s \left(\frac{k}{k_s}\right)^{n_s - 1}, \tag{1.53}$$

where n_s is the spectral index (close to 1), A_s is the normalization and k_s is the pivot wavenumber. Any power spectrum of a random field δ is given in terms of the primordial power spectrum via

$$P_\delta(k, t) = P_0(k)D_+(t)^2 T_\delta(k, t)^2, \tag{1.54}$$

where $T_\delta(k,t)$ is the transfer function. Since $P_0(k)$ relies on inflation physics, $T_\delta(k)$ tells how perturbations with different wave-numbers evolved differently from the very early Universe independently of initial conditions, since they *cross* the Hubble radius at different times as seen in Section 1.3.2. $T_\delta(k)$ depends also on the nature of δ (baryons, cold dark matter, radiation, etc.). Deriving the shape of $T_\delta(k)$ for each species reacquires to solve the Boltzmann equation in a perturbed FLRW metric, tracking the evolution of perturbation from the re-entering of k-modes in the Hubble radius with codes such as CLASS (Lesgourgues, 2011) or CAMB (Lewis et al., 2000). The D_+ term is the linear growth factor as derived in Section 1.3.2, that is affected by the background evolution of the Universe. Let's consider a spherical region of radius R. The "filtered" matter overdensity is given by

$$\delta_R(t) = \int d^3\vec{x}\delta_m(\vec{x}')W_R(\vec{x}), \tag{1.55}$$

where $W_R(\vec{x})$ is the top-hat function of radius R (normalized to 1). The variance of the filtered perturbation is

$$\sigma^2(R,z) = \langle|\delta_R|^2\rangle = \int \frac{d^3\vec{k}}{(2\pi)^3}|\overline{W}_R(k)|^2 P_{mm}(k,z), \tag{1.56}$$

where $\overline{W}_R$ is the Fourier transform of W_R and P_{mm} is the matter power spectrum. Following the above equation, $\sigma(r,z)$ tends to decrease with increasing filtering scale R. In the above expression, the power spectrum tells the contribution of each k-mode to the total variance within the sphere. From that, we define today's amplitude of fluctuations

$$\sigma_8 = \sqrt{\sigma^2(R = 8\ \mathrm{Mpc}.h^{-1}, z = 0)}. \tag{1.57}$$

The radius $R = 8$ Mpc.h^{-1} roughly corresponds to the scale of massive halos of matter in the Universe. The quantity σ_8 has a great influence over the growth of fluctuations through cosmic time and is often used as a reference for the normalization of the power spectrum instead of A_s. For instance, a larger value of σ_8 implies that fluctuations are larger, and therefore induces earlier structure formation. The parameter σ_8 can therefore be constrained from probes of structure formation such as galaxy clusters that will be described in Chapter 2.

1.3.4 Cosmic evolution and thermal history of the Universe

The different species in the Universe are represented by their average densities. We represent in Fig. 1.1 the joint evolution of the energy densities, namely the radiation, matter and dark energy density with respect to scale factor (left) and the variation of the scale factor with respect to time (right). In ΛCDM cosmology, radiation dominates the energy content in the Universe roughly before $z \sim 10^4$ giving that $\rho_{\rm tot} \propto a^{-4} \rightarrow a \propto t^{1/2}$ by solving the Friedmann equation (represented in blue in the two figures). Between $z \sim 10^4$ and $z \approx 0.3$ the matter dominates leaving $\rho_{\rm tot} \propto a^{-3} \rightarrow a \propto t^{2/3}$ (in orange) and dark energy (in green) from $z = 0.3$ to present (and future) days leaving an exponential expansion $a \propto \exp(\sqrt{\Lambda/3}t)$.

Moreover, within their average values, respective abundances of sub-species are modified through several mechanisms, by decaying or coupling to other species. These evolutions can be handled using relativistic quantum field theory. One key ingredient to understand particle interactions is to compare the rate Γ of a reaction, for instance: $\psi_i + \psi_j \rightleftharpoons \chi_k + \chi_l$, that depends basically on the cross-section of particles (velocity, fluid temperature, etc.) to the characteristic expansion rate of the Universe H. When the reaction's rate $\Gamma \gg H$, the reaction is "fast" enough to maintain the thermodynamic equilibrium compared to the expansion rate, and the energy partition of particles is given by Bose-Einstein distribution for bosons, and Fermi-Dirac for leptons. Since the expansion increases the physical distance between particles, it tends to

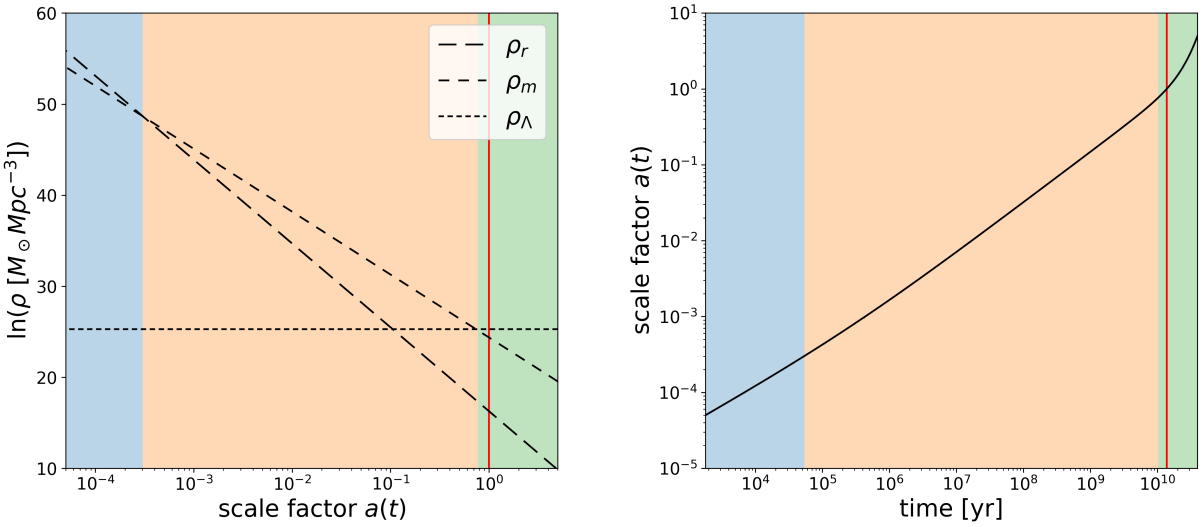

Figure 1.1: Joint evolution of energy densities and scale factor. Left: Evolution of the radiation, matter and dark energy densities with respect to the scale factor $a(t)$ (time goes from left to right). The radiation-dominated era is represented in blue, the matter-dominated era is represented in orange, and the dark energy-dominated era is represented in green. The red vertical line denotes the present time, i.e., $a(t_0 \approx 13.8\ \mathrm{Gyr}) = 1$. Right: Corresponding evolution of the scale factor with respect to cosmic time since the Big Bang.

slow down reactions. Solving the Boltzmann equation in an expanding Universe, the equation of evolution of the physical number density n_i of particle i is driven first by the interaction rate, but also by the background expansion where the collisions set in. It gives

$$\frac{d\ln(n_i a^3)}{d\ln a} = -\frac{\Gamma}{H}\left(1 - \left(\frac{n_i n_j}{n_k n_l}\right)_{\mathrm{eq}} \frac{n_k n_l}{n_i n_l}\right) \tag{1.58}$$

where terms in parenthesis are equilibrium values. We see that the thermodynamic equilibrium is maintained since reactions are fast enough compared to the expansion, i.e., when the term Γ/H is large enough. When $\Gamma \sim H$, the reaction transits out of equilibrium and the species decouples from the thermal bath due to the spatial dilution, the reaction *freezes* and stops. The species behaves *on its own* as if nothing was there, giving $n_i \propto a^{-3}$ (pure dilution due to expansion, obtained by mass conservation). The comoving number density of ψ_i remains constant with time. As the Universe follows adiabatic expansion, its temperature decreases and it can be shown that e.g. $T \propto a^{-1}$ for the radiation-dominated era, and $T \propto a^{-2}$ for the matter-dominated era. This modifies the interaction rate Γ, the ratio Γ/H will evolve to lower values for the different reactions that occur in the cosmological fluid. The freeze-outs of the different interactions will occur at different cosmological times.

In the following, we make a summary of the main epochs in the Universe, from the production of elementary particles to the formation of large scale structures. We will refer to the freeze-out epochs for each species that fills the cosmological fluid. We quote the typical redshift of each epoch and the background temperature of the thermal bath. In particular, since we focus on clusters of galaxies in this thesis, we detail a little bit more the inflation era, a very short period in the early Universe during which the primordial fluctuations of the cosmological fluid were created, sourcing the structure formation in the late Universe.

- **Pre-inflationary period** ($t < 10^{-36}$ s, $z \sim ?$, $T > 10^{16}$ GeV): This epoch is not well

understood, since our current physical theories have no relevant predictive power during this period. This is the place where physics in the very early Universe is assumed to be dominated by the quantum effects of Einstein's gravity. The three forces of the standard model of particle physics are assumed to be unified in a Grand Unified Theory (Georgi & Glashow, 1974), except for gravity.

- **Inflationary epoch and reheating**[3] ($t \sim 10^{-36} - 10^{-32}$ s, $z \sim ?$, $T \sim 10^{15} - 10^{9}$ GeV) According to inflation paradigm (Albrecht & Steinhardt, 1982; Guth, 1981; Linde, 1982), the Universe had a very rapid (namely exponential) phase of expansion prior to the Big Bang, during which the energy content of the Universe was dominated by cosmological constant-type of vacuum energy, much bigger than any other energy form, for instance described in the simplest single-field inflationary model by the *inflaton*. From the Euler-Lagrange equation, the *inflaton* field Φ follows

$$\ddot{\phi} + 3H\phi - V(\phi) = 0, \tag{1.59}$$

where the Hubble friction decelerates the *inflaton* motion through the minimum of its potential $V(\phi)$. Its energy-momentum tensor gives its energy density and pressure, and its equation of state is given by

$$w = \frac{p_\phi}{\rho_\phi} = \frac{\dot{\phi}^2 - 2V(\phi)}{\dot{\phi}^2 + 2V(\phi)}. \tag{1.60}$$

When $V(\phi) \gg \dot{\phi}^2$ (slow-roll approximation) then $w \approx -1$; Since the *inflaton* energy density dominates over other possible species during the inflation epoch, the equation of state of the effective cosmological fluid remains $w = -1$ and the *inflaton* acts as a cosmological constant, making the Universe expanding by a factor of the order of 10^{26} during a very short time. When the *inflaton* field reaches the bottom of its potential and slows down, then $V(\phi) = 0$ and $w = 1$, the inflation stops. After that, the *inflaton* will oscillate at the bottom of its potential and its energy will be converted into quark-gluon plasma and leptons. This period is known as *reheating*. Moreover, before inflation, the typical scales in the Universe satisfy $\lambda \ll 1/\mathcal{H}$ so that regions were in causal contact and then could have attained a uniform temperature. Inflation tells us that these regions must have been in contact with each other in the past. Then, comoving scales become super-horizon and re-enter the horizon after reheating. The pre-inflation causal contact solves the problem of the *Horizon* so that the Universe *thermalized* in its early life and explains why the CMB temperature is homogeneous on very large scales. Inflation also provides an answer to the *Flatness problem*, by stretching any initial spatial curvature to near flatness due to the very rapid expansion of the Universe, what we observe today. Moreover, the theory of inflation is typically invoked as a solution to the generation of primordial fluctuations as the seed for structure formation. In the very early universe, before its very rapid expansion phase, quantum fluctuations of the *inflaton* fill the microscopic field. These fluctuations are created on all length scales, with associated power spectrum (this point will be addressed in Section 1.3.3). During inflation epoch, $H \approx$ cst, and the scale factor

$$a(t) \propto exp(\sqrt{H}t). \tag{1.61}$$

The typical size of the Hubble horizon $1/\mathcal{H}$ then decreases very fast and most of comoving scales become super-horizon, i.e., $\lambda \gg 1/\mathcal{H}$ and the quantum fluctuations of *inflaton* are frozen, such that the causal physics cannot act on super-horizon perturbations as seen in

[3] The following is based on pedagogical review of inflation from Baumann (2009).

Section 1.3.2. At the end of inflation, the modes re-enter the Hubble radius and evolve again, now imprinted in the clustering of standard matter and radiation (detailed right after). These small inhomogeneities then act as a source for the initial in-homogeneous gravitational potential, being magnified to cosmic size due to the expansion of the Universe, and shaping the initial matter and radiation energy distribution. The end of inflation and reheating marks the beginning of the *standard* Big Bang cosmology, and the Universe is super-cooled during this period. Moreover, the strong interaction becomes distinct from the electroweak interaction (GUT transition: Glashow (1959); Weinberg (1967); Salam (1968)) and quarks, leptons and their anti-particles are created. Particles interact by exchanging gluons (between quarks), electroweak bosons W_1, W_2, W_3, and B and via the Higgs interaction through the exchange of H^+, H^-, H^0 and H^{0*} bosons.

- **Baryogenesis** ($t \sim$?): According to relativistic quantum physics, each particle (denoted by s^-) is created with a "anti"-partner (s^+), and both annihilate in high energy photons through $s^+ + s^- \rightleftharpoons \gamma + \gamma$. Contrary to relativistic quantum physics that predicts a perfect balance between particles and anti-particles, baryogenesis (see the pedagogical review of Cohen et al. (1998)) assumes that there exists an unknown process that causes the baryonic asymmetry, i.e., that there are slightly more particles than anti-particles formed. Then, matter and anti-matter particle content will quickly annihilate, leaving an excess of matter in the Universe. The freeze-out of the "anti-matter annihilation" reaction will depend on the energy of the particles.

- **Electro-Weak transition** ($t \sim 10^{-12}$ s): The mutual interactions between particles lead to thermal equilibrium. Moreover, the electromagnetic and weak interactions are still coupled, and particles are massless. Before the electroweak symmetry breaks, the particle interactions are ruled by the electroweak and Higgs interactions. Via the Higgs mechanism, the electromagnetic interaction carried by the photons and weak interactions carried by the massless bosons W^+, W^-, and Z_0 become distinct. Also, every elementary particle in the cosmological fluid that interacts with the Higgs field becomes massive, except for the photons. The weak and strong forces become short-distance interactions compared to electromagnetism. The fundamental forces have taken their current form. The temperature is still too high for quarks to combine into composite particles.

- **QCD transition** ($t \sim 10^{-5}$ s, $z \sim 10^{12}$, $T \sim 150$ MeV): The Quantum ChromoDynamics (QCD) describes the strong interaction between quarks and gluons. The study of the QCD transition dates back to the 80's. Iso et al. (1986) proposed that as the Universe continues to cool down, the QCD strong interaction between quarks and gluons becomes predominant, and quarks coalesce in composite particles called hadrons such as three quark systems called baryons (building blocks of atomic nuclei) and quark anti-quark systems called mesons. The slight matter-antimatter asymmetry from the baryogenesis led to the elimination of anti-baryons, leaving a small excess of hadrons and leptons/anti-leptons dominating the matter content in the Universe.

- **Neutrino decoupling** ($t \sim 1$ s, $z \sim 6 \times 10^9$, $T \sim 1$ Mev): Neutrinos are coupled to the other species via $e^+ + e^- \rightleftharpoons \nu_e + \bar{\nu}_e$. As the Universe expands very fast and the temperature falls, the neutrino interaction rate becomes smaller than the expansion rate of the Universe. Then, at $t \sim 1$ s, the neutrinos decouple from the thermal bath and form the Cosmic Neutrino Background (CνB).

- **Electron-positron annihilation** ($t \sim 10$ s, $z \sim 2 \times 10^9$, $T \sim 0.5$ MeV): Electrons and positrons annihilate into photons. The temperature is too low to create new electron-positron pairs (reverse reaction of annihilation), leaving a small excess of electrons.

- **Big Bang Nucleosynthesis (BBN)** ($t \sim 4$ min, $z \sim 4 \times 10^8$, $T \sim 100$ keV): The first nuclei were formed. In a very short time, protons and neutrons collide to produce deuterium. Heavier elements such as helium, tritium and lithium are also produced.

- **Matter dominated era** ($t > 50$ kyr, $z \sim 3500$, $T \sim 0.75$ eV): From inflation to $t \sim 50$ kyr, the background expansion is dominated by radiation (see Fig. 1.1, blue region). At $t \sim 50$ kyr, matter starts dominating the expansion of the Universe (orange region in Fig. 1.1).

- **Recombination** ($t \sim 260 - 380$ kyr, $z \sim 1400 - 1100$, $T \sim 0.3$ eV) Through the reaction given by $e^- + p \rightleftharpoons H + \gamma$, protons and electrons first became bounded electrically in neutral hydrogen atoms H, whereas the reverse reaction $H + \gamma \rightleftharpoons e^- + p$ becomes energetically disfavored when the temperature becomes low enough. The photon fluid is still highly coupled to the plasma through Thomson scattering via the reaction $e^- + \gamma \rightleftharpoons e^- + \gamma$. At $t \sim 380$ kyr ($z \sim 1090$, $T \sim 0.25$ eV), the Thomson scattering becomes inefficient and the photons decouple from the hot electron fluid. The photon mean free path becomes comparable to cosmological scales and they start propagating freely. These photons can be observed today as they constitute the Cosmic Microwave Background (CMB, see Fig. 1.2).

- **Dark ages** ($t \sim 380$ kyr - 1 Myr): At this point, the Universe is almost transparent to electromagnetic radiations after the recombination and decoupling epochs. At the same time, massive hydrogen clouds start collapsing under gravity, but not enough to form stars and galaxies. There is no local source of "light", apart from the CMB and the diffuse 21 cm radio emissions from atoms of hydrogen. Dark and baryonic (neutral) matter start collapsing into huge and diffuse filaments and haloes, and early large structures gradually emerge in the Universe.

- **Reionization** ($t \sim 1 - 4$ Myr, $z \sim 10 - 30$, $T \sim 2 - 8$ meV): Before reionization, the Universe was "transparent". The reionization epoch starts with the arrival of the first stars and galaxies, formed by the gravitational collapse of hydrogen clouds. The ultraviolet radiation from the first stars then transformed the intergalactic medium from neutral to ionized. After reionization, galaxy clusters start forming within the most massive dark matter halos.

- **Dark energy dominated era** ($t > 9$ Gyr, $z \sim 0.4$, $T \sim 0.33$ meV): Galaxy clusters and super-clusters continue to form, progressively by accretion of smaller massive structures. The energy density passed an inflection point when dark energy starts being dominant in the Universe (blue region in Fig. 1.1), responsible for the recent accelerated expansion (whereas for the radiation and matter-dominated era the Universe was decelerating according to the Second Friedmann equation in Eq. (1.20)) until today ($t \approx 13.8$ Gyr, $z = 0$, $T \sim 0.24$ meV).

- **Future** ($t > 13.8$ Myr): Possible scenarios of the future evolution of the Universe are various. For instance, according to ΛCDM, the space-time metric will keep expanding indefinitely under the influence of dark energy, and the Universe will get colder and more dilute until large-scale structures decompose in elementary particles. In the opposite, the "Big Crunch" scenario proposes that the expansion would at some point be reversed and that the Universe would start contracting towards a very hot and dense state.

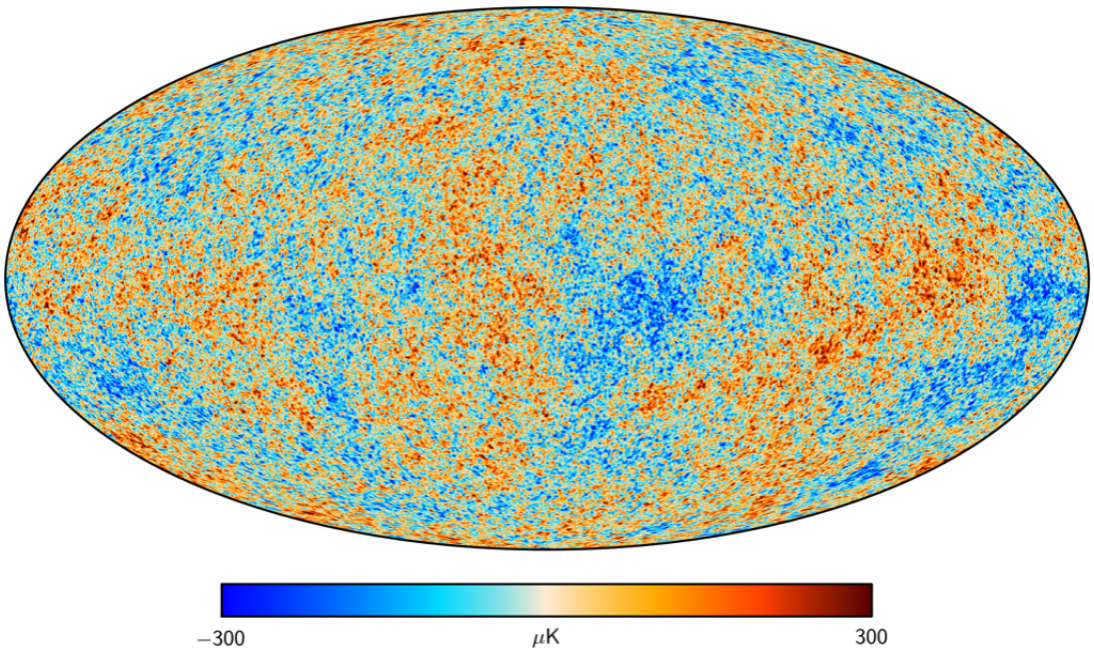

Figure 1.2: Map of CMB temperature anisotropies. Cold spots (in blue) represent overdensities, and hot spots (in red) correspond to underdensities measured by the *Planck* satellite (Planck Collaboration et al., 2020b).

1.4 Observational cosmology

After this detailed description of the theory supporting the cosmological model, we now turn to a short overview of the observational framework of cosmology. We primarily detect photons at different wavelengths coming from a variety of distant sources in a broad range of redshifts. From that, we need to identify "probes" that can be used to inform us about the content and the evolution of the Universe using ground-based or space telescopes. In this section, we briefly review the main cosmological probes, namely the mm light of the Cosmic Microwave Background in Section 1.4.1, the Type1a supernovae emission in Section 1.4.2, and the different tracers of the large scale structures in the Universe in Section 1.4.3. Galaxy clusters, which are the focus of this thesis, are one aspect of the large-scale structures and will be described in greater detail in the next chapter.

1.4.1 Cosmic Microwave Background

The CMB is a remnant radiation of the primordial Universe, left over after the recombination epoch at $z \approx 1100$, i.e., 380,000 years after the Big Bang. Photons stopped scattering off the hot electrons in the primordial baryon/photon plasma and started traveling freely and constitute now the Cosmic Microwave Background (photons were red-shifted when traveling from the last scattering surface to Earth, and today's energy distribution peaks in the microwave region of the spectrum). The spectrum of the CMB emission is close to a perfect thermal black-body spectrum, with a mean temperature of ∼2.725 Kelvin in all directions in the sky.

However, its temperature shows anisotropies at the $\sim 10^{-5}$ level (see Fig. 1.2) that are closely related to the density and metric perturbations in the very early Universe. These temperature fluctuations correspond to regions with slightly different densities and are the signature of the seeds for structure formation in the late Universe we observe today. Due to the anisotropic

scattering of photons at the recombination, the CMB is also slightly polarized at the 10^{-6} level, i.e., the orientation of light perpendicular to the direction of propagation of CMB photons is not completely random. The CMB polarization can be decomposed into both E and B-modes, which are both strongly sensitive in very different ways to gravity, especially to primordial scalar (sourcing E-modes) and tensor (sourcing B-modes) perturbations in the early Universe.

Most of the CMB temperature and polarization anisotropies are imprinted at the time when CMB photons were first emitted at recombination; these are called primary anisotropies. In particular, early B-modes would have been created during the inflation era by primordial gravitational waves (see the review of Kamionkowski & Kovetz (2016)). From the last scattering surface to our detectors on Earth secondary anisotropies appeared in CMB spectra (see Aghanim et al. (2008) for a review), since CMB photons had many interactions with their ionized environments (e.g. via Thomson scattering with free electrons during the re-ionization epoch) and emerging large scale structures in the late Universe (through light deflection by the foreground matter distribution, inducing secondary CMB B-modes, or via the Sunyaev-Zel'dovich effect with hot electron gas in galaxy clusters). Therefore, the CMB is a unique window to study the physics in the primordial plasma, as well as the thermal history and the mass distribution in the Universe. We refer the reader to Challinor (2012) for a review.

1.4.2 Type 1a supernovae

In some astrophysical scenarios, a white dwarf evolves into a binary system and gradually accretes the mass content of its companion. When the dwarf's mass exceeds the Chandrasekhar mass limit of $\approx 1.44\ M_\odot$, it triggers an explosion called a Type 1a supernova. Then, it is commonly assumed that Type 1a supernovae happen at the same star mass (the Chandrasekhar mass), and shine at a similar level of brightness, making them *standardizable* candles. Their characteristic luminosity is affected by the dilution of the energy density upon a sphere with increasing radius from the supernovae, as well as the redshift of photons when traveling through space, and this effect only depends on the redshift of the supernovae. The apparent magnitude m of a supernovae at redshift z that is observed on Earth is given by

$$m = M + 5\log_{10}\left(\frac{D_L(z)}{10\ \mathrm{pc}}\right), \tag{1.62}$$

where M is the absolute magnitude of the Type 1a supernova, and $D_L(z) = (1+z)^2 D_A(z)$ is the luminosity distance, that depends on cosmology through the scale factor. In practice, some corrections to m need to be accounted for, due to the extinction of the supernovae luminosity by dust lying in the host galaxy, but also intrinsic light curve shape and color variations. Therefore, the apparent luminosity of Type 1a supernovae is the most direct measurement of cosmological distances in the Universe, and thus its expansion and is usually called a *geometrical* probe. They are particularly used to probe the effect of dark energy and expansion rate in the late Universe. We refer the reader to Howell (2011) for a review.

1.4.3 Tracers of the Large-Scale Structures

The matter density field filling the Universe is not directly observable. However, its properties can be inferred from specific tracers. Here, we review the main tracers of the matter density field, namely the galaxy clustering, the cosmic shear and galaxy clusters. The latter are the focus of this thesis and will be more thoroughly described in Chapter 2.

Galaxy clustering: The information contained within the two-point correlation of galaxies enables us to measure the full shape of the galaxy power spectrum, the baryon acoustic oscillation

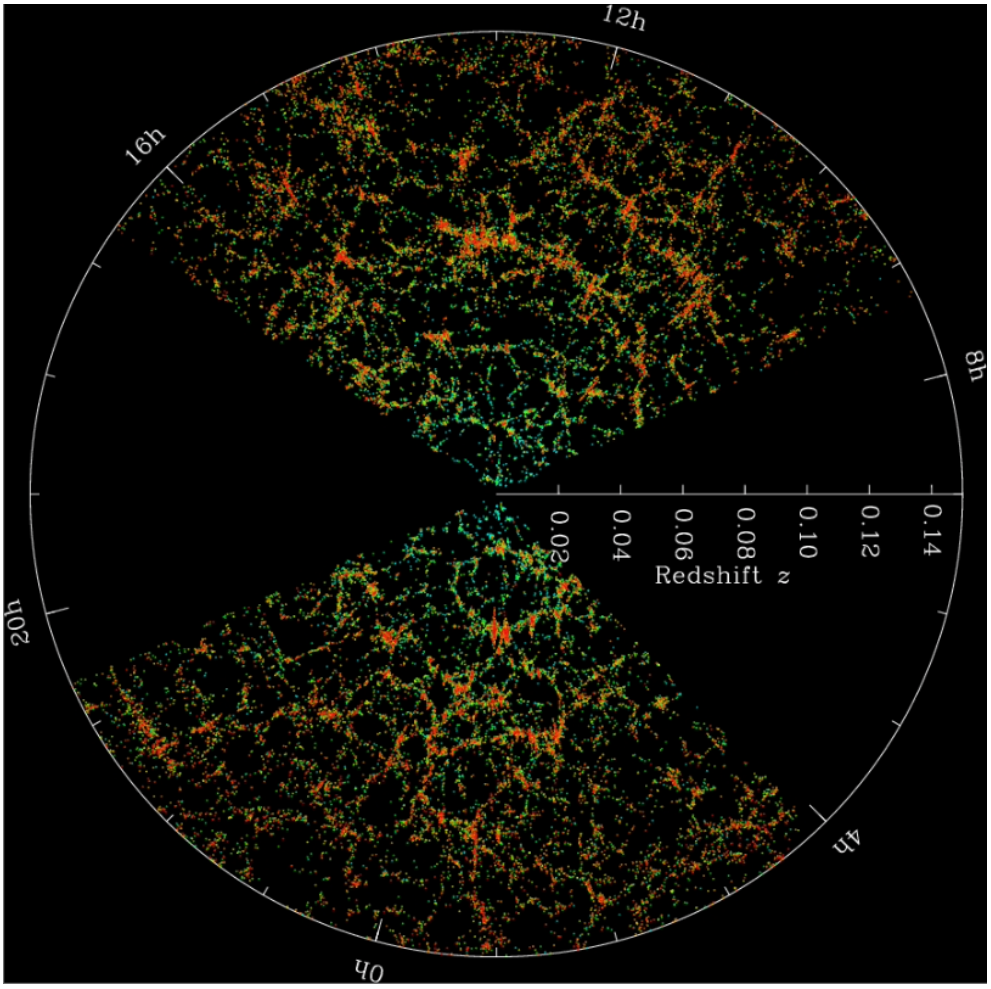

Figure 1.3: Redshift/angular distribution of SDSS galaxies. Each point represents a galaxy (redder: made of older stars). Image taken from https://classic.sdss.org.

and the effect of redshift-space distortion that we detail below.

- *Galaxy power spectrum*: Galaxies in the Universe are not randomly distributed, and follow the clustering of the matter distribution (for instance, see the SDSS galaxy sample in Fig. 1.3). In Fourier space, the observed power spectrum of the galaxy distribution can be related to the matter overdensity field in the Universe via the general form (Kurki-Suonio, 2021; Desjacques et al., 2018)

$$P_g(k,z) = b_g^2(k,z) P_{\rm mm}(k,z), \tag{1.63}$$

  where $b_g(k,z)$ is a scale-dependent bias that connects the galaxy overdensity for the mode k to that of matter (i.e., $\delta_g(k,z) = b_g(k,z)\delta_{\rm m}(k,z)$) and encodes all physical processes that make galaxies imperfect tracers of matter. Equivalently, the galaxy correlation function can be complemented by higher-order correlation functions that retain a lot of complementary information on the non-linear gravitational collapse occurring from the dark ages to today, and that imprints non-Gaussianities in the matter density field.

- *Baryon acoustic oscillation*: In the primordial Universe just before recombination, the competition between gravity (attracting matter toward over-dense regions) and the out-

ward pressure generated by the energy released by photon-matter interactions created oscillations in the matter/baryon density fields. Over-dense shells of baryons and photons propagated like acoustic waves in the primordial plasma. At recombination the baryons travel freely, and the typical size of these acoustic sound waves *froze* and imprinted the baryon spatial distribution (Bassett & Hlozek, 2010; Alcock & Paczynski, 1979), now exhibiting a typical bump called the BAO peak in the galaxy two-point correlation function at comoving separation r_s

$$r_s = \int_0^{\eta_\star} d\eta c_s(\eta) \approx \ 150 \text{ Mpc}, \tag{1.64}$$

as the largest comoving distance traveled by sound waves at speed $c_s(z)$. This bump corresponds to the length of the sound horizon and can be observed in the correlation functions of all baryon tracers of the matter density field, which then started to fall by gravity in these over-densities right after recombination. The BAO peak was first observed by SDSS (Eisenstein et al., 2005) in the large-scale correlation function of 47,000 Luminous Red Galaxies. It can be measured in both transverse $s_\perp$ and radial $s_{||}$ comoving separation respectively from its angular and redshift apertures $\Delta\theta$ and Δz such as

$$s_\perp = D_A(z)\Delta\theta(1+z) \quad \text{and} \quad s_{||} = \Delta z H(z)^{-1}. \tag{1.65}$$

Comparing these to quantities to the 3-dimensional acoustic length r_s is a great playground to test the consistency of cosmological models (the so-called Alcock-Paczynski test). It is often used as a standard ruler, making it a unique way to probe the effect of expansion on distance measurements, especially dark energy in the late Universe.

- *Redshift-space distortion*: The radial distance in cosmology is encoded in the *redshift*. In ΛCDM, the redshift of a galaxy z is given in terms of the Hubble recession velocity giving the cosmological redshift $z_{\rm H}$, and its peculiar motion $\vec{v}$ relative to the Hubble flow (Hamilton, 1998). Its reconstructed redshift-distance is given by $s = r(z_{\rm H}) + \vec{r}(z_{\rm H}) \cdot (\vec{v}/c)$ where $r(z_{\rm H})$ is the comoving radial distance at *Hubble* redshift $z_{\rm H}$. The peculiar velocity of a galaxy is sourced by the large-scale in-homogeneities in the mass distribution via the continuity equation

$$\nabla \cdot \vec{v} = -aH(a)f(a)\delta_{\rm mm}(\vec{x}, a), \quad \text{where} \quad f(a) = \frac{d\ln D_+}{d\ln a} \tag{1.66}$$

is the linear growth rate and H is the Hubble parameter. For instance, the small motions of matter within a galaxy cluster will cause the galaxies at the same cosmological distance to have slightly different redshifts. The peculiar velocities of galaxies may cause the galaxy to appear displaced along the line-of-sight, namely closer to us when the galaxy's peculiar motion is directed toward us ($\vec{r} \cdot \vec{v} < 0$), and reversely ($\vec{r} \cdot \vec{v} > 0$). The observed galaxy power spectrum is no longer isotropic and depends on the transverse component of the wave-number k, as well as on the typical speed at which structures grow. Compared to the "static" galaxy power spectrum in Eq. (1.63), the observed galaxy power spectrum follows

$$P_g(\vec{k}, z) = b_g^2(1 + f\mu^2/b_g)^2 P_{\rm mm}(\vec{k}, z) \quad \text{where} \quad \mu = k_{||}/|\vec{k}|. \tag{1.67}$$

The anisotropic correction $(1 + f\mu^2/b_g)^2$ is called redshift-space distortions and affects the pattern of clustering of galaxies in the redshift space, and can be used to infer the growth of structure from spectroscopic galaxy redshift surveys.

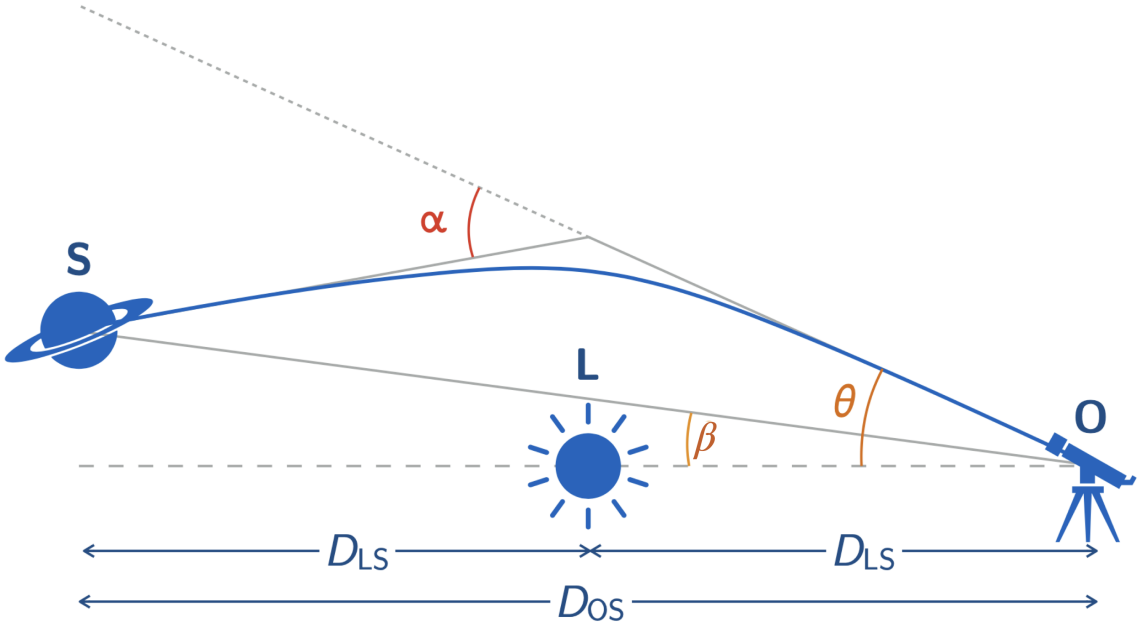

Figure 1.4: Deflection of light-rays coming from a distant source **S** by a single lens **L**, reaching the observer in **O**. Adapted from Doux et al. (2017).

Weak gravitational lensing: In the presence of matter fluctuations along its path, the trajectory of a photon will be deflected. Then, the observed position of the source inferred by tracing back photons considering a straight line will differ from its true position. Fig. 1.4 illustrates the simplest case, where the gravitational potential of a single lens **L** deflects the light-ray emitted by a source **S** at angular position $\vec{\beta}$ and reaching the observer **O**. The observed position of the source is given by

$$\vec{\beta} = \vec{\theta} - \frac{D_{\rm LS}}{D_{\rm OS}}\vec{\alpha}, \tag{1.68}$$

where $D_{\rm LS}$ and $D_{\rm OS}$ are respectively the distance between the lens and the source, and between the observed and the source. In other words, the *observed* angular position on the sky $\vec{\beta}(\vec{\theta}, z)$ of a source located at redshift z (equivalently at a comoving radial distance $r(z)$) will be different from its true angular position $\vec{\theta}$ that would have been observed in a homogeneous Universe (if there was no massive lens between the source and the observer). The light coming from a distant source at redshift z is deflected by the potential of the various correlated large-scale structures between the emitter and the observer. To relate $\vec{\beta}$ and $\vec{\theta}$ we can solve the geodesic equation for the photon from z up to $z = 0$. Making the Born approximation that is valid when the deflections are small, we can compute each contribution to the local deflection of the photon along its *unperturbed* path given by $\chi(r')\vec{\theta}$ (where $r' \in [0, r]$). Solving the geodesic equation, we have that the observed position β is then given by

$$\vec{\beta}(\vec{\theta}, z) = \vec{\theta} - \nabla_\perp \psi(\vec{\theta}, z) \quad \text{where} \quad \psi(\vec{\theta}, r) = \frac{2}{c^2}\int_0^r dr' \frac{\chi(r - r')}{\chi(r')\chi(r)}\Phi(f_K(r')\vec{\theta}, r'), \tag{1.69}$$

where $\nabla_\perp$ is the gradient with respect to the two projected angular directions on the sky plane, ψ is the weighted gravitational potential "seen" by the photon along its unperturbed path (called the lensing potential, derived from the true potential Φ) and χ is the transverse comoving dis-

tance. The lensing of background sources is usually studied through two different observational features.

- *Magnification of point sources*: Weak lensing preserves the surface brightness I of an observed galaxy, i.e., $I_{\rm true}(\vec{\theta}) = I_{\rm obs}[\vec{\beta}(\vec{\theta})]$, but modifies its apparent solid angle spanned by the source. Then, the galaxy's apparent flux is *magnified*, i.e., fainter sources appear brighter. At the same time, their number density is *diluted* since the apparent solid angle is enlarged. For a flux-limited survey, the *net* galaxy number density depends on these two competing effects (how many more sources are magnified depending on the "un-lensed" magnitude distribution of galaxies versus the lensing dilution of the surface density).

- *Deformation of source galaxy images*: Considering a source galaxy image to be an ellipse with intrinsic ellipticity $\epsilon^{\rm int}$, its observed ellipticity $\epsilon^{\rm obs}$ is given by

$$\epsilon^{\rm obs} = \frac{\epsilon^{\rm int} + g}{1 + g^* \epsilon^{\rm int}}, \tag{1.70}$$

  where $g = \gamma/(1-\kappa)$ is the reduced shear, and γ and κ are respectively the shear and the convergence, both depending on the lensing potential Ψ (see full details in Appendix B). Then, the measurement of galaxy shapes can be directly related to the projected matter distribution, the latter inducing a coherent deformation of source images. The observed ellipticity is then an unbiased tracer of the shear field such as $\langle \epsilon^{\rm obs} \rangle \approx \gamma$ for $\gamma, \kappa \ll 1$ (see details in Appendix B). The lensing of galaxy shapes is usually combined with the galaxy clustering into the so-called "3×2pt" probe, namely three 2-point correlation functions: the galaxy-galaxy clustering $\langle \delta_g \delta_g \rangle$, the shear field around galaxies $\langle \delta_g \gamma \rangle$ and the shear-shear correlation function $\langle \gamma\gamma \rangle$.

We refer the reader to the review of Schneider (2006); Baker (2017); Bartelmann & Schneider (2001) but also on the online "Notes on Gravitational Lensing" of S. J. Hilbert.

Galaxy clusters: Galaxy clusters are the largest matter density peaks in the Universe and represent the most advanced step in the hierarchical process of structure formation. Clusters of galaxies are unique tracers of the matter density field, since they reveal both the dark matter and the baryon content in the Universe. Since galaxy clusters lie at the intersection of the cosmic web filaments, their 2-pt correlation function/power spectrum can be used as an alternative manner to access the statistical properties of the underlying matter density (e.g., see Fedeli et al. (2011); Marulli et al. (2021)). Assuming that the galaxy cluster content is fairly representative of the baryon budget in the Universe, cluster baryon fraction can also be used to constrain cosmology (Allen et al., 2008; Ettori et al., 2009). They can also be studied through their abundance and their lensing signal, that are the two focuses of this thesis, and they will be more thoroughly presented in Chapter 2.

Other probes: So far we discussed galaxy clustering, weak gravitational lensing and galaxy clusters which are the main cosmological probes of the late Universe used in optical surveys. Here below we discuss more briefly other complementary probes such as cosmic voids, Lyman-α forests and CMB lensing.

- *Cosmic voids*: They are vast under-dense regions of space. They are great laboratories to study the nature of gravity and dark energy, by measuring their abundance as well as the void-galaxy cross-correlation function, which contains a wealth of information about the density and peculiar motion of matter in their surrounding regions. See Correa (2023) for a review.

- *Lyman-α forests*: The Lyman-α forest is a series of absorption lines in the spectra of distant quasars, highly luminous active galactic nuclei (AGN) observed over a large redshift range up to $z \sim 7$. This absorption arises from the electronic transition of neutral hydrogen (occurring at $\lambda = 1216$ angstrom) lying in the foreground inhomogeneous structures along the photon path. A change in the absorbed flux is closely related to a change in the gas density in the Universe, dictated by both small and large-scale mass distributions in the Universe. So the Lyman-α forest offers a unique window to probe the thermal history of the intra-galactic medium, but also the properties of the large-scale structures measuring with great precision the matter clustering. We refer the reader to Weinberg et al. (2003) for a review.

- *CMB lensing*: The observed mapping of the CMB temperature is perturbed due to the lensing of CMB photons traveling the Universe from the last scattering surface. It follows the Liouville theorem $T_{\rm CMB,obs}(\vec{\beta}) = T_{\rm CMB}(\vec{\theta})$ where $\vec{\beta} = \vec{\theta} - \nabla_\perp \psi_{\rm CMB}(\vec{\theta})$ (see Eq. (1.69)), meaning that photons are deflected by the convergence field $\kappa_{\rm CMB}(\vec{\theta}) = \kappa(\vec{\theta}, z_\star = 1100)$ where $z_\star$ is the redshift at recombination epoch. The lensing of the CMB induces secondary anisotropies in both CMB temperature and polarization, that can be studied to infer properties of the matter distribution up to very high redshifts. See Lewis & Challinor (2006) for a review.

The various cosmological probes that we presented in this section allow the community to study the Universe at many different epochs, from the last scattering surface to the growth of very large scale structures. Combining these many sources of cosmological information can lead to improve significantly the constraints on our cosmological model. The "concordance" between the constraints from cosmological probes has emerged as a great argument in favor of the ΛCDM paradigm. As an example, Fig. 1.5 from Suzuki et al. (2012) illustrates the "concordance" between early and late probes namely the Type1a supernovae (blue), baryon acoustic oscillations (green) and the CMB (orange) for the inference of Ω_m and Ω_Λ. This figure dates back 10 years but allows us to illustrate the power of multi-probe analysis, i.e., how combining the different cosmological probes decreases significantly the uncertainty of relevant parameters, as shown in gray.

1.4.4 The experimental landscape: a brief overview

The various cosmological probes described before can be used to refine cosmological models. In that context, a wide variety of ground-based and space experiments browsing a wide range of wavelengths are designed to study the laws of the Universe, from the ΛCDM paradigm to many extensions, such as inflation, dynamical dark energy, neutrinos, etc. Next-generation surveys will benefit from high precision, width and depth to target multiple tracers out to high redshift and over larger volumes. In the following paragraphs, we draw a non-exhaustive portrait of current and near-future cosmological experiments.

CMB experiments: First, the Universe can be observed through the millimeter light of the CMB, which gathers a wealth of information on its early life, during the radiation era, notably of the primordial power spectrum generated before inflation. With higher resolution and sensitivity the *Planck* mission (Planck Collaboration et al., 2020a) that succeeded the Wilkinson Microwave Anisotropy Probe (WMAP, Bennett et al. (2013)) provides today's most accurate constraint on ΛCDM cosmology by mapping the level of CMB anisotropies and its polarization. Upcoming CMB experiments such as the Simons Observatory (Ade et al., 2019) in 2023, the

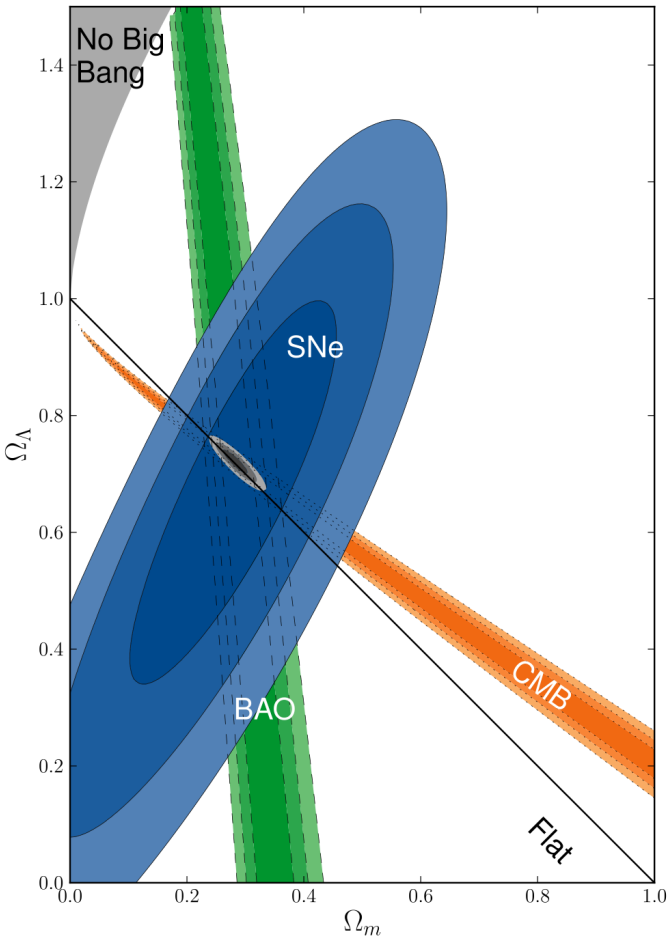

Figure 1.5: 68.3%, 95.4%, and 99.7% confidence regions of posteriors in the $\Omega_m - \Omega_\Lambda$ plane from Type1a supernovae detected in the HST Cluster Supernova Survey (blue contours), baryon acoustic oscillation measured on the clustering of SDSS DR7 galaxies (Percival et al., 2010) (green) and CMB measurements of WMAP (Komatsu et al., 2011) (orange). Figure from Suzuki et al. (2012).

space mission *liteBIRD*[4] (LiteBIRD Collaboration et al., 2023) in 2027, and the ground-based experiment CMB-S4 (Abazajian et al., 2016) in 2027 will put tight constraints on standard ΛCDM cosmology, but will also address fundamental questions about its numerous extensions. For instance probing inflation, by searching for Primordial Non-Gaussianities (PNG) that are predicted in a wide range of magnitude by a plethora of multi-field inflation models that generate non-zero three-point correlation function of CMB anisotropies. They will also search for B-mode polarization of CMB anisotropies, generated by primordial gravitational waves during inflation. CMB experiments also extract a wealth of information on the cosmological scenario from the mapping of the large-scale structure, via the CMB lensing by foreground structures from the last scattering surface to the late Universe. But also through the scattering of CMB photons by the hot electron gas around galaxy clusters, that we discuss in the next chapter (e.g., South Pole Telescope (SPT, Ruhl et al. (2004)), Atacama Cosmology Telescope (ACT, Fowler et al. (2007))) whose properties are highly sensitive to the growth rate late expansion history,

[4]Lite (Light) satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection

matter content. They are also excellent candidates for the search for massive neutrinos, that tend to smooth the formation of structure at small scales.

Galaxy surveys: In optical and near infra-red bands, galaxies can be used to trace the presence of matter in-homogeneity in the late Universe. The late Universe is highly in-homogeneous since structures have collapsed gravitationally and expansion is dominated by the presence of dark energy. Redshift-space distortion (RSD) or baryon acoustic oscillations (BAO) can be used to constrain the structure growth scenario and dark energy, as studied by the Sloan Digital Sky Survey (SDSS, Alam et al. (2021)). By observing millions of galaxies and quasars the Dark Energy Spectroscopic Instrument (DESI, Levi et al. (2019)) is measuring the BAO peak and RSD and will also measure the effects of inflation through the PNG that have specific imprints in the large-scale clustering of galaxies. Moreover, the late Universe can be probed by measuring the weak gravitational effect of foreground structures on source images. It can be combined with the galaxy clustering, into the so-called 3×2pt analysis (three two-point correlation functions) such as with the Dark Energy Survey (DES, The Dark Energy Survey Collaboration (2005)), the Hyper Suprime-Cam (HSC, Aihara et al. (2018)) and the Canada France Hawaii Telescope (CFHT, Heymans et al. (2012)). During the next decade, the *Euclid* space mission (Laureijs et al., 2011), the Vera C. Rubin observatory (LSST Science Collaboration et al., 2009) that we will present in Chapter 4, or *Roman* Space Telescope (Akeson et al., 2019) will measure the weak lensing by large-scale structures with unprecedented precision, but also the abundance of galaxy clusters and cluster lensing (these aspects will be detailed in the next chapter) that will enable to unveil the hidden properties of dark matter and dark energy. At radio wavelength, the Square Kilometre Array (SKA, Weltman et al. (2020)) in 2027 will provide significant improvements in our understanding of the accelerating Universe and the inflationary era.

The next decade will see the emergence of experiments conducting very large cosmological surveys, that will either confirm further ΛCDM or comfort the cracks that are appearing.

1.5 Challenges of modern cosmology

The experiments mentioned previously will help address the main challenges of modern cosmology. We divide the different challenges into two parts, the first one presents the "unknowns" of the current cosmological model such as the nature of inflation, dark matter and dark energy. The second refers to the observational tensions in ΛCDM that have emerged in the last ~ 10 years. Other unsolved problems of the ΛCDM are in the forefront of research in cosmology, high-energy astrophysics and particle physics, from both theoretical and observational point of view, such as the unknown neutrino masses, the observed abundances of species predicted by the Big Bang Nucleosynthesis (the "missing baryon problem"), etc., that will not be addressed in this section.

1.5.1 The unknowns

- *Inflation and primordial fluctuations:* As mentioned before, the theory of inflation is typically invoked as a solution to many problems of the standard Big Bang cosmology, e.g., that early quantum fluctuations were the seed for the formation of large-scale structures. Numerous degrees of freedom of the theory are unknown, and open questions still remain about the detailed *physics* of inflation, e.g., the shape of the inflaton potential, possible multi-field inflation, the shape of the primordial power spectrum, and how did inflation ended. For instance, the presence of CMB B-modes would corroborate the inflation scenario.

- *Dark energy:* Dark energy is invoked to explain the late accelerated expansion of the universe, thus understanding its nature and properties is crucial for modern cosmology. The ΛCDM paradigm assumes that dark energy acts as a constant energy density and is represented by the cosmological constant Λ. It is commonly assumed that dark energy corresponds to the zero-point energy of vacuum suggested by quantum field theory, which is predicted to be 120 orders of magnitude higher than of that from cosmological observations. This issue is known as the "cosmological constant problem" (Weinberg, 1989) and as "the largest discrepancy between theory and experiment in all of science" (Adler et al., 1995). In cosmological applications, going beyond the Λ hypothesis permits to test a variety of dark energy models and try to answer fundamental questions about the late expansion. Dark energy can be assumed to be "dynamical", i.e., as a fluid that can cluster with itself and with the other cosmological fluid, with proper interactions with dark matter. Therefore, dark energy would have a significant impact on the matter clustering over a variety of scales. Moreover, as favored by a variety of physically motivated dark energy models, its equation of state is allowed to differ from -1 and to have possible time evolution, which can follow the Chevallier-Polarski-Linder parametrization (Chevallier & Polarski, 2001; Linder, 2003)

$$w(a) = w_0 + w_a(1-a), \tag{1.71}$$

  where the coefficients w_0 and w_a ponder the evolution of the equation of state through cosmic time. That turns to $w_0 = -1$ and $w_a = 0$ for a pure cosmological constant.

- *Dark matter:* It accounts for about 85% of the matter content in the Universe and has a strong influence on the formation of large-scale structures and matter clustering in the Universe. However, its nature is unknown. Dark matter remains difficult to detect since it interacts with its environment exclusively through gravity. In high energy particle physics, the search for dark matter is very active, trying to measure a variety of scatterings or couplings of beyond-standard-model dark matter candidates (e.g., Weakly Interactive Massive Particle (WIMPS), axions) and standards (neutrinos) with standard particles in laboratories. But also by looking at cosmic rays or gamma rays, trying to detect anomalous components originating in dark matter annihilation, and identify potential dark matter candidates (see the review of Feng (2010) for dark matter candidates from Particle Physics). However, there is still no evident detection. It is also argued that the "dark matter" observational effects could be explained by the presence of massive astrophysical compact halo object (MACHO) made of baryons, such as hypothetical primordial black holes created soon after the Big Bang, faint neutron stars, etc. The necessity of a "dark matter" in ΛCDM could also be explained by some modifications of the standard Einstein's general relativity (MOdified Newtonian Dynamics (MOND, Milgrom (1983)), where the presence of supplemental non-baryonic matter is not invoked. Thus, the nature of the dark matter keeps being one of the most intriguing problems in fundamental physics and cosmology.

- In this paragraph, we discuss some other theoretical unknowns of current cosmological models that are the target of modern cosmology, namely the non-linear power spectrum, the impact of baryons, and the intrinsic alignment of galaxies.

  Non-linear power spectrum: In the linear regime, when the density perturbation $\delta = (\rho(\vec{x}) - \bar{\rho})/\bar{\rho} \ll 1$, the evolution of the matter density field can be derived using linear perturbation theory, that is all we need to study the CMB. However, most of the field of structure formation, galaxy formation, and collapsed objects in the late Universe concerns the understanding of the non-linear regime. When $\delta \sim 1$, the linear perturbation theory

is no longer valid and modes start to couple to each other, and the growth rate is no longer sufficient to describe the evolution of the density field (Peebles, 1980). In general, no analytic solutions for the non-linear power spectrum exist but fitting formulas of the non-linear power spectrum have been developed based on simulations (Smith et al., 2003; McDonald et al., 2006). Although their precisions are sufficient for most of the current applications, future surveys would have statistical uncertainties so low that fitting formulas with improved accuracy is highly desired.

Intrinsic alignment of galaxies: The upcoming high-quality shear data provided by the next-generation galaxy surveys requires an accurate understanding of systematic effects. One of the major astrophysical effects that contaminate the measure of cosmic shear is the intrinsic alignment of source galaxies. Galaxies are subject to tidal fields originating from their large-scale environment during their formation and evolution (see e.g., White (1984); Catelan et al. (2001)). Then, their shapes become correlated with the large-scale gravitational field. Their observed shapes can be decomposed following $\epsilon_{\rm obs} = \epsilon_{\rm int} + \gamma + \gamma^I$ where γ^I represents the intrinsic alignment contribution. It is vital to understand the intrinsic alignment and to isolate it from the *standard* cosmic shear signal, to extract the cosmological information from weak lensing observable in an unbiased way, and to do so is not trivial.

Impact of baryons: The understanding of the small-scale clustering of matter is essential in the prediction of weak lensing observable. The total matter power spectrum needs to be known at least up to $k = 10\ h.\mathrm{Mpc}^{-1}$ to achieve the desired accuracy in the inference of cosmological parameters (Huterer & Takada, 2005) in the context of the next generation of weak lensing data. In the past, it was sufficient to model the total matter power spectrum analytically (Takahashi et al., 2012) or using dark matter-only simulations (Heitmann et al., 2014), but baryonic effects can have a significant impact on the distribution of matter (van Daalen et al., 2011). The main effect to be modeled is the suppression of clustering at scales of a few Mpc, caused by gas injection by active galactic nuclei (AGNs). Hydrodynamical simulations have been used to provide these predictions (van Daalen et al., 2011; Vogelsberger et al., 2014; Springel et al., 2018), that differ from one another in the numerical methods used and the implementation of baryonic physics[5]. The precision of upcoming cosmological constraints will rely on the modeling of the impact of baryons on the clustering of matter at small scales.

1.5.2 Observational tensions in ΛCDM

Some discordance between the early Universe versus late-time probes appeared in the past decade, such as the disagreement between the inferred values of H_0 and S_8, that we detail below (see the review by Abdalla et al. (2022)) and a summary of which is given in Figs. 1.6 and 1.7.

- *The H_0 tension:* In a nutshell (see the very complete review by Valentino et al. (2021)), the cosmological analyses that are based on the direct local distance ladder measurements (*direct* measurement of the local expansion rate, at $z \leq 1$) favor higher values for H_0 than that from the CMB light emitted at $z = 1100$ (providing *indirect* measurement of H_0 by assuming a cosmological model). For instance, Planck Collaboration et al. (2020a) CMB analysis led to a Hubble constant $H_0 = 67.27 \pm 0.60\ \mathrm{km.s}^{-1}.\mathrm{Mpc}^{-1}$, whereas Riess et al. (2021) measured $H_0 = 73.2 \pm 1.3\ \mathrm{km.s}^{-1}.\mathrm{Mpc}^{-1}$ from Type1a supernovae, or else Denzel

[5]Which can in turn result in different predictions for the clustering of matter at small scales, where the baryon feedback is important.

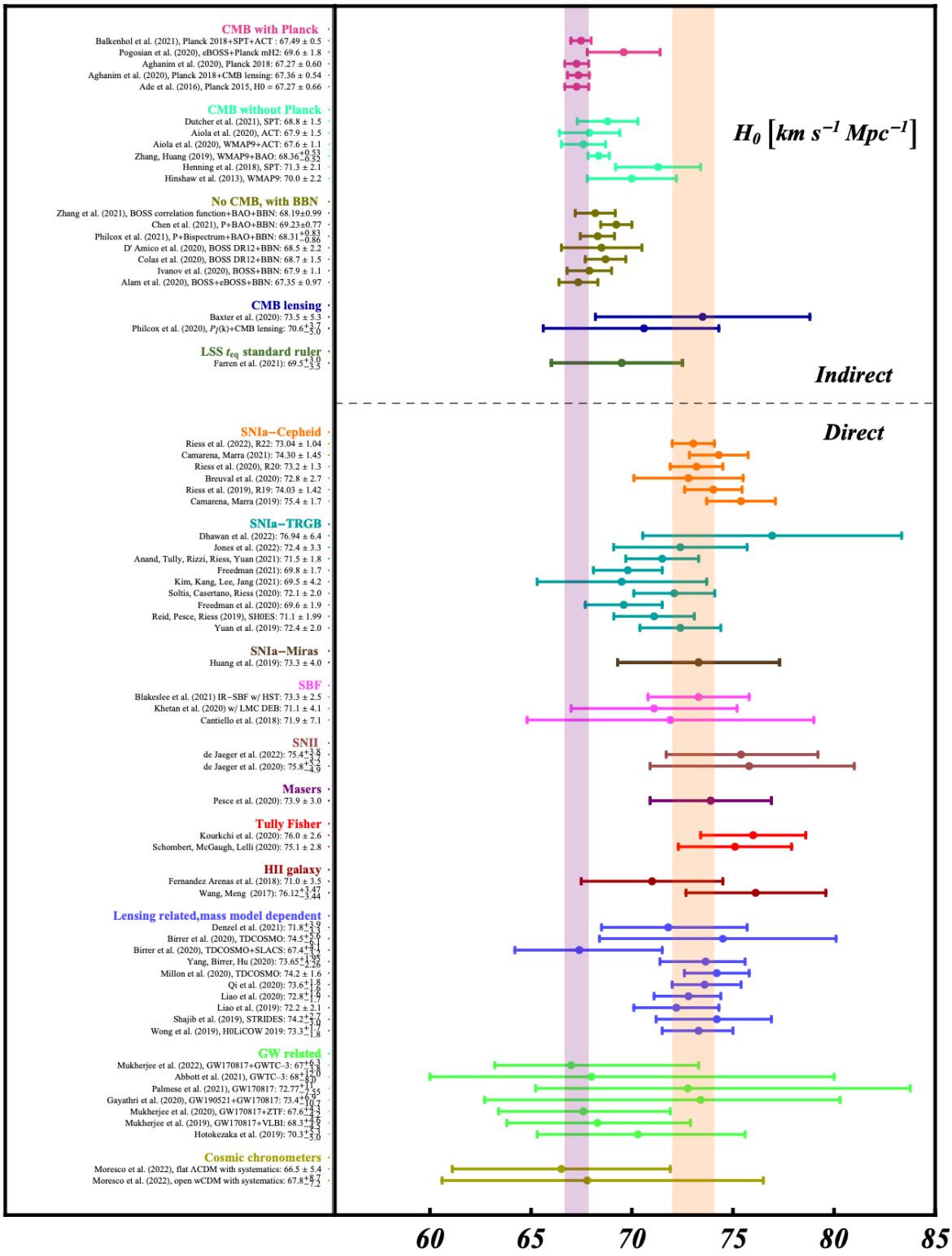

Figure 1.6: H_0 constraints for a variety of cosmological analyses and corresponding 68% errors. Figure from Abdalla et al. (2022).

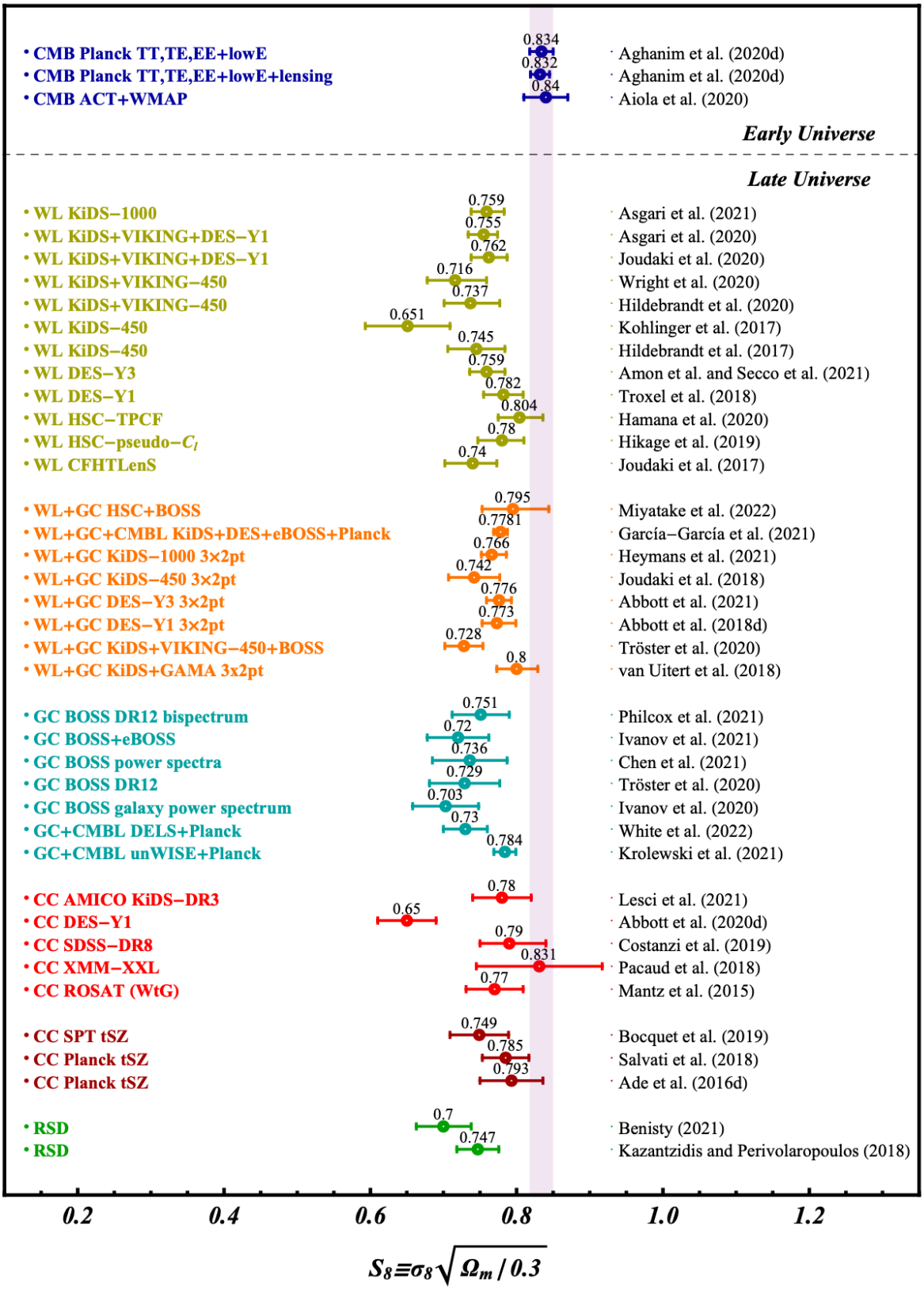

Figure 1.7: S_8 constraints and corresponding 68% errors. Figure from Abdalla et al. (2022).

et al. (2021) measured $H_0 = 71.8 \pm 3.9$ km.s^{-1}.Mpc^{-1} from time delay measurements[6]. The situation is even more clear in Fig. 1.6 from Abdalla et al. (2022) that shows a collection of H_0 estimates, where CMB constraints (above the dashed line) are lower than that obtained from distance measurements (below the dashed line) by few sigmas. The H_0 discrepancy is one of the major anomalies when comparing current CMB constraints to that obtained from late Universe probes, and the upcoming increasing precision of the data on both sides presents strong challenges to search for the possible solution to these problems, such as multiple observational unrelated errors, but also hints at a possible failure in the standard ΛCDM paradigm.

- *The S_8 tension:* The parameter $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$ quantifies the level of structure formation. The cosmological analyses based on the large-scale structures lying in the late Universe (namely the weak gravitational lensing, galaxy clustering and cluster abundance, at redshift $z \leq 1-2$) favor weaker S_8 compared to that provided by early Universe probes such as CMB measurements. For instance, the 3×2pt KiDs analysis (Heymans et al., 2021) measured that $S_8 = 0.783 \pm 0.024$, then highlighting a 3σ tension with respect to the *Planck* CMB constraints $S_8 = 0.834 \pm 0.016$ (Planck Collaboration et al., 2020a). An overview of the S_8 tension is shown in Fig. 1.7, where CMB constraints are represented above the black dashed lines (for "early Universe"), and S_8 constraints from late large-scale structures ("late Universe") are shown below the dashed lines, for example from galaxy clusters (red and purple), that is the main focus of this thesis, but also for the other large-scale structure probes (galaxy weak lensing: olive, galaxy weak lensing and clustering: orange, galaxy clustering: cyan and redshift space distortion: green). This tension has added another issue over the validity of ΛCDM. Providing solutions to the S_8 tension is one of the most prominent challenges in modern cosmology.

1.6 Conclusions

The standard cosmological model is built on quite simple hypotheses and has permitted a variety of predictions that were corroborated by numerous cosmological probes, such as the expansion of the Universe, the Cosmic-Microwave background, as well as the emergence of today's large-scale structures from the growth of initial density perturbations. Reaching high statistical accuracy on cosmological parameters is the main purpose of the precision cosmology experiments in the next decade, that will address fundamental questions about the nature of more than 90% of the energy content that is still unknown, and will try to solve the surprising persistent discrepancies that have appeared between different cosmological probes known as the H_0 and S_8. This thesis focuses on the cosmological insights we can gain by using galaxy clusters, especially by measuring their abundance and their weak gravitational lensing effect, that we detail in the next chapter.

[6] From gravitational lensing, light rays are deflected by gravity. The lens causes a deflection such that a light ray emitted by a point source will be delayed compared to its unperturbed virtual path. The time delay depends on the difference between the source position and the true position of the emitter, as well as on the mass model of the lens. In practice, it is not possible to measure an absolute time delay (we do not observe the virtual unperturbed path) but we measure relative time delays between multiple images created by strong lensing (one single source image can be lensed in multiple images). This difference is sensitive to the distance ratio $D_L D_S / D_{LS} \sim H_0^{-1}$ from Eq. (1.17) (Shiralilou et al., 2020), where D_L, D_S and D_{LS} are respectively the distance to the lens, to the source and between the lens and the source.

Chapter 2

Galaxy clusters and cosmology

Contents

Clusters of galaxies form in the densest regions of the cosmic web. Their formation is a highly non-linear process dominated by gravity and is sensitive to the properties of both dark energy and dark matter. The primary cosmological probe based on galaxy clusters is the measurement of their abundance as a function of mass and redshift. The upcoming cosmological analyses will benefit from the large catalogs of galaxy clusters to refine cosmological models at unprecedented level.

We introduce galaxy clusters in Section 2.1. They lie in the most massive collapsed dark matter halos in the Universe, that can be related to the matter density fluctuations through the halo model that we present in Section 2.2. In Section 2.3, we turn to the principle and the landscape of cluster abundance cosmology of the last 20 years. The abundance of galaxy clusters has provided competitive and complementary constraints to early Universe probes on various cosmological scenarios, and we emphasize the importance of an accurate and robust calibration of the mass-proxy relation. In Section 2.4 we present the effect of weak gravitational lensing by galaxy clusters, as a powerful tool to calibrate the mass-richness relation of optically detected clusters of galaxies. We conclude in Section 2.5.

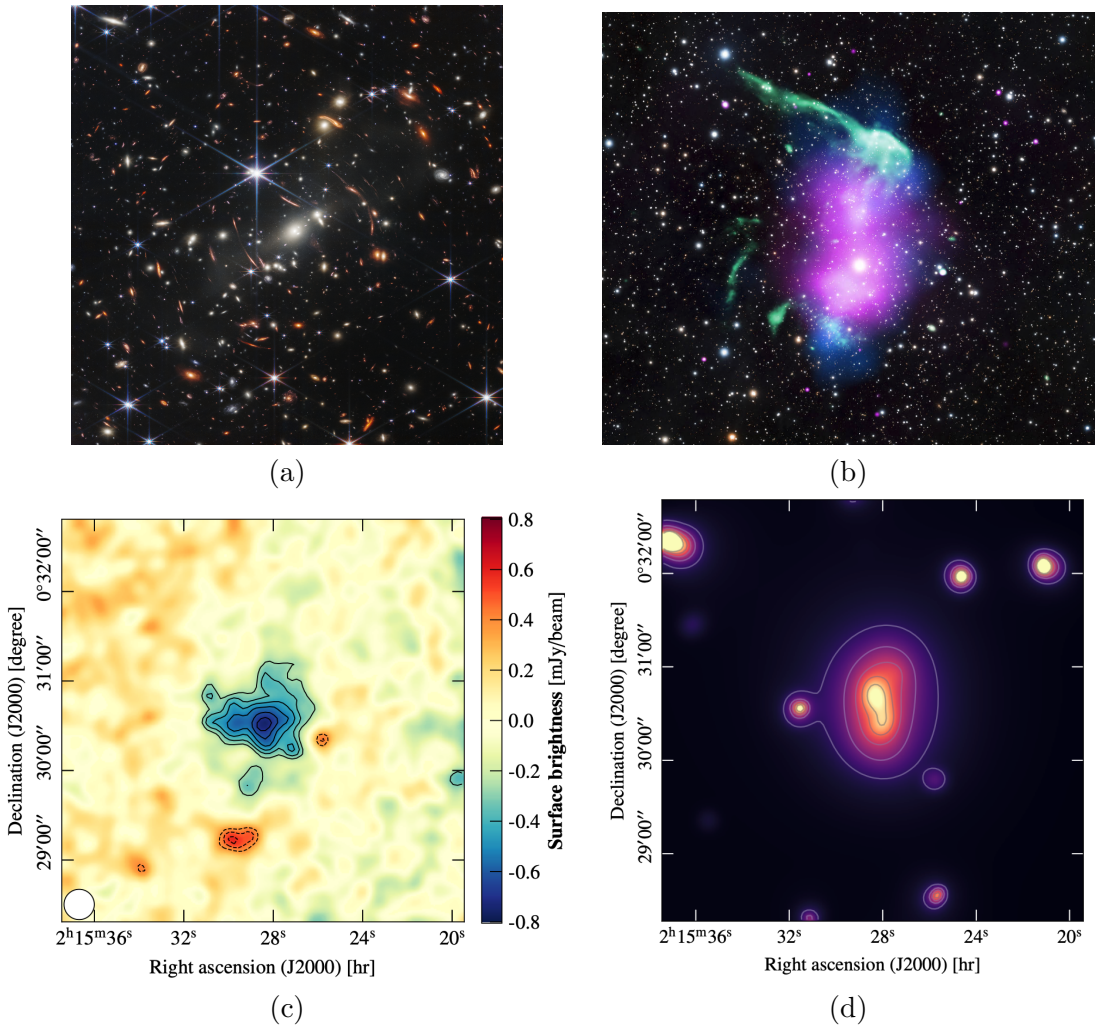

Figure 2.1: Multi-wavelength view of several clusters of galaxies. (a): Optical and near-IR view of the SMACS 0723 Galaxy cluster observed by the James Webb Space Telescope. Figure from https://www.nasa.gov. (c) is the SZ map of the ACT-CL J0215.4+0030 cluster in the 150 GHz band measured by the NIKA2 camera. (b) Optical (white, from Subaru Telescope), Radio (green, from Very Large Array telescope) and X-rays (purple, from Chandra) images of the "Toothbrush Cluster" RX J0603.3+4214. Figure from https://www.nasa.gov. (d) is the X-ray luminosity map of the same cluster measured by XMM-Newton. Images (c) and (d) from Kéruzoré et al. (2020).

2.1 What is a galaxy cluster?

A cluster of galaxies is the biggest and the most massive stable structure in the Universe that consists of tens, hundreds or even thousands of luminous galaxies and many more dwarf galaxies that are bound by gravity, with typical total mass comprised between $\sim 10^{13} M_\odot$ and $\sim 10^{15} M_\odot$, and size of few Mpc. Moreover, galaxy clusters are *recent* objects, since they are at the endpoint of structure formation from initial post-inflationary fluctuations, and retrace the late history of the Universe, typically below redshift $z = 2$.

They are also multi-wavelength emitters showing typical spectral signatures. They can be observed through the optical and near-infrared light of the luminous stars (Zwicky, 1937) in and around their member galaxies, that compose $\sim 3\%$ of their total mass (e.g., see Fig. 2.1 (a)). The main baryon contribution in galaxy clusters comes from the diffuse hot electron gas the Intra-Cluster Medium (ICM), which amounts to about 12% of their total mass. It can be measured at millimeter wavelength through the inverse Compton scattering of CMB photons off the thermal electrons, predicted by Zeldovich & Sunyaev (1969) and detected in many clusters (see e.g. Carlstrom et al. (2002)). We show an example of a Sunyaev-Zeldovich map in Fig. 2.1 (c). The ICM also emits at X-ray energies (Fig. 2.1 (d)), by thermal Bremsstrahlung of electrons, first detected by Cavaliere et al. (1971). The presence of radio emission (Fig. 2.1 (b)) in galaxy clusters also reveals cosmic rays and intense magnetic fields in the ICM (van Weeren et al., 2019). Finally, dark matter contributes the most, about 80%, being the most important source for gravitational stability. Mapping the total mass distribution within clusters, including dark matter, is possible by detecting the gravitational lensing effect of the cluster's potential on background galaxy images, that will be presented in detail in Section 2.4.

Since they arise from the collapse of initial perturbations in the matter density field under gravity, they are great laboratories for investigating gravitational structure formation, galaxy formation as well as thermodynamics of the ICM (see the very complete review from Kravtsov & Borgani (2012)). As well as being very complete astrophysical objects, they can be used to gain a wealth of information on the cosmological scenario through their abundance in the Universe, as we will see in the next sections.

2.2 Theoretical framework for cluster cosmology

The most used theoretical framework for cluster cosmology is based on the halo approach of large-scale structures also known as the "halo model" (Cooray & Sheth, 2002). According to this model, all the matter is assumed to be contained within the collapsed, virialized halos of different mass that correlates spatially to each other. Then the matter density at any $\vec{x}$ can be expressed in terms of $\rho_{\rm cl}(\vec{x}|\vec{x}_k, m_k)$, the individual matter density of halo with label k with mass m_k such as

$$\rho_{\rm m}(\vec{x}) = \sum_{k=1} \rho_{\rm cl}(\vec{x}|\vec{x}_k, m_k), \tag{2.1}$$

where the sum runs over all halos in the Universe. Then, the halo distribution can be used to describe the statistical properties of the matter density field. We introduce in this section the relevant cosmological-related quantities, such as the redshift and mass distribution of dark matter halos (i.e., describing the distribution of m_k), their clustering properties (the spatial distribution of halos through their position x_k), the mass profiles of dark matter halos $\rho_{\rm cl}$. These quantities provide a very precise description of the matter distribution in the Universe through the halo model.

2.2.1 Mass-redshift distribution of dark matter halos

The mass and redshift distribution of halos is given by the differential number of halos per redshift and mass

$$\frac{d^2N(m,z)}{dmdz} = \Omega_S \frac{dV(z)}{dzd\Omega_S}\frac{dn(m,z)}{dm}. \tag{2.2}$$

The first term denotes the partial comoving volume per redshift and comoving angle and is given by

$$\frac{dV(z)}{dzd\Omega_S} = D_H \frac{(1+z)^2 D_A^2}{H(z)/H_0}, \tag{2.3}$$

where D_H is the Hubble distance, D_A is the angular diameter distance and $H(z)$ is the Hubble parameter at redshift z.

The last term is the *halo mass function* (HMF) predicting the comoving differential number density of dark matter halos per mass range. Because structure grows with respect to the level of matter fluctuations, the clustering of the matter density field and the total matter content in the Universe, the evolution of the halo mass function is highly sensitive to cosmology. A general form of the halo mass function is

$$\frac{dn(m,z)}{dm} = \frac{\rho_m(z)}{m}\frac{d\ln\sigma}{dm} f(\sigma), \tag{2.4}$$

where $\sigma \equiv \sigma(m)$ is the variance of the density field, which is related to the matter power spectrum, filtered on a scale R enclosing an average mass $m = V_R\rho_m(z)$ as defined in Section 1.3.3, and $f(\sigma)$ is the multiplicity function (its meaning is described in Appendix C). Press & Schechter (1974) (PS) were the first to propose a model for the multiplicity function based on the spherical collapse of filtered Gaussian matter density fluctuations exceeding a critical threshold. PS formalism was refined by Sheth et al. (2001) that included ellipsoidal collapse. These aspects are discussed in detail in Appendix C. The actual dynamics of gravitational collapse and the process of dark matter halo formation is even more complex since it accounts for non-spherical collapse beyond elliptical, but also interactions between neighboring halos, the impact of baryon physics, feedback of Active Galactic Nuclei (AGN), etc. Since these early analytical works, the calibration of the multiplicity function has been widely performed by using a series of simulations (CDM N-body or hydro-dynamical simulations), where collapsed regions are identified using halo finders. From a variety of simulations but also theoretical approaches evoked before, the multiplicity function is found to be almost *universal* over a wide range of mass and redshift (see e.g., Jenkins et al. (2001); Evrard et al. (2002)). Thus, Despali et al. (2015); Tinker et al. (2008); Tinker et al. (2010); Watson et al. (2013); Bocquet et al. (2015) have devised more accurate parameterizations than the PS formalism, reaching $\sim 5-10$ % accuracy in describing the simulated dark matter halo mass distribution on a wide range of mass and redshift.

We represent some of the widely used halo mass functions from the literature in the Fig. 2.2 (top panel). Halo mass functions are represented as a function of the overdensity mass of the halo $M_{200\mathrm{m}}$ (detailed in the Section 2.2.3) and follow the Planck Collaboration et al. (2020a) cosmology. The full lines (top panel) represent the halo mass functions at $z = 0.3$, and dashed lines at $z = 0.8$. We see that when the Universe was younger, the comoving number density of halos was smaller, especially for high mass halos that formed by accreting smaller structures through cosmic time. All the simulations used to calibrate these halo mass functions used dark matter only particles, except Bocquet et al. (2015) (pink line) that included the impact of baryons. The low panel represents the relative difference of all these mass functions (with only dark matter particles) to the Bocquet et al. (2015) halo mass function (including baryons). The lower plot shows how much the assumption of halo formation in the Universe (notably the ingredients used in the simulations) may bias our understanding of the true underlying halo mass

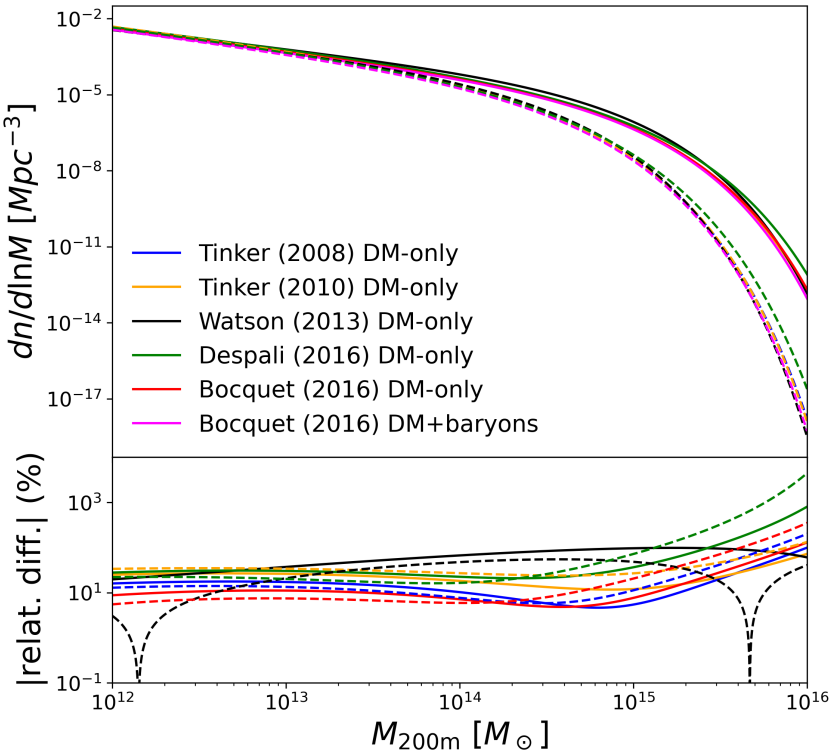

Figure 2.2: Top panel: Halo mass functions Tinker et al. (2008); Tinker et al. (2010); Watson et al. (2013); Despali et al. (2015); Bocquet et al. (2015) at redshift $z = 0.3$ (full line) and $z = 0.8$ (dashed line). The lower panel represents the relative difference to the Bocquet et al. (2015) halo mass function (including dark matter and baryons).

distribution in our Universe, since halo mass functions may differ from one to each other by 10% (low mass) to 1,000% above a few 10^{15} $M_\odot$. The halo mass function is the key ingredient for cosmological analysis based on the abundance of collapsed large-scale structures in the Universe, that we will present in Section 2.3.1.

2.2.2 Halo clustering

The halo overdensity at fixed mass m relates to that of matter such as

$$\delta_h(\vec{x}|m) = \mathcal{B}_h[\delta_m(\vec{x})|m] \approx b_{\rm h}(m)\delta_m(\vec{x}), \tag{2.5}$$

where $\mathcal{B}[\cdot]$ is a functional of the matter density field (Desjacques et al., 2018). The current paradigm for dark matter halo clustering is that at linear scales, the halos are "biased" tracers of the matter density field. It gives the last expression in the above equation that is valid at sufficiently large scales, and where $b_{\rm h}$ is the halo bias. This approximation is very useful to link the halo power spectrum to the matter power spectrum. Its mass dependence denotes that low-mass clusters correlate differently to the underlying matter density field compared to high-mass halos, with also a possible redshift evolution. For two halos with respective masses m_1 and m_2, their 2-point correlation function is proportional to the matter-matter correlation function, such as

$$\xi_{\rm hh}(r|m_1, m_2) = b_{\rm h}(m_1)b_{\rm h}(m_2)\xi_{\rm mm}(r). \tag{2.6}$$

In most of cosmological models and simulations, if $m_1 > m_2$, then $b_{\rm h}(m_1) > b_{\rm h}(m_2)$. Most halo bias models are calibrated using numerical simulations (see for example Bhattacharya et al. (2011); Tinker et al. (2010))

2.2.3 Halo mass density profile

The matter density field around the cluster is then given by

$$\rho(r|m) = \overline{\rho}_{\rm m}(1 + \xi_{\rm hm}(r|m)) \quad \text{where} \quad \xi_{\rm hm}(r|m) = \langle \delta_h(\vec{x}_1|m)\delta_{\rm m}(\vec{x}_2)\rangle. \tag{2.7}$$

where $r = |\vec{x}_1 - \vec{x}_2|$, δ_h is the halo number density perturbation and δ_m is the local matter perturbation. After calculation (Cooray & Sheth, 2002) the halo-matter cross-correlation function reaches the expression

$$\xi_{\rm hm}(r|m) = \frac{\rho_{\rm cl}(r|m)}{\overline{\rho}_{\rm m}} + b(m)\xi_{\rm mm}(r). \tag{2.8}$$

$\xi_{\rm mm}(r)$ is the matter-matter correlation function and $b(m)$ is the halo bias. We denote two terms; the 1-halo term $\xi^{\rm 1h}_{\rm hm}(r|m) \propto \rho_{\rm cl}(r|m)$ is the contribution of the single halo to the surrounding matter density field, and depends on single halo properties, the 2-halo term $\xi^{\rm 2h}_{\rm hm}(r|m) = b_h(m)\xi_{\rm mm}(r)$ indicates the contribution from the surround large scale structures, and is sensitive to the matter power spectrum. Dark matter halo density profiles $\rho_{\rm cl}$ depend strongly on the non-linear mechanisms of structure formation through gravity, and are thus very difficult to predict. Under the assumptions of radial matter distribution, the dark matter radial density profile of a cluster at redshift z is defined in terms of its overdensity mass M_Δ relative to the background density field $\bar{\rho}$

$$M_\Delta = \frac{4\pi r^3_\Delta}{3}\Delta\bar{\rho}, \tag{2.9}$$

which corresponds to the enclosed mass within a sphere with radius r_Δ in which the average matter density is equal to Δ times the background density field $\bar{\rho}$ (that can be either the mean mass density $\rho_m(z)$ or the critical density $\rho_c(z)$). Let's note that the mass definition is not only important to define the mass profile but also for the evaluation of the halo mass function and bias. The mass definition is also crucial to fit functional halo mass functions from the observed halo mass distribution in simulations, since the mass of collapsed object in N-body simulations and detected by halo finders depends strongly on the halo finder characteristics (e.g., linking length between two dark matter simulated particles).

The simplest version is the Singular Iso-thermal Sphere (SIS) profile, which tells that the equilibrium distribution of a gas of self-gravitating particle is an iso-thermal sphere, with Maxwellian particle velocity distribution (Hoffman & Shaham, 1985). Considering the infinitesimal mass element $dm \propto \rho(r)r^2dr$, we have that $\rho_{\rm SIS}(r) \propto r^{-2}$. Many effects may cause deviations from the simplest SIS profile[1], due to the high complexity of structure formation from gravitational collapse (aspherical collapse, impact of baryons, high angular momentum of mass distribution, etc.), that induces a scale-dependent slope $\gamma(r)$ of the profile $\rho(r) \to r^{-\gamma}$.

In practice, ρ is defined using parametric functions that are adjusted to match the observed matter density profile in dark matter simulations. There exist a variety of fitting functions (Navarro et al., 1997; Hernquist, 1990; Diemer & Kravtsov, 2014; Einasto, 1965) that provide good description of the collapsed halos in many simulations (see e.g. Wang et al. (2020) that tested NFW and Einasto profiles in simulations over 20 orders of magnitude in mass). The different profiles will be presented later in Chapter 4. For example, the Navarro et al. (1997)

[1] In addition to be nonphysical because of its singularity at $r = 0$.

(NFW) profile is one of the most widely used profiles in cluster studies. It is given by

$$\rho_{\rm cl}^{\rm nfw}(r) = \rho_s \left(\frac{r}{r_s}\right)^{-1} \left(1 + \frac{r}{r_s}\right)^{-2}, \tag{2.10}$$

where r_s is the scale radius and ρ_s is the scale density. The profile is defined by two inner and outer slopes, such that if $r \ll r_s$, $\rho_{\rm cl}^{\rm nfw} \propto r^{-1}$ and if $r \gg r_s$, $\rho_{\rm cl}^{\rm nfw} \propto r^{-3}$. It is usually defined by the overdensity mass M_Δ and the halo concentration $c_\Delta = r_\Delta / r_s$, that denotes the level of concentration of the mass distribution in the innermost regions of the halo. This definition is common to other halo profile modelings. The halo mass and concentration are also found to be correlated in simulations (Diemer & Kravtsov, 2014; Duffy et al., 2008; Bhattacharya et al., 2011), as a direct consequence of the hierarchical accretion history, such as massive halos appear less concentrated and reversely. Series of parametric concentration-mass relations are calibrated on simulations to account for this tight correlation, we will discuss them in Chapter 4.

2.3 The abundance of galaxy clusters as a probe for cosmology

2.3.1 Bayesian inference with cluster counts

We introduced in Section 1.4 how observations can be used to constrain the cosmological models. Probing cosmological parameters consists in deriving, from a set of observations denoted by $\widehat{d}$, the *posterior* distribution of selected cosmological parameters $\vec{\Theta}$ that define cosmological models (e.g. Ω_Λ, Ω_m, σ_8, n_s, etc.), given by the Bayes theorem

$$\mathcal{P}(\vec{\theta}|\widehat{d}) \propto \mathcal{L}(\widehat{d}|\vec{\theta})\pi(\vec{\theta}), \tag{2.11}$$

where $\mathcal{L}$ is the *likelihood* function, that gives the probability for observing the data $\widehat{d}$ when $\vec{\Theta} = \vec{\theta}$. It is the key ingredient of any probe, linking the *sensitivity* of observations to the targeted cosmological models i.e. strong sensitivity to cosmology means strong variation of the likelihood with respect to $\vec{\theta}$. The function of parameters π is the *prior* and denotes the *a priori* knowledge of the parameters $\vec{\Theta}$. Once the likelihood and prior are known, the posterior distribution can be used to compute the mean values and respective errors of some cosmological parameters (see review of statistical methods in cosmology by Verde (2010)). This section is dedicated to present the main features of cluster abundance cosmology, i.e. gaining cosmological insights from the count of clusters in the Universe.

As presented in Section 2.1, galaxy clusters are the most striking elements of the large-scale structures, and may be used to understand the growth of structures at cosmological scales. When zooming out, galaxy clusters lie at the intersections of the cosmic web filaments, so the distribution of galaxy clusters in the Universe (masses, positions, etc.) is fairly connected to the distribution of halos in the Universe, highly sensitive to cosmology (that was described in Section 2.2).

The abundance of detected clusters is usually measured within the interval of redshift $z_1 < z < z_2$ and *proxy* interval $\mathcal{O}_1 < \mathcal{O} < \mathcal{O}_2$. In optical surveys, clusters are detected via the light of their member galaxies that are bound with gravity. We can define the *richness* for each cluster, as the number of member galaxies above some luminosity threshold that is represented by $\mathcal{O} = \lambda^2$. From observational counts in several richness-redshift bins $\widehat{d} = \{\widehat{N}_i\}$, the joint

[2]Cluster proxy in millimeter and X-rays are respectively the integrated Compton parameter and X-ray surface brightness, that depends mostly on the hydro-dynamical properties of clusters (e.g., the cluster thermal pressure, electron density and temperature).

posterior distribution from cluster abundance (CA) cosmology is given by

$$\mathcal{P}_{\rm CA}(\vec{\theta}|\{\widehat{N}_i\}) \propto \mathcal{L}_{\rm CA}(\{\widehat{N}_i\}|\vec{\theta})\pi(\vec{\theta}), \tag{2.12}$$

where $\mathcal{L}_{\rm CA}(\{\widehat{N}_i\}|\vec{\theta})$ is the cluster abundance likelihood, and denotes the overall statistical properties of cluster count and its sensitivity to a change in $\vec{\theta}$. The cluster abundance likelihood can take several forms depending on the regime under scrutiny, and we explore this aspect further in Chapter 5.

When decomposed, the cluster likelihood is fully described by its moments.

- For a given redshift-richness bin, the first-order moment is given by

$$N_i(\vec{\theta}) = \mathrm{E}_{\mathcal{L}_{\rm CA}}[\widehat{N}_i|\vec{\theta}], \tag{2.13}$$

where $\mathrm{E}_{\mathcal{L}}[\cdot]$ denotes the ensemble average over all possible sampling of the data (i.e., over several realizations of our Universe), as introduced in Section 1.3.3. It corresponds to the *mean* cluster abundance within the redshift-richness interval i at fixed $\vec{\theta}$. First, the partial abundance of observed halo/galaxy clusters per richness, mass and redshift range is given by the general form

$$\frac{d^3N(\lambda,m,z)}{d\lambda dm dz} = \Omega_S \underbrace{\frac{d^2V(z)}{dz d\Omega_S}}_{\text{Volume}} \underbrace{\frac{dn(m,z)}{dm}}_{\text{HMF}} \underbrace{P(\lambda|m,z)}_{m\text{-}\lambda\text{ relation}} \underbrace{\Phi(m,\lambda,z)}_{\text{Selection function}} . \tag{2.14}$$

In the above equation, the most cosmology-dependent quantities are the halo mass function and the comoving volume. As presented in detail in Section 2.2, the mass and redshift distributions of halos, whose galaxy clusters are the optical analogous, is a very sensitive probe to cosmology, i.e. (i) the growth rate scenario of fluctuations, (ii) the non-linear physics of gravitational collapse that are both encoded in the matter power spectrum, but also to (iii) the late expansion of the Universe driven by dark energy diluting the number density of clusters. It is modified to account for observational properties, through P and Φ in the above equation are associated to both astrophysical properties of galaxy clusters themselves and the cluster detection method.

  - **Cluster scaling relation** The function $P(\lambda|m,z)$ is the cluster scaling relation, or equivalently the mass-richness relation (it can be expressed as $P(m|\lambda,z)$). This relation links the cluster proxy to the underlying halo mass and represents the fact that, due to the stochastic process of cluster formation and baryon physics occurring in its innermost regions, the cluster richness λ is a statistical variable for a halo with fixed mass m. Usually, this function includes some scatter, implying that for a given halo mass and redshift, the richness of a halo is not derived from a one-to-one relation. From variety of observations (e.g., SZ: Arnaud et al. (2005), X-rays: Maughan (2007) and optical: McClintock et al. (2019)) and simulations, the mean observable $\langle\mathcal{O}|m\rangle$ is generally assumed to follow a power law relation (Evrard et al., 2014). For the richness λ, we have that

$$\langle \ln\lambda|m\rangle = \int d\lambda \ln\lambda P(\lambda|m,z) = \ln\lambda_0 + \beta\ln(m/m_0). \tag{2.15}$$

The actual precision on cosmological parameters obtained from cluster abundance is currently limited by our understanding of the mass-richness relation (see review by Pratt et al. (2019)), and more particularly by our ability to estimate cluster masses (see the next Section 2.3.3). At optical wavelengths, this relation can be calibrated from the weak gravitational lensing of galaxy clusters, that we will present in Section 2.4.

– **Selection function** Φ is a function of the cluster properties and survey conditions. The selection function also depends on the cluster finder algorithm that is used to find clusters, that can be affected by a series of internal systematic and observational effects. It encodes how many clusters are effectively detected compared to the parent population that is present in the survey, and its understanding is a prerequisite to any statistical analysis of abundance datasets. We will address this point in Chapter 4 when considering the selection function of cluster finders in the simulated data of the LSST Dark Energy Science Collaboration.

Other "systematic" effects of the cluster detection method need to be carefully accounted for. First, cluster redshift and richness are both derived from the cluster member galaxies. The "observed" richness $\lambda_{\rm obs}$ and redshift $z_{\rm obs}$ may be biased and/or scattered around their "true" values. For simplicity, we do not account for this in the above equation assuming that cluster richness and redshift are perfectly known (i.e., $z_{\rm obs} = z$ and $\lambda_{\rm obs} = \lambda$), but they are important to reach accurate constraints. Finally, we need to marginalize Eq. (2.14) over the mass dimension, to account for all possible values of halo mass down to the minimum mass $m_{\rm lim}$ reached by the cluster sample (this quantity is usually not known, and is sometimes taken to be 0). Finally, the total cluster abundance prediction is then given by

$$N_i(\vec{\theta}) = \int_{z_1}^{z_2} dz \int_{\lambda_1}^{\lambda_2} d\lambda \int_{m_{\rm lim}}^{+\infty} dm \frac{d^3N(\lambda, m, z)}{d\lambda dm dz}. \tag{2.16}$$

Bolliet et al. (2020) quantified the sensitivity of the total cluster count $N_{\rm tot}$ to the growth of structure (i.e. the two parameters Ω_m and σ_8) for SZ-detected cluster cosmology, such that it is well approximated by

$$N_{\rm tot} = \sum_i N_i \propto \Omega_m^{2.9} \sigma_8^{9.2}, \tag{2.17}$$

telling that, cluster abundance is a steep function of cosmological parameters; for instance, in a Universe with high Ω_m, the formation of structure tends to increase significantly the cluster count. This trend is also verified for the amplitude of matter fluctuations since increasing fluctuations tends to form more and more clusters.

- The second moment of the likelihood gives the covariance between two different richness-redshift bins, i.e. the intrinsic level of scattering of the data as well as their mutual correlation and is given by

$$\Sigma_{ij} = \mathrm{E}_{\mathcal{L}_{\rm CA}}[(\widehat{N}_i - N_i)(\widehat{N}_j - N_j)|\vec{\theta}], \tag{2.18}$$

which can be decomposed such as (Hu & Kravtsov, 2003)

$$\Sigma_{ij} = N_i \delta_{ij}^K + N_i N_j b_i b_j S_{ij}^{\rm SSC}. \tag{2.19}$$

The first term is the intrinsic variance of a counting experiment known as the Poisson shot noise. Poisson covariance between bins is null since the Poisson process describes uncorrelated counts. The second term is known as the Super Sample Covariance (SSC) and is proportional to $N_i N_j$. It corresponds to the impact of the clustering properties of halos within and beyond the survey geometry, i.e., scales with the halo bias b_i, and the intrinsic fluctuation and clustering amplitude of the matter density field in $S_{ij}^{\rm SSC}$. This contribution increases the single variance of any cluster count in addition to shot noise and correlates counts between them.

So, we defined the first and second moment of the likelihood, that fairly describe the sensitivity (strong dependence to cosmological quantities e.g. see a review by Bartlett (1997); Allen et al. (2011)) and precision of cluster abundance cosmology through its covariance. According to the Fisher formalism (Fisher, 1935; Cramer, 1952), it is possible to forecast the performance of the cluster abundance statistics to constrain cosmological parameters. For a set of measured abundances $\{\widehat{N}_i\}$, the precision (the inverse variance) of the cosmological parameters that are obtained from a Bayesian analysis can be obtained from the first and second moment of the cluster abundance likelihood. As presented in detail in Appendix C, the parameter precision increases with the responsivity of the abundance cosmological model in Eq. (2.16) to a change in parameters (see Eq. (2.17)) that refers to the first moment of the abundance likelihood. Second, accounting for the second moment, we can show that the parameter precision increases with the total predicted abundance $\widehat{N}_i$ in each bin, proportional to the sky area Ω_S, however being lowered by the super-sample covariance term $S_{ii}^{\rm SSC} N_i \sim$ cst with respect to Ω_S. So, large galaxy surveys, such as the Rubin/LSST (LSST Science Collaboration et al., 2009), *Euclid* (Laureijs et al., 2011) can then significantly contribute to increasing the precision of cosmological parameters by providing complete, pure and deep cluster catalogs on a very large fraction of the sky. However, this statistical precision may not be fully exploited if some systematic effects are not under control. As presented in the next section, high-quality data for weak gravitational lensing is essential to calibrate the mass-observable relation for precision cosmology (see the very complete review of the impact of the calibration of the cluster mass-proxy relation on cosmology in Pratt et al. (2019), also see e.g., the DES cluster abundance analysis in Abbott et al. (2020)), as well as the understanding of all cluster detection systematic effects. Also, the impact of uncertainties related to the calibration of the halo mass function itself may impact significantly cosmological constraints for large future surveys (Artis et al., 2021).

Even though we will focus in this thesis on the sensitivity of cluster abundance to standard ΛCDM cosmology (see Section 1.3.1), clusters can also be very useful to constrain various extensions of the standard model, such as the equation of state of dark energy in wCDM cosmology, since dark energy impact strongly how structures grow (see e.g. Lee (2012); Linder & Jenkins (2003); Huterer et al. (2015); Wang & Steinhardt (1998)). Galaxy clusters are also excellent candidates for the search of massive neutrinos, since they imprint specific features in the matter clustering at small scales, and may have a strong impact on the abundance of very massive halos (Böhringer & Chon, 2016; Bolliet et al., 2020). Departures from general relativity such as modified gravity models predict enhanced gravitational potentials on cluster scales, affecting the abundance of rare massive halos (e.g., $f(R)$ models, see Cataneo et al. (2015); Ferraro et al. (2011)). Being very sensitive to the initial configuration of density perturbations that source the formation of structures, clusters can put strong constraints on primordial non-Gaussianity predicted by a plethora of inflationary models (as shown by Robinson et al. (2000); Trindade et al. (2012)).

2.3.2 The cluster scaling relation

Every cosmological analysis based on cluster abundance is conceptually very simple: from a proxy-selected cluster survey, we need to determine the efficiency of the cluster detection (the term Φ in Eq. (2.14)) as well as the relation between the cluster observable and the underlying halo mass $M_{\rm true}$, that connects to the cosmological model via the halo mass function (the term P in Eq. (2.14)). To do so, several approaches have been studied in the literature and used in the cluster abundance analyses that we present in the next Section 2.3.3. The mass-proxy relation can be calibrated *prior* to the inference of the cosmological parameters by using cluster mass estimates, or can be fitted jointly with cosmology by using the count response (but not

exclusively) to a change in scaling-relation parameters. Each analysis usually explores all of these methods depending on the characteristics of the dataset.

The first method requires to measure the mass of clusters. First, the "total" cluster mass can be obtained from their weak gravitational lensing effects on background galaxies unveiling their total projected mass distribution (see Section 2.4)

$$\Sigma(R) = \int dz \rho(x, y, z), \tag{2.20}$$

where $R = \sqrt{x^2 + y^2}$, z is the line-of-sight coordinate and ρ is the cluster mass density, allowing to estimate the weak lensing mass $M_{\rm WL}$ as it is done in optical surveys e.g., DES, KiDs, SDSS. Still at optical wavelengths, the mass $M_{\rm dyn}$ can be obtained by measuring the velocity dispersion $\sigma_{\rm dyn}$ of member galaxies (needing spectroscopic data). For instance, from Munari et al. (2013), the dynamical mass is linked to the galaxy velocity dispersion via[3]

$$\frac{M_{\rm dyn}}{10^{15}\ M_\odot} = \left(\frac{\sigma_{\rm dyn}}{A}\right)^{1/\alpha}, \tag{2.21}$$

where $A = 1177$ km.s^{-1} and $\alpha = 0.364$. Moreover, the thermodynamic properties of the ICM such as the X-ray luminosity of the hot electron gas (sensitive to electron density $n_{\rm gas}$ and temperature $T_{\rm gas}$) provide the cluster X-ray mass profile (Ettori et al., 2013)

$$M_{\rm X}(< r) = -\frac{kT_{\rm gas}(r)r}{\mu m_p G}\left(\frac{\partial \ln T_{\rm gas}(r)}{\partial \ln r} + \frac{\partial \ln n_{\rm gas}(r)}{\partial \ln r}\right), \tag{2.22}$$

under the assumption of the Hydro-Static Equilibrium (HSE). In the above equation, k is the Boltzmann constant, m_p is the proton mass and μ is the mean molecular weight of the electron gas. The HSE mass can also be inferred from the Sunyaev-Zeldovich effect observed at millimeter wavelength, that is sensitive to the thermal pressure of the electron gas $P_{\rm gas}$. In combination with X-ray observations (providing $n_{\rm gas}$), the associated SZ mass profile is given by (see e.g., Muñoz-Echeverría et al. (2023))

$$M_{\rm SZ}(< r) = \frac{1}{\mu m_p G}\frac{r^2}{n_{\rm gas}(r)}\frac{\partial P_{\rm gas}(r)}{\partial r}. \tag{2.23}$$

However, HSE masses are generally "biased" since we assume the cluster to be at hydro-static equilibrium. The link to the true mass is generally described by the HSE bias $b_{\rm HSE}$ such as $M_{\rm HSE} = (1 - b_{\rm HSE})M_{\rm true}$, where $M_{\rm true}$ is generally considered to be the lensing mass (it is generally accepted that $b_{\rm WL} \approx 0$, again, we will address this point in Section 2.4). The HSE bias is still a major challenge in cluster cosmology based on HSE masses, since HSE masses compared to weak lensing masses gives $1 - b_{\rm HSE} = 0.8$ (compatible with simulations) when the fitted HSE bias from the joint analysis of *Planck* CMB anisotropies and cluster count give $1-b_{\rm HSE} \approx 0.6$ (see the recap from Douspis et al. (2019)). Thanks to possible sky overlap between a variety of surveys, cluster masses can be determined with high precision by combining multi-wavelength follow-up observations carrying independent systematic effects, e.g., weak lensing follow-up observations of sub-sample of clusters can be used as a prior for analysis based on HSE mass estimates (e.g., *Planck*-XVII, SPT-H16/B19, WtG, *eROSITA*, see Table 2.1). For instance, the SPT-B19 analysis used HSC weak gravitational lensing data for 32 of their 343 SZ-detected clusters and *Chandra* X-ray follow-up observations for 82 clusters.

[3]Galaxies are bound by gravity, i.e., they trace the cluster gravitational potential. The cluster mass can be obtained by applying the Jeans equation for a steady-state spherical system (at virial equilibrium).

So far we have discussed the inference of cluster masses. For richness-detected clusters (in optics, e.g., DES, KiDs or SDSS) that will be the focus of this thesis, the cluster mass-proxy relation can be obtained "internally" from cluster weak lensing measurements (e.g., see Simet et al. (2017) on SDSS data, or Melchior et al. (2017); McClintock et al. (2019) on DES data, Murray et al. (2022) with HSC data). Otherwise, it can be constrained jointly to cosmological parameters combining the cluster abundance and the mean mass of clusters within richness-redshift bins considering the likelihood

$$\mathcal{L}_{\rm tot} = \mathcal{L}_{\rm CA}(\{N_i\}|\vec{\theta}) \times \mathcal{L}_{\rm MM}(\{M_i\}|\vec{\theta}) \tag{2.24}$$

where $\mathcal{L}_{\rm MM}$ is the cluster mean mass likelihood which relates to the cosmological and proxy-mass relation parameters via the expected mass

$$M_i(\vec{\theta}) = \frac{1}{N_i}\int_{z_1}^{z_2} dz \int_{\lambda_1}^{\lambda_2} d\lambda \int_{m_{\rm lim}}^{+\infty} dm\ m \frac{d^3N(\lambda, m, z)}{d\lambda dm dz}. \tag{2.25}$$

where the mean masses $\{M_i\}$ within richness-redshift bins are inferred from weak gravitational lensing produced by an ensemble of clusters within each richness-redshift bin (as it is done in DES, KiDs and SDSS cluster analyses).

2.3.3 Experimental landscape

We gather in Table 2.1 some details about the past, ongoing and future cosmological analyses based on the abundance of clusters detected at millimeter wavelength (ACT, *Planck*, SPT), optical/near-IR (DES, KiDS, SDSS) or X-ray (XMM-XXL, *eROSITA*[4], WtG). The first column corresponds to different surveys/experiments from which the cluster catalogs are obtained[5]. The second column lists the different analyses; for a given survey, the analyses may differ from one another[6] by e.g., the methodology, the combination to other probes/wavelengths, the amount of data since the start of the experiment/survey, etc. The last column corresponds to the number of clusters that are (or will be) used in the different analyses, this number depends on the lowest mass of the cluster sample, the detection efficiency and the properties of proxy limiting the redshift spanned by the survey (Weinberg et al., 2013).

So far we have focused on the methodology used in many cluster abundance cosmological analyses, especially addressing the problem (and solutions) of the cluster mass calibration for precision cosmology in the previous Section 2.3.2. We present in Fig. 2.3 (left panel) some constraints on the S_8 parameter obtained from cluster-based analyses listed in Table 2.1, here SDSS-A20, WtG, ACT-H13, SPT-B19, *Planck*-XXIV and DES-Y1 (in red). The 1σ compatibility between SDSS (optical), SPT, *Planck* (millimeter) and WtG (X-ray) constraints demonstrate the consistency between cluster analyses at different wavelengths. Complementary to single-probe analysis, cluster abundance can also be combined with other probes to break degeneracy

[4]The work by Chiu *et al.* is a preliminary analysis on a reduced portion of the sky carried by the *eROSITA* Final Equatorial Depth Survey (eFEDS) that comprised 455 clusters after selection. The *eROSITA* full sky survey sample will contain $\sim 10^5$ X-ray clusters.

[5]The exceptions are, first, the *Planck*×SPT analysis, where Salvati *et al.* constrained cosmology with the abundance of 439 *Planck* clusters with $0 < z < 1$ (full-sky observations) and 343 SPT clusters between $0.25 < z < 1$ over 2,500 deg^2 of the south sky. The second exception is the DES×SPT analysis, where Costanzi *et al.* used the count of $\sim$ 7000 redMaPPer clusters from DES between $0.2 < z < 0.65$, and 141 SPT clusters with $z > 0.65$ to constrain cosmology (a subsample of SPT clusters with $z < 0.65$ is matched to DES optical clusters to improve the constraints of the mass-observable relations).

[6]For instance, the SDSS-C19 combined the abundance with stacked lensing masses, while the SDSS-F23 analysis with stacked lensing profiles directly.

Survey	Analysis	Ω_S	$z_{\rm min}$ - $z_{\rm max}$	$N_{\rm tot}$
ROSAT	WtG, Mantz et al. (2015)	400	0 - 0.5	224
XMM-XXL	Pacaud et al. (2018)	50	0.05 - 1	178
**eROSITA*	eFEDS, Chiu et al. (2023)	140	0.1 - 1.2	455
ACT	S11, Sehgal et al. (2011)	455	0.16 - 1	9
	H13, Hasselfield et al. (2013)	504	0.15 - 0.8	15
Planck	XX, Ade et al. (2014)	26,000	0 - 1	189
	XXIV, Ade et al. (2016)	26,000	0 - 1	439
	×SPT, Salvati et al. (2022)		0 - 1	782
SPT	H16, de Haan et al. (2016)	2,500	0.25 - 1.7	377
	B19, Bocquet et al. (2019)	2,500	0.25 - 1.75	343
	C22, Chaubal et al. (2022)	2,500	0.25 - 1.75	343
*CMB-S4	$>$ 2029	27,000	0 - 3	$> 10^5$
*Simons Obs.	$>$ 2024	16,000	0 - 2	$> 2.10^4$
SDSS	R10, Rozo et al. (2010)	7,398	0.1 - 0.3	10,810
	M13, Mana et al. (2013)	7,500	0.1 - 0.3	13,823
	C19, Costanzi et al. (2019)	11,000	0.1 - 0.3	7,000
	A20, Abdullah et al. (2020)	11,000	0.045 - 0.125	756
	P23, Park et al. (2023)	11,000	0.1 - 0.33	8,379
	S23, Sunayama et al. (2023)	11,000	0.1 - 0.33	8,379
	F23, Fumagalli et al. (2023)	11,000	0.1 - 0.3	6,964
DES	Y1, Abbott et al. (2020)	1,800	0.2 - 0.65	7,000
	×SPT, Costanzi et al. (2021)			7,000
KiDS	DR3, Lesci et al. (2022)	377	0.1 - 0.6	3,652
**Euclid*	$>$ 2024	15,000	0.2 - 2	$> 2.10^5$
*LSST	$>$ 2024	18,000	0.1 - 1.2	$> 3.10^5$
**Roman*	$>$ 2026	2,000	0 - 2	$> 4.10^4$

Table 2.1: Overview of the past (starting from $\sim$ 2010), ongoing and future surveys and corresponding abundance-based cosmological studies. The first column corresponds to the different surveys and the different cosmological analyses are listed in the second column.The third column is the sky area of the cluster survey/catalog (in deg^2), the fourth column is the redshift range of the cluster catalog, and the last column gives the total number of clusters used in the analysis (or that will be used). The analyses are sorted following the wavelength sensitivity of the corresponding experiment. X-ray surveys are in the first horizontal block, millimeter surveys in the second block and optical surveys in the third block. Future surveys are denoted by *.

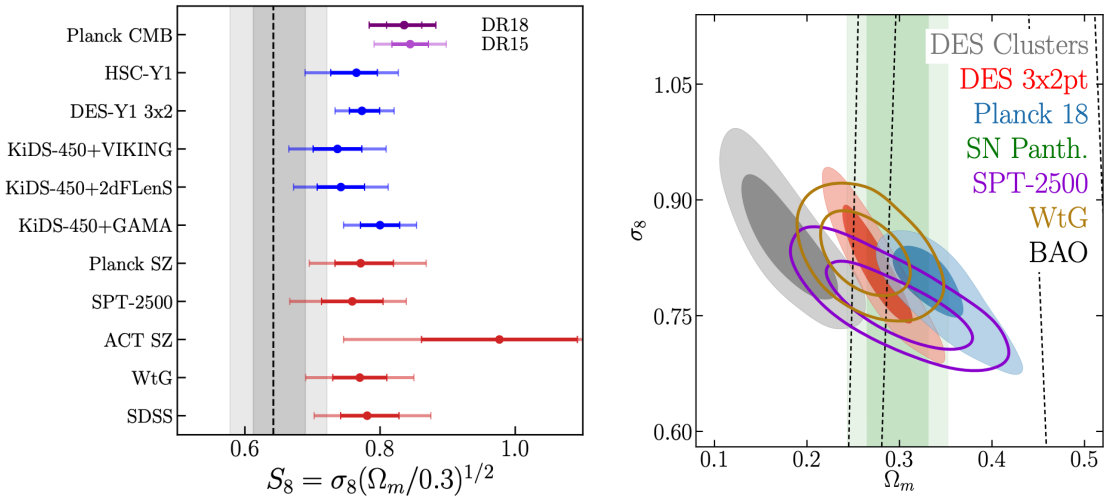

Figure 2.3: Left: S_8 constraints for a variety of cosmological analyses. DES-Y1 cluster results are represented by the grey shaded region, showing that DES-Y1 cluster analysis has a significantly lower value of S_8 than other probes. Abundance-based analyses are colored in red. Blue results are obtained from 3×2pt analysis, and CMB *Planck* results are in purple. Right: Joint posterior distribution of Ω_m and σ_8 for different cosmological analyses. DES 3×2pt and cluster analysis are respectively in red and grey, *Planck* CMB constraints are in blue. Both figures from Abbott et al. (2020).

between parameters and improve constraints, e.g., the joint fit of SPT cluster abundance and *Planck* CMB anisotropies decreases the errors on S_8 by 20% in $\nu\Lambda$CDM cosmology (cosmology with massive neutrinos) by notably breaking the degeneracy between neutrino masses and σ_8. Still, in Fig. 2.3 (left panel), cluster analyses are also consistent at $\sim 1\sigma$ with all 3×2pt results represented in blue. However, low-redshift probes (clusters + 3×2pt) prefer lower values of S_8 (except for ACT clusters, which preferred higher value of S_8) compared to CMB constraints in purple at the $\sim 1.5/2.5\sigma$ levels (where ACT constraints are 1σ consistent), known as the S_8 tension, that we already discussed in Section 1.5. So far it is hard to draw a global portrait of cluster abundance cosmology since various inference setups are usually considered for a single dataset, as well as various combinations with other probes to gain higher precision and overcome calibration issues.

As part of these different abundance analyses, we now present and discuss more in detail the DES-Y1 results that were published at the beginning of my thesis, and were in severe disagreement with other studies. The recent DES-Y1 cluster analysis (Abbott et al., 2020), combining cluster abundance and *stacked* cluster lensing masses, found significantly lower value of $S_8 = 0.65 \pm 0.04$ to that of other probes, represented by the gray shaded region in Fig. 2.3 (left panel), and in particular a 5.6σ tension with *Planck* CMB results on the joint $\sigma_8 - \Omega_m$ posterior, plus a 2.4σ tension with their 3×2pt analysis (see Fig. 2.3, right panel). They found that their low mass clusters (or equivalently low richness clusters with $\lambda < 30$) are the main drivers of the tension since they have a much smaller lensing signal than expected. This leads to an incorrect interpretation of stacked lensing data (again, see Section 2.4), that affect significantly the constraint of the mean cluster mass for low richness values. Removing the low richness clusters enabled to solve the internal inconsistency tension between their 3×2pt and cluster analyses, but increased significantly the errors on the parameters. Because of the size of the discrepancy, the authors do not argue on possible new cosmological phenomena beyond the

CDM paradigm or strong observational systematic that should explain this tension, but insist on a possible un-modeled weak lensing systematic due to projection and triaxiality effects, as well as selection effect bias (about 13% of the error budget in the weak lensing mass calibration), that describes the trend of the cluster finder to detect preferentially clusters of galaxies with properties (here the richness) that are correlated with their weak lensing signal at fixed mass (for instance elongated along the line of sight). This effect appears to significantly impact the lensing signal of low mass clusters, and after having tested on mocks, they found that it lowers mass estimates by $20 \sim 30\%$. The surprising DES results emphasize the complexity of using weak gravitational lensing for mass calibration with photometric selected clusters and accounting for these systematic is crucial for DES-Y3 and other surveys that use lensing mass calibration[7]. It is one of the projects I am contributing to in the context of the DESC (see Chapter 4).

The next-generation surveys (mm: CMB-S4, Simons Observatory, SPT-3G, AdvACT, optical: Rubin LSST, *Euclid*) will benefit from very large cluster surveys, and cosmological analysis will benefit from very large statistical power to decrease errors on cosmological parameters (see Table 2.1, second block, where the number of clusters will gain one or two orders of magnitude compared to past and ongoing studies). Thanks to these high statistics and high-quality data for cluster mass calibration at different wavelengths, wide field and deep cluster surveys will play a crucial role in improving the precision of the cluster mass-proxy relation. To extract most of the statistical power of future surveys, minute handling of systematics is necessary.

2.4 Weak gravitational lensing by galaxy clusters

As mentioned before, the calibration of the cluster scaling relation can be performed using weak lensing. Here, we present the basis of weak gravitational lensing by galaxy clusters. We present in Section 2.4.1 the theoretical framework to predict the lensing signal around clusters, and we present in Section 2.4.2 the statistical methods that are generally employed to gain insights on cosmological parameters and mass-related properties from stacked cluster lensing measurements.

2.4.1 Predicting the stack lensing signal

In this section, we first introduce the formalism of the *effective* weak lensing shear produced by an ensemble of clusters, following a similar "ensemble" approach in redshift-proxy bins as presented in Section 2.3.1. As introduced in Section 1.4.3, the observed shapes of galaxies are unique tracers of the shear and convergence fields in the Universe. We recall the equation that links the observed galaxy ellipticity to its intrinsic shape and the local shear and convergence of the foreground projected matter distribution along the line-of-sight, given by

$$\epsilon^{\rm obs} = \frac{\epsilon^{\rm int} + g}{1 + g^* \epsilon^{\rm int}}, \quad \text{where} \quad g = \frac{\gamma}{1-\kappa} \tag{2.26}$$

is the reduced shear, and $\epsilon^{\rm int}$ is the galaxy intrinsic ellipticity, the one observed without lensing distortion. In the weak lensing regime, the convergence $\kappa \ll 1$ and the shear $\gamma \ll 1$, so $\epsilon^{\rm obs} \approx \epsilon^{\rm int} + \gamma$. If the intrinsic ellipticities have no particular orientation, then the *average*

[7] In particular, the DES×SPT analysis by Costanzi *et al.* showed that combining DES-Y1 abundance and SPT multi-wavelength data (to constrain both cosmology and the mass-observable relations) yield consistent cosmological results to other probes. This work supports the conclusion of the DES-Y1 analysis which suggests the presence of unknown systematics in the modeling of the stacked weak lensing signal of low-richness clusters, impacting the reconstruction of cosmological parameters when calibrating the mass-richness relation with mass estimates derived from DES stacked weak lensing profiles.

observed ellipticity at a given location in the sky is an unbiased estimator of the shear that verifies

$$\langle \epsilon^{\rm obs} \rangle \approx \gamma. \tag{2.27}$$

The weak lensing effect (produced by the combination of γ and κ, only γ in the weak lensing regime) can be measured on galaxies located in the background of clusters, that represent the densest matter density peaks along the path of a photon coming from distant sources. In practice, we express the two degrees of freedom $\gamma_1(\vec{\theta})$ and $\gamma_2(\vec{\theta})$ in the tangential/cross reference frame, centered on a given arbitrary location $\vec{\theta}_0$, that is usually taken to be the center of a galaxy cluster. Moreover, noting that any location in the sky $\vec{\theta}'$ can be decomposed $\vec{\theta}' = \vec{\theta}_0 + \vec{\theta}$, where the offset separation $\vec{\theta}$ is defined by its modulus $\theta = |\vec{\theta}|$ and its polar angle ϕ, we can define the *tangential* and *cross* shear such as

$$\gamma_+(\vec{\theta}_0 + \vec{\theta}|\vec{\theta}_0) = -{\rm Re}[\gamma {\rm e}^{-i2\phi}] = -\gamma_1 \cos(2\phi) - \gamma_2 \sin(2\phi), \tag{2.28}$$

$$\gamma_\times(\vec{\theta}_0 + \vec{\theta}|\vec{\theta}_0) = -{\rm Im}[\gamma {\rm e}^{-i2\phi}] = -\gamma_2 \cos(2\phi) + \gamma_1 \sin(2\phi). \tag{2.29}$$

Now, let's derive the *effective* lensing shear around a set of clusters, for simplicity, we consider an ensemble of clusters with known masses, since the derivation for the shear profile for proxy-selected clusters is straightforward and will be explored in Chapter 4. Consider a population of clusters with redshifts $z \in [z_1, z_2]$ and masses $m \in [m_1, m_2]$. The total number of clusters is given by

$$N_{\rm tot} = \int_{z_1}^{z_2} dz \int_{m_1}^{m_2} dm \, \frac{d^2N(m,z)}{dmdz}. \tag{2.30}$$

The above equation is equivalent to Eq. (2.16) but only accounting for mass and redshift bins. The cluster angular number density is given by $n_h(\vec{\theta}) = \bar{n}_h(1 + \delta_h(\vec{\theta}))$, where $\bar{n}_h = N_{\rm tot}/\Omega_S$, and $\delta_h(\vec{\theta})$ is the projected overdensity of clusters at location $\vec{\theta}$. From Jeong et al. (2009), the average tangential/cross shear is given by the cross-correlation between the cosmological shear field with the cluster spatial distribution

$$\langle \gamma^h_{+/\times} \rangle(\vec{\theta}) = \langle \delta_h(\vec{\theta}_0) \gamma_{+/\times}(\vec{\theta}_0 + \vec{\theta}|\vec{\theta}_0) \rangle. \tag{2.31}$$

In the above equation, switching γ to κ in Fourier space (see Appendix B) gives explicitly the correlation between the cluster overdensity δ_h and the convergence field κ. Making the azimuth average along the polar angle ϕ of the separation $\vec{\theta}$, we find (see the full derivation in Jeong et al. (2009))

$$\langle \gamma^h_+ \rangle(\theta) = \frac{1}{N_{\rm tot}} \int_{z_1}^{z_2} dz \int_{m_1}^{m_2} dm \, \frac{d^2N(m,z)}{dmdz} \langle \Sigma^{-1}_{\rm crit}(z) \rangle \Delta\Sigma(\theta, z|m) \tag{2.32}$$

$$\langle \gamma^h_\times \rangle(\theta) = 0. \tag{2.33}$$

The averaged cross shear $\langle \gamma^h_\times \rangle$ is always zero for any type of mass distribution (see Appendix K). Second, for the average tangential shear $\langle \gamma^h_+ \rangle$, the different terms that appear in the expression are as follows;

The halo mass and redshift distribution $d^2N/dmdz$ has already been discussed in Section 2.2 and is given in terms of the halo mass function, the comoving volume as well as the sky area.

The lensing efficiency is also known as the inverse effective critical surface mass density and is given by

$$\langle \Sigma^{-1}_{\rm crit}(z) \rangle = \begin{cases} \int_z^{z_{\rm lim}} dz_s p_s(z_s) \Sigma^{-1}_{\rm crit}(z_s, z) & \text{if } z_s > z, \\ 0 & \text{if } z_s < z. \end{cases} \tag{2.34}$$

where $p_s(z)$ is the redshift probability density function of background source galaxies and $\Sigma_{\rm crit}^{-1}(z_s, z)$ (see Appendix B) is the (inverse) critical surface density given by

$$\Sigma_{\rm crit}(z_s > z, z) = \frac{c^2}{4\pi G}\frac{\chi(r_s)}{\chi(r)\chi(r_s - r)}\frac{1}{1+z}. \tag{2.35}$$

The expression ensures that lenses at redshift z only affect the shapes of galaxies located in their background, i.e., $z_s > z$, and the lensing *strength* depends on the relative distance between the lens, the source and the observer, encoded in $\langle\Sigma_{\rm crit}^{-1}(z)\rangle$.

The excess surface density: At fixed cluster mass, redshift and source population, the excess surface density describes how the properties of the local mass distribution at cluster scales "shape" the lensing effect on background images (whereas the lensing efficiency tells how "strong" it is with respect to the geometrical layout of the lens-source pair). It is linked to the halo-matter power spectrum $P_{\rm hm}(k|z, m)$ at mass m and in the flat sky approximation, it is given by (Jeong et al., 2009)

$$\Delta\Sigma(\theta, z|m) = \frac{\rho_m}{\chi(z)^2}\int \frac{ldl}{2\pi} P_{\rm hm}(k_l(z), z|m) J_2(l\theta), \tag{2.36}$$

where $k_l(z) = \frac{l+1/2}{\chi(z)}$, J_2 is the second order Bessel function and χ is the comoving transverse distance at redshift z, and ρ_m is the background matter density. $\Delta\Sigma(\theta, z|m)$ is usually expressed in terms of two more familiar physical quantities, namely

$$\Delta\Sigma(\theta, z|m) = \Sigma(<\theta, z|m) - \Sigma(\theta, z|m), \tag{2.37}$$

where $\Sigma(\theta, z|m)$ is the azimuth averaged *projected* matter density at angular separation θ from the cluster center, and $\Sigma(<\theta, z|m)$ is the mean projected density below θ. The matter density around cluster is given by $\rho(r|m) = \rho_m(1 + \xi_{\rm hm}(r|m))$ (see Section 2.2) where $\xi_{\rm hm}(r|m)$ is the halo-matter correlation function. The azimuth averaged *projected* surface density along the line-of-sight is defined by

$$\Sigma(\theta, z|m) = \int \frac{d\varphi}{2\pi}\Sigma(\vec{\theta}, z|m) \quad \text{where} \quad \Sigma(\vec{\theta}, z|m) = \int dy \rho(\vec{r}|m) \tag{2.38}$$

is the *local* surface mass density, calculated as the projection along the line-of-sight y of the three-dimensional matter density (here, $\vec{r} = (x, y, z)$ denotes the three-dimensional position). As for $\Delta\Sigma$, the quantity Σ can be related to the halo-matter power spectrum such as

$$\Sigma(\theta, z|m) = \frac{\rho_m^0}{\chi(z)^2}\int \frac{ldl}{(2\pi)} P_{\rm hm}(k_l(z), z|m) J_0(l\theta). \tag{2.39}$$

where J_0 is the zero-th order Bessel function (also obtained by changing J_2 to J_0 in the expression of $\Delta\Sigma$). Similarly, the mean projected density $\Sigma(<\theta, z|m)$ is defined as the mean of the comoving projected density within a circle with radius θ such as

$$\Sigma(<\theta, z|m) = \frac{2}{\theta^2}\int_0^\theta d\theta'\theta'\Sigma(\theta', z|m). \tag{2.40}$$

Moreover, from Eq. (2.37), $\Delta\Sigma(\theta, z|m)$ is linear in terms of the halo-matter power spectrum, i.e., if $P_{\rm hm} = P_1 + P_2$, then $\Delta\Sigma = \Delta\Sigma_1 + \Delta\Sigma_2$. Following Section 2.2, the halo-matter power spectrum can be decomposed into two different contributions. First the 1-halo term $P_{\rm hm}^{1h}$ (equivalently $\xi_{\rm hm}^{1h}$) and second the 2-halo term $P_{\rm hm}^{2h}$ (equivalently $\xi_{\rm hm}^{2h}$), which denote respectively the contribution

of the single halo matter density to the lensing effect, and the matter perturbations coming from the connected neighboring halos. For instance, the NFW (Navarro et al., 1997) or Einasto (Einasto, 1965) parameterizations are widely used to model the 1-halo term, i.e., how matter organizes at small scales independently to neighboring halos, when assuming that the single halo mass distribution is almost spherical.

However, we have assumed in Eq. (2.29) that we identify correctly the cluster center (i.e., its "true" value θ_0). The "observed" center may deviate from its "true" value by an offset separation $\vec{\theta}_{\rm off}$ (Oguri & Takada, 2011). In addition, all the clusters do not share the same offset angle (in both direction and modulus), so the offset modulus should follow a given probability function $p(\theta_{\rm off})$ depending on the quality of the detection, where we can reasonably assume that $p(\phi_{\rm off}) = {\rm cst}$ (the offset polar angle is randomly oriented). This leads to a modified excess surface density profile $\Delta\Sigma^{\rm off}_{\rm 1h}$ accounting for this off centering[8]. Consider that only a given fraction $f_{\rm cen}$ of clusters with mass m in the survey have "true" assigned centers, we have that

$$\Delta\Sigma(\theta, z|m) = f_{\rm cen}\Delta\Sigma^{\rm cen}_{1h}(\theta, z|m) + (1 - f_{\rm cen})\Delta\Sigma^{\rm off}_{1h}(\theta, z|m) + \Delta\Sigma_{2h}(\theta, z|m). \tag{2.42}$$

After re-scaling $\Delta\Sigma(\theta, z|m)$ by the inverse critical density (that has no radial dependency) and integrating over the mass-redshift distribution of halos, the resultant tangential shear in Eq. (2.32) follows similar decomposition. Another significant systematic is the cluster member dilution of the lensing signal. If cluster member galaxies are accounted for as part of the source (background) galaxies, the lensing signal is diluted since the member galaxies are not lensed. The dilution of the lensing profile by member galaxies is usually accounted for in the modeling by the boost factor $(1 - f_{\rm cl})$ where $f_{\rm cl}$ is the fraction of cluster members misidentified as being source galaxies[9], such that the measured profile has to be compared to $(1 - f_{\rm cl})\Delta\Sigma_{\rm th}$.

We represent in Fig. 2.4 the profile $\Delta\Sigma(R, z = 0.4|M_{\rm 200c} = 10^{14}\ M_\odot)$, where $R = D_A(z)\theta$ is the *physical* separation from the cluster center. We choose a NFW mass distribution with concentration $c_{\rm 200c} = 6$, the miscentering distribution is assumed to follow the Rayleigh distribution

$$p(\theta_{\rm off}) = \frac{\theta_{\rm off}}{\sigma^2_{\rm off}} \exp -\frac{1}{2}\left(\frac{\theta_{\rm off}}{\sigma_{\rm off}}\right)^2, \tag{2.43}$$

with $\sigma_{\rm off} = 0.013$ deg, and by choosing that the fraction of centered halos is $f_{\rm cen} = 0.8$. We see that the total profile (in red) is mostly affected by the mis-centering correction (in blue dashed line) and boost factor correction for $R \leq 1$ Mpc, that smooths the lensing signal to lower values compared to the well-centered profile with no contamination (blue line). At radius $R \geq 5 - 6$ Mpc, the profile is dominated by the 2-halo term (in orange) that boosts the lensing due to correlated neighboring structures that deform galaxy shapes stronger than the single halo for $R > 10$ Mpc.

Several other effects impact cluster lensing, and will affect the lensing signal, and should ideally be taken into account. For instance, the NFW or Einasto models are calibrated on N-body simulations with dark matter only particles, so they do not account for the impact of baryons that are usually confined in the innermost regions and contribute to boosting the lensing signal at low radius (Lee et al., 2018; Cromer et al., 2022). Moreover, in our example we only consider spherical profiles (off-centered or correctly centered), whereas the cluster matter distribution has, in general, a more complex triaxial shape (Stapelberg et al., 2022) and this

[8]The off-centered $\Delta\Sigma$ is derived from the corrected off-centered halo-matter power spectrum $P^{1h}_{\rm hm,off}(k_l(z), z|m)$ that is given by

$$P^{1h}_{\rm hm,off}(k_l(z), z|m) = P^{1h}_{\rm hm}(k_l(z), z|m) \int_0^{+\infty} d\theta_{\rm off} p(\theta_{\rm off}) J_0(l\theta_{\rm off}). \tag{2.41}$$

[9]The boost factor is radially dependent since it is more likely to misidentify cluster members as background sources close to the cluster center rather than at larger radius.

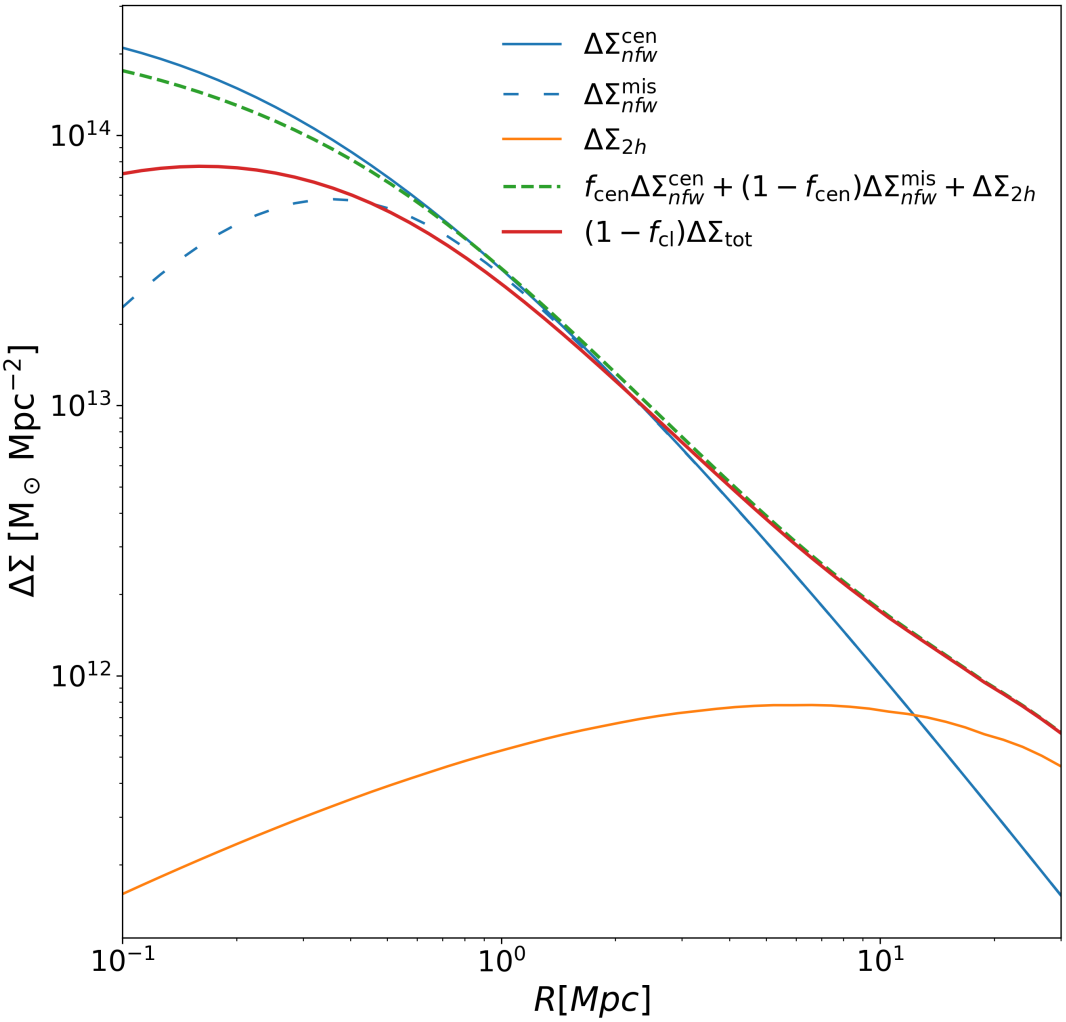

Figure 2.4: The different contributions to the excess surface density profile at fixed mass. The correctly centered NFW 1-halo profile is in blue, and the off-centered profile is represented by the dashed blue line. The 2-halo term is in orange. The total excess surface density, without accounting for the boost factor correction is in green dashed line (we used a power law boost model used in Melchior et al. (2017)). The corrected total profile is represented in red.

needs to be accurately accounted for precision cluster lensing studies (Zhang et al., 2023; Feroz & Hobson, 2011)[10], as well as the presence of substructures in and around clusters (Clowe et al., 2004).

2.4.2 Measuring stacked lensing signal

We presented how the lensing signal around an ensemble of clusters relates to the underlying mass density of clusters, from a modeling point of view. We address in this section how to properly estimate this lensing effect from real data.

For a cluster (denoted by the letter "l" for lens) at redshift z_l, we first identify a catalog of background source "s" galaxies ($z_l < z_s$). For each source galaxy, we define the tangential ϵ_+ and cross $\epsilon_\times$ ellipticities with respect to the cluster center angular position

$$\epsilon_+ = -\mathrm{Re}[\epsilon e^{-i2\phi}] = -\epsilon_1 \cos(2\phi) - \epsilon_2 \sin(2\phi), \tag{2.44}$$

$$\epsilon_\times = -\mathrm{Im}[\epsilon e^{-i2\phi}] = -\epsilon_2 \cos(2\phi) + \epsilon_1 \sin(2\phi), \tag{2.45}$$

where ϕ is the polar angle from the l-cluster center to its ls-background source galaxy. So, for a cluster at redshift z_l, we can identify all galaxies located in its background within an annulus comprised between $R - \Delta R/2$ and $R + \Delta R/2$ with $R = \theta D_A(z_l)$ (i.e., the physical projected separation from the cluster center in Mpc). Let's consider N_l clusters with redshifts $z \in [z_1, z_2]$ and richness $\lambda \in [\lambda_1, \lambda_2]$. For each l-lens, we identify the N_{ls} background galaxies[11] within the ΔR region. The estimator of the *stacked* excess surface density for the N_l clusters is (Sheldon et al., 2004; Shirasaki & Takada, 2018)

$$\widehat{\Delta\Sigma}(R) = \frac{1}{N_l} \sum_{l=1}^{N_l} \sum_{s=1}^{N_{ls}} \bar{w}_{ls} \Sigma_{\mathrm{crit},ls} \epsilon_{ls}^{+} \quad \text{where} \quad \sum_{s=1}^{N_{ls}} \bar{w}_{ls} = 1. \tag{2.46}$$

The weights $\bar{w}_{ls}$ and $\Sigma_{\mathrm{crit},ls}$ both account for possible sources of systematics affecting the measure of galaxy ellipticities and redshifts. Galaxy shapes and redshifts are measured from their photometric properties, i.e., from the spectral and spatial properties of their observed flux in different optical bands (we will see more in detail these two aspects in Chapter 4). Shapes and redshifts of galaxies are then measured with some bias and scatter around their true value due to a variety of systematic effects, since the observed flux is affected by the light absorption by the atmosphere (for ground-base telescope), the noise from the telescope, etc. Then, the redshift of a background galaxy is not *perfectly* known,. Instead of a single value z_s, the redshift of a galaxy is described by a photometric probability density function $p_{\mathrm{photoz}}(z_s)$. For simplicity[12], we assume here that the error related to the shape measurement method is simply given by $\sigma_{\mathrm{meas}}(\epsilon_{ls}^{+})$. The quantity $\Sigma_{\mathrm{crit},ls}$ is the photometric *effective* critical surface mass density of the lens-source system, averaged over the distribution $p_{\mathrm{photoz,s}}(z_s)$ of the galaxy with index s, such as

$$\Sigma_{\mathrm{crit},ls}^{-1} = \int_{z_l}^{+\infty} dz_s \; p_{\mathrm{photoz}}(z_s) \Sigma_{\mathrm{crit}}(z_s, z_l)^{-1}. \tag{2.47}$$

In Eq. (2.46), the weights $\bar{w}_{ls}$ are chosen to maximize the signal-to-noise ratio for the estimator (as shown in Sheldon et al. (2004)) and can be written as the product of two terms, namely

[10] We will address this point for elliptical galaxy clusters in Chapter 6.

[11] The misidentification of foreground galaxies as background galaxies yields to an attenuation of the signal as mentioned before.

[12] In reality, shape measurements are biased and need to be calibrated (see Chapter 4).

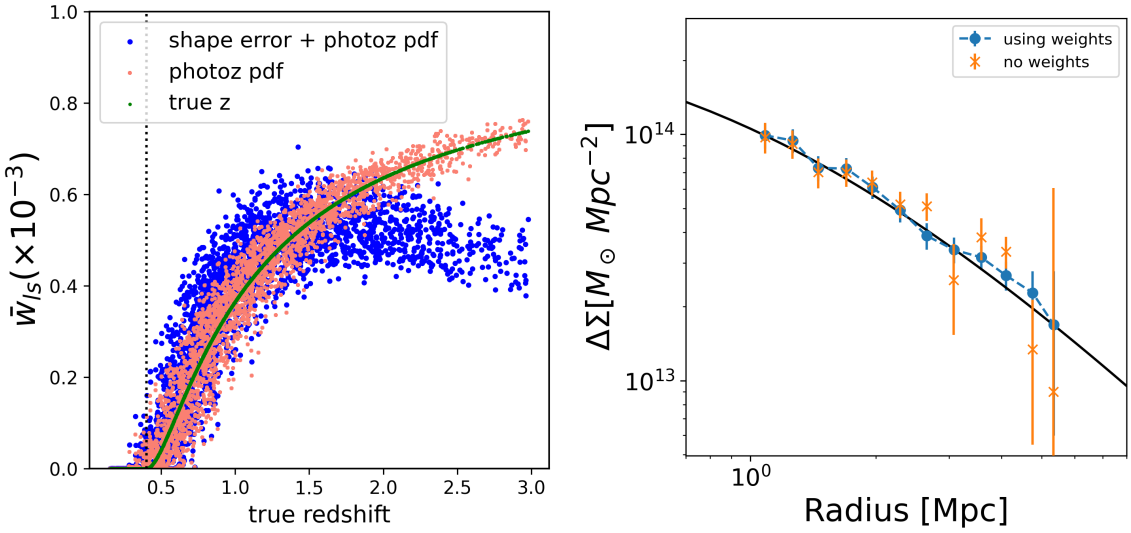

Figure 2.5: Left: lensing weights as a function of the true galaxy redshifts. the case with no shape errors and perfect redshift is shown in green, with no shape errors and photometric redshifts in red, with shape errors and photometric redshifts in blue. Right: Excess surface density profile computed without using (orange) and using (blue) lensing weights. The fiducial halo model is represented in black full lines.

$w_{ls} = w_{ls}^{\rm geo} w_{ls}^{\rm shape}$ where

$$w_{ls}^{\rm geo} = \Sigma_{{\rm crit},ls}^{-2}, \tag{2.48}$$

$$w_{ls}^{\rm shape} = \frac{1}{\sigma_{\rm rms}^2(\epsilon_s^+) + \sigma_{\rm meas}^2(\epsilon_s^+)}, \tag{2.49}$$

and where the normalized weight $\bar{w}_{ls}$ is given by

$$\bar{w}_{ls} = \frac{w_{ls}}{\sum_{s=1}^{N_{ls}} w_{ls}}. \tag{2.50}$$

The quantity $\sigma_{\rm rms}(\epsilon_s^+)$ is the standard deviation of the tangential component of the ellipticity, whereas $\sigma_{\rm meas}(\epsilon_s^+)$ denotes the error on shape measurement. In the ideal case where galaxy redshifts and shapes are perfectly known, these weights reduce to $w_{ls}^{\rm geo} = \Sigma_{{\rm crit},ls}^{-2}$ and $w_{ls}^{\rm shape} = 1/\sigma_{\rm rms}^2(\epsilon_s^+)$. Finally, we can re-write the stacked excess surface density in Eq. (2.46) such as

$$\widehat{\Delta\Sigma}(R) = \frac{1}{\sum_{l=1}^{N_l} W_l} \sum_{l=1}^{N_l} W_l \widehat{\Delta\Sigma}_l(R), \tag{2.51}$$

where

$$W_l = \sum_{s=1}^{N_{ls}} w_{ls} \quad \text{and} \quad \widehat{\Delta\Sigma}_l(R) = \sum_{s=1}^{N_{ls}} \bar{w}_{ls} \Sigma_{{\rm crit},ls} \epsilon_{ls}^+ \tag{2.52}$$

is the *per cluster* weak lensing profile. In this formalism, the stacked excess surface density appears as a weighted mean of individual lensing profiles. As an example, we show in Fig. 2.5 (left) the cluster-galaxy lensing weights $\bar{w}_{ls}$ for a mock cluster at redshift $z = 0.4$ (denoted by the vertical dashed line) and a population of background galaxies. We gradually increase the realism

of the mock data. We start with ideal weights (green points) when galaxy redshifts and shapes are perfectly known (no photometric redshift and no shape measurement error) so they appear as a continuous line. In that context, the weights increase with the redshift, meaning distant galaxies contribute more to the estimated lensing signal (they are more lensed than galaxies located closer to the cluster). When considering imperfect photometric redshifts (in red), the weights become scattered around their true values. This demonstration was made with still idealistic mock data. More effects need to be taken into account. As for the photometric redshifts, shapes are measured with some error, that usually increases with redshift (it is more difficult to measure the shape of distant galaxies compared to closer ones). The weights accounting for both photometric redshift and shape measurement are represented in blue. Even if distant galaxies are more lensed, their measured shapes have larger errors, and then their weights are lowered compared to their ideal values. To show the importance of the lensing weights in the lensing shear measurement, we show in Fig. 2.5 (right) the estimated excess surface density profile for a single cluster with and without accounting for the weights (i.e., computing the standard average of the individual estimators $\Sigma_{\rm crit,ls}\epsilon_{ls}^{+}$ in radial bins, versus using Eq. (2.52)). We see that accounting for lensing weights makes the profile less scattered than the weight-free profile. So lensing weights are crucial to maximize the signal-to-noise ratio for precise lensing measurement.

2.4.3 Covariance of the stacked lensing signal

We have seen in Section 2.4.1 that the lensing signal can be used to trace the underlying matter density field around the cluster. This signal can be used in a variety of ways, i.e., from "simply" estimating the effective cluster mass in a stack to more complex analyses, containing jointly the parameters of the cluster scaling relation and the cosmology. Both will be developed in Chapter 4. To infer information about some parameters $\vec{\Theta}$ that appear in the prediction of the stacked lensing profile (e.g., the richness-mass relation, the cosmological parameters, the halo properties, etc.), we use once again the Bayes theorem such that the parameter posterior is given by

$$\mathcal{P}(\vec{\theta}|\widehat{\Delta\Sigma}) \propto \mathcal{L}(\widehat{\Delta\Sigma}|\vec{\theta})\pi(\vec{\theta}), \tag{2.53}$$

where $\mathcal{L}$ is the cluster lensing *likelihood* function, as the probability for observing the data $\widehat{\Delta\Sigma}$ knowing that the parameters $\vec{\Theta} = \vec{\theta}$. Generally, the cluster lensing likelihood is taken to be a multivariate Gaussian probability distribution, that reduces to

$$\mathcal{L}(\widehat{\Delta\Sigma}|\vec{\theta}) \propto \exp -\frac{1}{2}\chi^2 \quad \text{where} \quad \chi^2 = (\Delta\Sigma_l(\vec{\theta}) - \widehat{\Delta\Sigma})^T \mathbf{C}^{-1}(\Delta\Sigma(\vec{\theta}) - \widehat{\Delta\Sigma}). \tag{2.54}$$

In this expression $\Delta\Sigma(R_i|\vec{\theta}) = \mathrm{E}_{\mathcal{L}}[\widehat{\Delta\Sigma}(R_i)|\vec{\theta}]$ is the *predicted* stacked excess surface density profile for parameters $\vec{\theta}$ (as introduced in Section 2.4.1, modeled with respect to the halo mass function, halo profiles, richness-mass relation, selection function, etc. that will be discussed more in details in Chapter 4). $\mathbf{C}$ is the covariance of the stacked profile given by

$$(\mathbf{C})_{ij} = \mathrm{E}_{\mathcal{L}}[(\widehat{\Delta\Sigma}_i - \langle\widehat{\Delta\Sigma}_i\rangle)(\widehat{\Delta\Sigma}_j - \langle\widehat{\Delta\Sigma}_j\rangle)], \tag{2.55}$$

where the average runs over all possible values of $\widehat{\Delta\Sigma}$ and $\widehat{\Delta\Sigma}_i = \widehat{\Delta\Sigma}(R_i)$. The scatter (diagonal element) and correlations (off-diagonal elements) of $\mathbf{C}$ originate from a variety of phenomena, and following the decomposition in Wu et al. (2019); Gruen et al. (2015),

$$\mathbf{C} = \mathbf{C}_{\rm shape} + \mathbf{C}_{\rm uncorr-lss} + \mathbf{C}_{\rm corr-lss} + \mathbf{C}_{\rm intr}. \tag{2.56}$$

First, $\mathbf{C}_{\rm shape}$ denotes the intrinsic scatter in the shape of background galaxies (shape noise[13]) as well as the limited sample of clusters and background galaxies (shot noise) and is the main

[13]Even if shapes are perfectly measured, the intrinsic dispersion of galaxy shape will always give non-zero variance

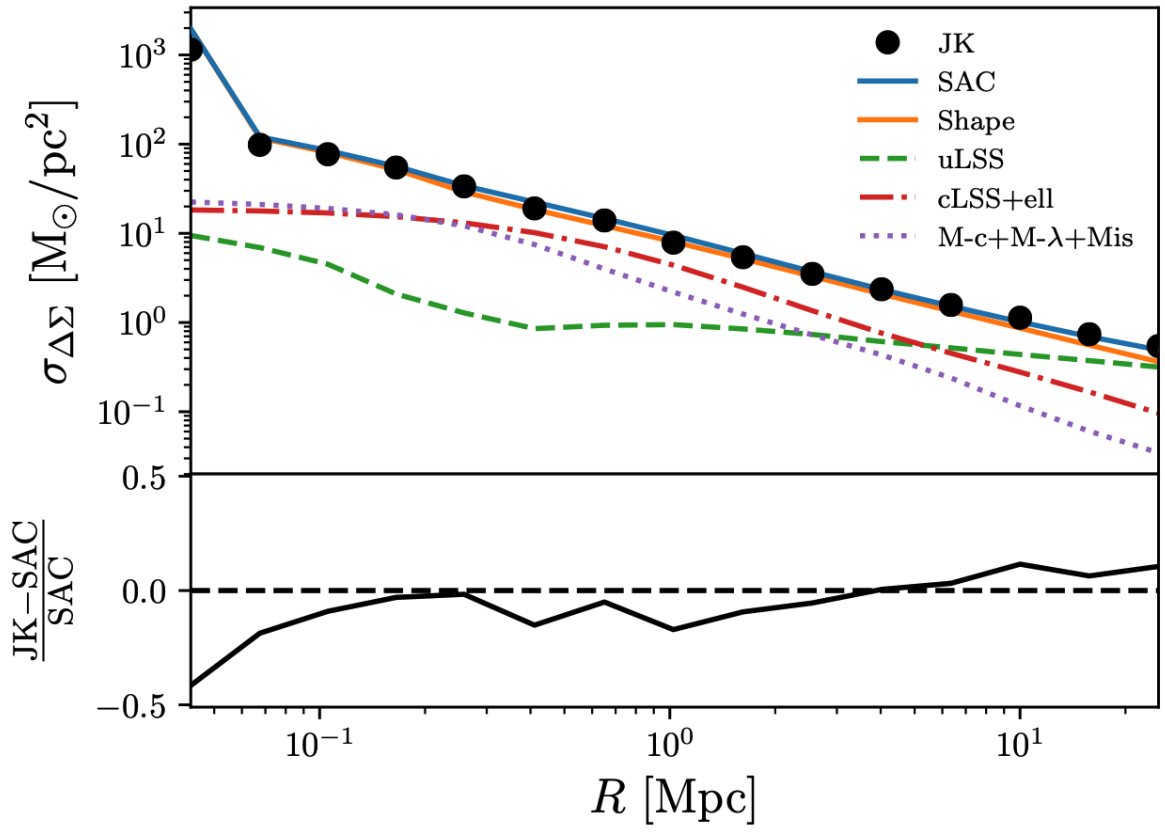

Figure 2.6: The different contributions to the variance of the stacked excess surface density. "Shape" denotes the contribution from the shape noise of galaxy shape. "uLSS" is for uncorrelated large-scale structures, "cLSS+ell" corresponds to the correlated large-scale structures and the intrinsic variation in the host halo ellipticity, "M-c+M-λ+Mis" stands for the variance in the mass-concentration relation, the richness-mass relation and for the mis-centering. The full variance is labeled with "SAC". Figure taken from McClintock et al. (2019).

source of scatter in the measure of $\Delta\Sigma$. The more clusters there are in the stack, the more background galaxies are used, and the less the variance of the stacked profile will be. All the other contributions introduce more variance and correlations between two radial bins, that drop rapidly by approximately an order of magnitude for large-scale separations and slower for small-scale separations (Wu et al., 2019). The $\mathbf{C}_{\rm uncorr-lss}$ term denotes the contribution from the uncorrelated large-scale structures, i.e., random structures along the line-of-sight (in addition to the detected cluster) that introduce a scatter in the measured convergence (then the shear) when it is measured in angular bins (see Hoekstra (2003)). It is, as for $\mathbf{C}_{\rm shape}$, independent of the intrinsic cluster properties. $\mathbf{C}_{\rm corr-lss}$ quantifies the impact of the stochastic variation of correlated halos around a cluster, associated to the variation of the 2-halo term, and depends on both matter power spectrum and halo bias. $\mathbf{C}_{\rm intr}$ tells the level of scattering of the lensing profile due to the variation in intrinsic halo properties, i.e. the variation of concentration at fixed mass, the halo ellipticity and orientation. It may also denote the scatter in individual masses of richness-selected clusters within the stack.

Fig. 2.6 shows the prediction[14] of each contribution to the variance of the stacked excess surface density profile for a stack of 762 redMaPPer-detected clusters[16] in DES (McClintock et al., 2019). The full variance is labeled with "SAC" and the black dots represent the variance estimated directly from the data, which we detail hereafter. From the figure, we can roughly say that the shape noise, as well as intrinsic variations of halo properties, will dominate the variance at small scales ($R \leq 1$ Mpc), the shape noise will contribute the most at middle scales and the LSS contribution starts dominating the variance at larger scales. The labels small/middle/large scale depend on the importance of shape noise, for instance, Wu et al. (2019) showed that[17] for $\bar{n}_s = 20$ arcmin^{-2}, the LSS contributions start dominating the lensing variance to radius larger $\sim 6-7$ Mpc, but this result also depends on the cluster redshift, cluster mass, etc. There are other sources of scatter minor to the ones already enumerated, e.g., the error induced by the presence of substructures in halos, the photometric redshift of background galaxies, super-sample covariance, etc.

In practice, all possible contributions to the covariance (even those that are not easy to derive analytically) can be captured using re-sampling techniques on real data, that we will consider in this thesis. Here below we give the details to compute the delete-1 jackknife covariance matrix (Shirasaki et al., 2017; Phriksee et al., 2020). First, we divide the sky area covered by the galaxy cluster survey sample in $N_{\rm JK}$ equal-sized sub-areas. From $k \in [1, N_{\rm JK}]$, we remove the clusters that belong to the k-th jackknife region and we compute $\widehat{\Delta\Sigma}_k$ profile. The jackknife covariance is given by

$$(\mathbf{C}^{\rm JK})_{i,j} = \frac{N_{\rm JK}-1}{N_{\rm JK}} \sum_{k=1}^{N_{\rm JK}} [\widehat{\Delta\Sigma}_k(R_i) - \overline{\Delta\Sigma}(R_i)][\widehat{\Delta\Sigma}_k(R_j) - \overline{\Delta\Sigma}(R_j)] \tag{2.57}$$

[14] In their paper, McClintock et al. (2019) used the "Semi-Analytical-Covariance" (SAC) technique presented in Gruen et al. (2015). They computed the "uLSS" and "cLSS+ell" terms assuming the halo model for the large-scale structure distribution (including the matter power spectrum, halo bias, and the halo mass function), as well as all possible variations in the intrinsic ellipticity of the host cluster, correlated with neighboring halos. Then, they used mock galaxy catalogs mimicking the scatter in the mass-concentration relation, the richness-mass relation, and the mis-centering for the "M-c+M-λ+Mis" term, and they obtained the "Shape" term by rotating galaxies and reevaluating the stacked excess surface density profile. In the literature, several methods for the prediction of the covariance have been used. For instance, it can be done analytically, by using the `cluster-lensing-cov`[15] package from Wu et al. (2019).

[16] With redshifts $z \in [0.2, 0.35]$ and richness $\lambda \in [20, 30]$.

[17] For HSC (covering 1400 deg^2), the typical source number density is $\bar{n}_s \sim 30$ arcmin^{-2}, higher than KiDS ($\bar{n}_s \sim 7$ arcmin^{-2}) and DES ($\bar{n}_s \sim 6$ arcmin^{-2}). For comparison, LSST will provide $\bar{n}_s \sim 35$ arcmin^{-2}.

with the mean

$$\overline{\Delta\Sigma}(R_i) = \frac{1}{N_{\rm JK}} \sum_{k=0}^{N_{\rm JK}} \Delta\Sigma^k(R_i). \tag{2.58}$$

The jackknife method allows us to measure all combined contributions to the variance based on real data in a simpler way than full derivation. For instance, see Fig. 2.6 that compares the diagonal elements of the jackknife variance estimate with 100 sub-regions to the SAC variance obtained in McClintock et al. (2019). However, the jackknife method introduces some noise in the estimation of the covariance matrix that adds un-physical bias to the recovered parameters, the level of "re-sampling noise" can be adjusted by choosing an appropriate number of sub-regions and even modifying the number d of deleted regions at each iteration, referred as the "delete-d" jackknife covariance matrix (Escoffier et al., 2016).

The bootstrap method is often used to estimate the cluster lensing covariance matrix (see e.g. Parroni et al. (2017)). For each stack, we create $N_{\rm boot}$ ensembles of clusters by selecting the same number of clusters but taking them with replacement. We compute $N_{\rm boot}$ times the stacked $\widehat{\Delta\Sigma}$ profile for each bootstrap resampling, we compute the covariance matrix

$$(\mathbf{C}^{\rm boot})_{i,j} = \frac{1}{N_{\rm boot}-1} \sum_{k=0}^{N_{\rm boot}} [\Delta\Sigma^k(R_i) - \overline{\Delta\Sigma}(R_i)][\Delta\Sigma^k(R_j) - \overline{\Delta\Sigma}(R_j)], \tag{2.59}$$

where

$$\overline{\Delta\Sigma}(R_i) = \frac{1}{N_{\rm boot}} \sum_{k=0}^{N_{\rm boot}} \Delta\Sigma^k(R_i). \tag{2.60}$$

Moreover, the *sample* covariance matrix (Phriksee et al., 2020) for a given stack of clusters with excess surface density profiles $\Delta\Sigma^k$ is defined by

$$(\mathbf{C}^{\rm sample})_{i,j} = \frac{1}{N_{\rm cluster}-1} \sum_{k=0}^{N_{\rm cluster}} [\Delta\Sigma^k(R_i) - \overline{\Delta\Sigma}(R_i)][\Delta\Sigma^k(R_j) - \overline{\Delta\Sigma}(R_j)], \tag{2.61}$$

where

$$\overline{\Delta\Sigma}(R_i) = \frac{1}{N_{\rm cluster}} \sum_{k=0}^{N_{\rm boot}} \Delta\Sigma^k(R_i). \tag{2.62}$$

We will compare these three approaches, namely the jackknife, the bootstrap and the sample covariance in Chapter 4.

2.5 Conclusions

Galaxy clusters are powerful tracers of the matter density field in the Universe, and their primary cosmological probe is their mass and redshift distribution that provides a wealth of information on the growth of structure and background cosmology. The performance of cluster abundance cosmological analysis mostly relies on the accuracy of the mass-proxy relation, which is the current limiting factor. To overcome this issue, the mass-proxy relation can be calibrated using cluster mass measurments. At optical wavelength, the weak gravitational lensing of galaxy clusters is a unique window to probe the mass distribution of galaxy clusters through the coherent deformation of the shape of galaxies located in their background. In the next chapters of this manuscript, we detail the different results in different aspects of cluster cosmology obtained in this thesis. We first present the preparation of the cosmological analysis with cluster abundance of the Vera Rubin LSST data in Chapters 3 and 4, in Chapter 5 we discuss many aspects of likelihoods for cluster abundance cosmology and we focus on a methodology to improve the constraint of weak lensing mass of individual massive clusters in Chapter 6.

Chapter 3

The Rubin observatory, the LSST and DESC

Contents

From the experimental landscape presented in Chapter 1, the Legacy Survey of Space and Time to be conducted by the Vera C. Rubin Observatory is arguably the most ambitious ground-based project in terms of scale and science objectives. In this chapter, we introduce the Rubin observatory and its main instrumental characteristics in Chapter 3. We present the LSST that will serve the community with a gigantic dataset to study a variety of scientific topics, ranging from the Solar system to cosmology. This is described in Section 3.2. The work presented in this thesis was pursued as part of the Dark Energy Science Collaboration (DESC), which is dedicated to the cosmological analysis and interpretation of the Rubin LSST data and that I present in Section 3.3.

3.1 The Vera C. Rubin Observatory

The Vera C. Rubin Observatory (named for the American astronomer Vera Florence Cooper Rubin) has been under construction since 2011 on Mount Cerro Pachon in Chile (see Fig. 3.1). The main building will host the Simonyi Survey Telescope. The main facility is complemented by the Rubin Auxiliary Telescope (AuxTel) that is installed a short distance away (in the background of Fig. 3.1). It will provide important complementary information such as atmospheric transmission to correct the Rubin LSST catalogs.

The lower part of the observatory in Fig. 3.1 holds the control and camera clean and the mirror coating rooms. Fig. 3.2 shows the inside of the 30-m dome (upper part of the observatory). The telescope mount, in the center of the image, has a compact design to slew very fast between the different pointing directions typically in 5 seconds. It holds a 3-mirror optical setup with an 8.4 meter diameter primary mirror, that covers a 9.6 $\deg^2$ field-of-view ($\sim$50 full moons). The effective focal length of the optical system is 10.31 meters.

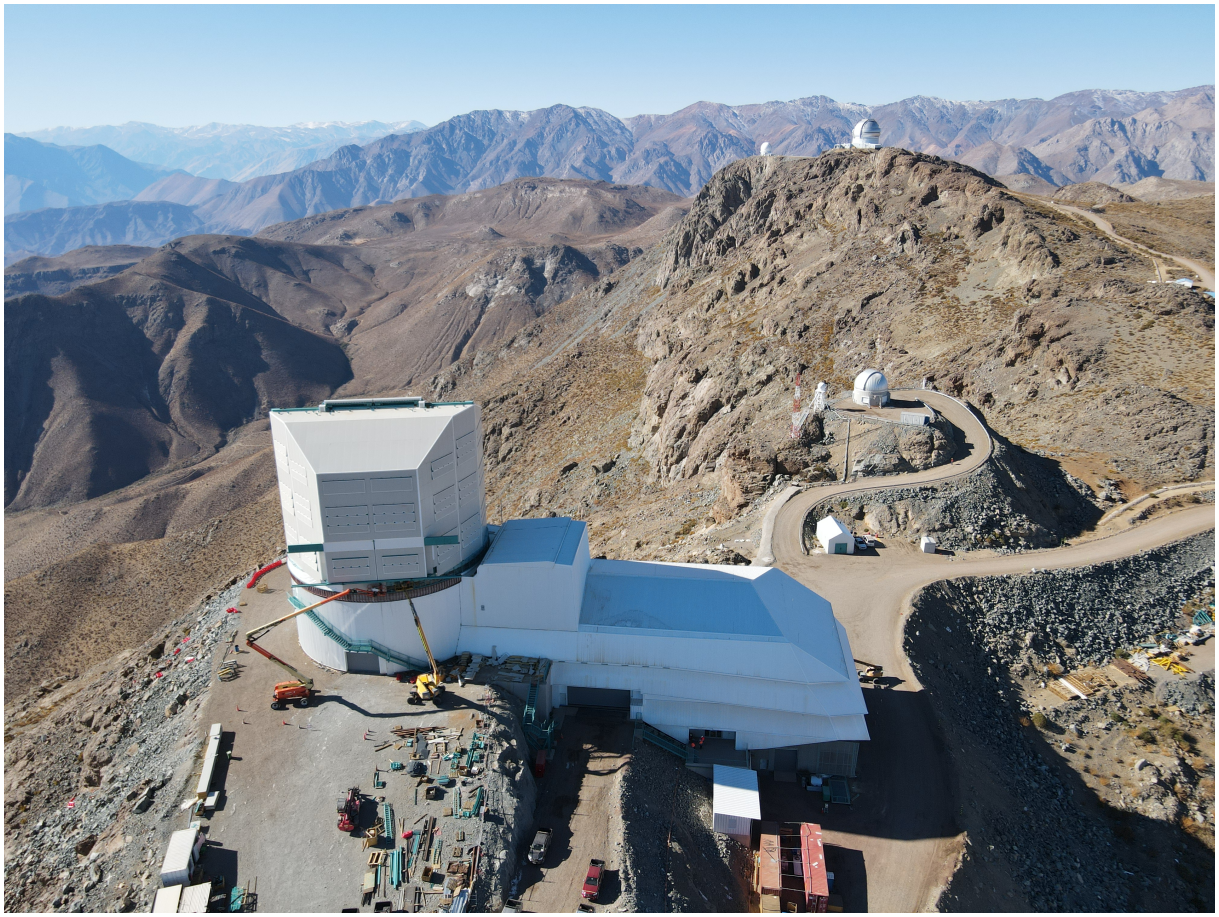

Figure 3.1: The Vera Rubin observatory under construction. Credits `https://www.lsst.org/`.

The Rubin camera[1] (LSSTCam) is represented schematically in Fig. 3.3 and sits at the center of the telescope mount at the intersection of the four arms. The focal plane (the 64 cm diameter meshed zone) is equipped with a 3.2 Gigapixel camera. The rest of the camera consists of six filters (with a diameter of 75 cm and a weight of up to 40 kg) in six optical bands (u, g, r, i, z, y) and three field corrector lenses (two of them are represented in Fig. 3.3). The filters are placed on the carousel that surrounds the camera, so they are ready to be installed in front of the focal plane by an automatic filter exchange system with a precision of a tenth of a millimeter (in Fig. 3.3, the red filter is ready to be used and the blue one is installed on the carousel). The carousel can only hold five of the six filters, so the sixth is installed on the carousel by a filter loader that temporarily replaces one of them. The filter exchange system is under the responsibility of IN2P3, and the result of a collaboration between five laboratories[2], and has been validated and installed on the camera in late 2020. Fig. 3.4 shows a front picture of the real camera, assembled at SLAC National Accelerator Laboratory in California, where we can see the focal plane through the different lenses. The COVID-19 pandemic has seriously impacted the schedule, and the system's first light will occur in $\sim$ July 2024.

[1]The largest CCD camera ever built.

[2]The researchers involved in the construction of the filter changer were rewarded with the CNRS collective Crystal medal in 2021, two of them work at LPSC, Grenoble. In addition to the filter changer, the IN2P3 is involved in the construction of the Camera Calibration Optical Bench (CCOB, wide and thin beams) that is used to calibrate the camera pixels at the per mil level, the test bench has been developed at LPSC, Grenoble.

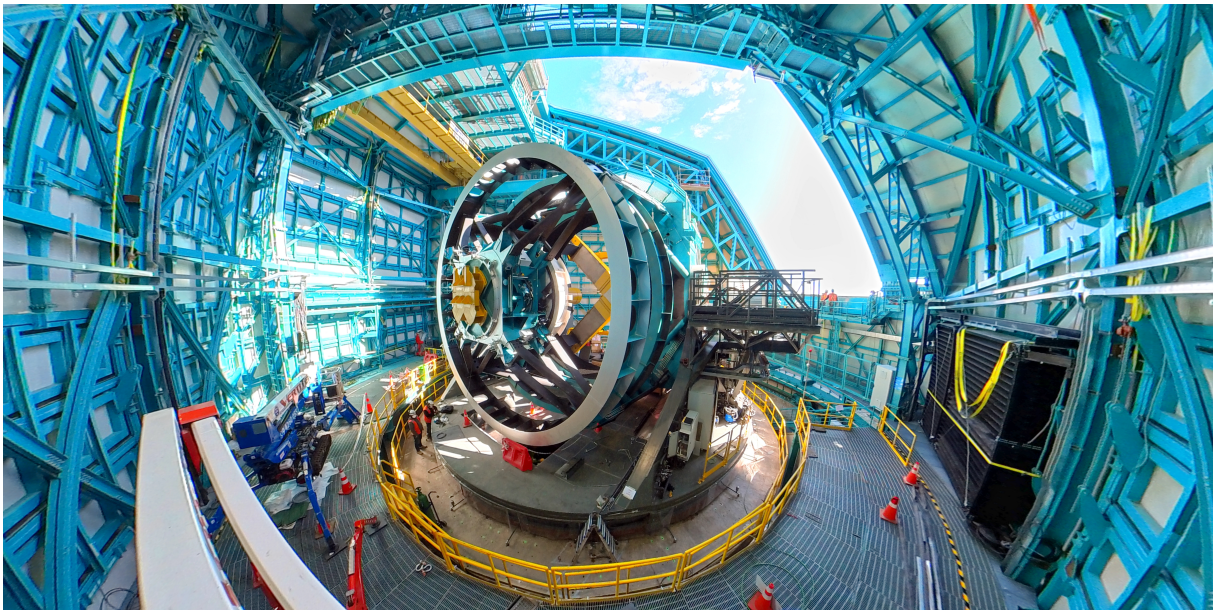

Figure 3.2: Picture of the inside of the observatory dome. Credits `https://www.lsst.org/`.

3.2 The Legacy Survey of Space and Time (LSST) and science goals

During the next ten years, the Rubin Observatory will be carrying out the Legacy Survey of Space and Time (LSST), by imaging the entire visible sky (18.000 deg^2 of the southern hemisphere sky). During this long period, each patch of sky will be visited approximately 825 times. The observing strategy (the cadence) is still being defined, but visits will last 34 seconds (15 seconds exposure + 2 seconds readout + 15 seconds exposure + new pointing as reading). The limit magnitude of detection will reach 24.5 in the $r-$band for a single exposure, and r <27.8 for the full stacked data by the end of the survey. Thanks to the 8-meter diameter mirror and the repeated exposures, the LSST magnitude limits are larger than the previous large photometric surveys e.g., SDSS (r <20-22), HSC (r <26.4), DES (r <23-24). Roughly 90% of the observation time will be used for surveying uniformly the 18.000 deg^2 of the southern sky on the *Wide-Fast-Deep* (WFD) mode, with the remaining time spent on *Deep Drilling Fields* (DDF), dedicated to very deep and fast time domain surveys of specific regions of the sky. The LSST will deliver a 15 TB dataset per night, 5.5 million images over 10 years, and tables of a few 37 billion objects (stars and galaxies), representing significant computing challenges from data storage to accessibility for the international scientific community. The IN2P3 will provide 40% of the computational resources required to process the data at the IN2P3's Computing Center[3] (or CC-IN2P3 in Villeurbanne, Lyon). At the time of this writing, the full survey operations are planned to begin around March 2025 (the calendar is constantly evolving).

The LSST will permit the scientific community to address many fundamental questions in cosmology and astrophysics[4], thanks to an unparalleled level of precision, as detailed hereafter.

- **Cosmology, dark matter and dark energy**: The large catalogs of galaxy positions,

[3]The web-page of the CC-IN2P3 can be found here: `https://cc.in2p3.fr`.

[4]For a more extensive review of the LSST science, I refer the reader to the science book LSST Science Collaboration et al. (2009).

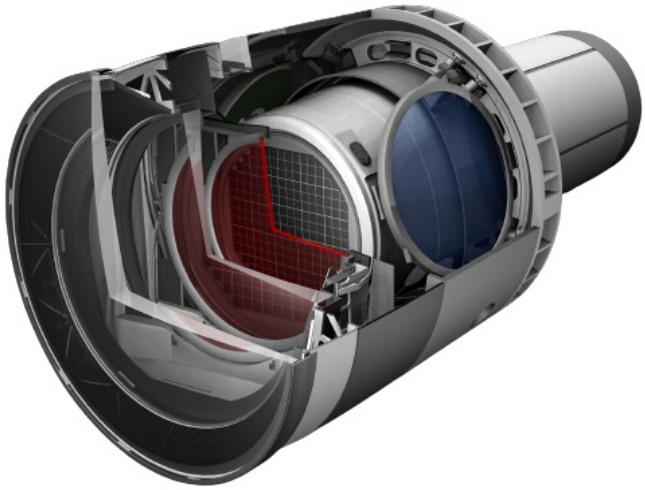

Figure 3.3: Rendering of the camera, from right to left, starting from the focal plane: the 3.2 Gigapixel camera, the filter (here, it is the red filter, the blue is installed on the carousel), two of the three field corrector lenses. The corrector lens in the forefront has a diameter of 1.65 m. Credits `https://www.lsst.org/`.

shapes and redshifts will significantly enhance our comprehension of dark matter and dark energy and their impact on the recent evolution of the Universe. This will be achieved by measuring the galaxy's two-point correlation, thus allowing us to probe the clustering of matter at various scales, as well as the weak gravitational lensing of billions of galaxies. Another key cosmological probe the LSST will exploit is cluster abundance, since hundreds of thousands of galaxy clusters will be detected in the LSST galaxy field up to a redshift of 1.2. LSST galaxy catalogs will also be used extensively for the galaxy cluster weak lensing mass calibration in Chapter 4 and in Chapter 6. Moreover, the LSST will measure thousands of Type Ia supernovae and will supplement previous datasets with thousands of new detections to probe the expansion rate with an unprecedented level of precision.

- **The Milky Way**: The LSST data will also refine the image of our home galaxy, the Milky Way, and of its neighboring galaxies (part of the Local Group) cataloging the positions and colors of a few tens of billions stars which will provide significant insights into the structure of star-forming regions within our Galaxy's disk, and the accretion history and overall structure of the Milky Way.

- **Observing the changing sky**: Thanks to its observing strategies, Rubin will explore the sky in the time-domain down to faint apparent magnitudes, and will dramatically increase the statistics of transient phenomena that are ubiquitous in the Universe. From this, the community will gain new insights into the physics of these various objects e.g. variable stars, electromagnetic counterparts of gravitational wave events or gamma-ray burst afterglows.

- **The Solar System**: The promised unprecedented capability of the faint time domain observation will allow LSST to take an inventory of orbiting systems, and millions of small moving bodies will be detected. LSST data will provide a wealth of information about the

Figure 3.4: Front picture of the integrated LSST Camera located in the clean room at SLAC (front lens $\sim$ 1.6 m). Credits `https://www.lsst.org/`.

formation history of our Solar System and planets (the Rubin observatory will also detect asteroids whose orbits would potentially cross the orbit of the Earth).

3.3 The Dark Energy Science Collaboration (DESC)

Numerous science collaborations have been established to tackle the various scientific topics outlined in the previous section (e.g., dark energy, solar system, Milky Way, galaxies, AGN, transients, strong lensing and informatics & statistics). Among those, the Dark Energy Science Collaboration (DESC) was created in June 2012 and is the largest scientific collaboration by far.

3.3.1 Collaborative context

DESC is separate from but has strong links with the Rubin project. The management of the DESC collaboration includes a management team, advisory board, collaboration council, membership committee, and publication board, among other aspects. For this section, we will discuss only the science-related aspects of the collaboration. DESC has more than 1,100 members coming from 15 countries, and I am one of the $\sim$250 full members[5].

The DESC is structured into several working groups (WG, listed in the diagram in Fig. 3.5)

[5]Full members are DESC members whose involvement and contributions to the collaboration have been recognized as important enough to be granted voting rights and to stand in some elections of the collaboration. Some responsibilities are also generally covered by full members (e.g., co-convener of DESC working groups, etc.).

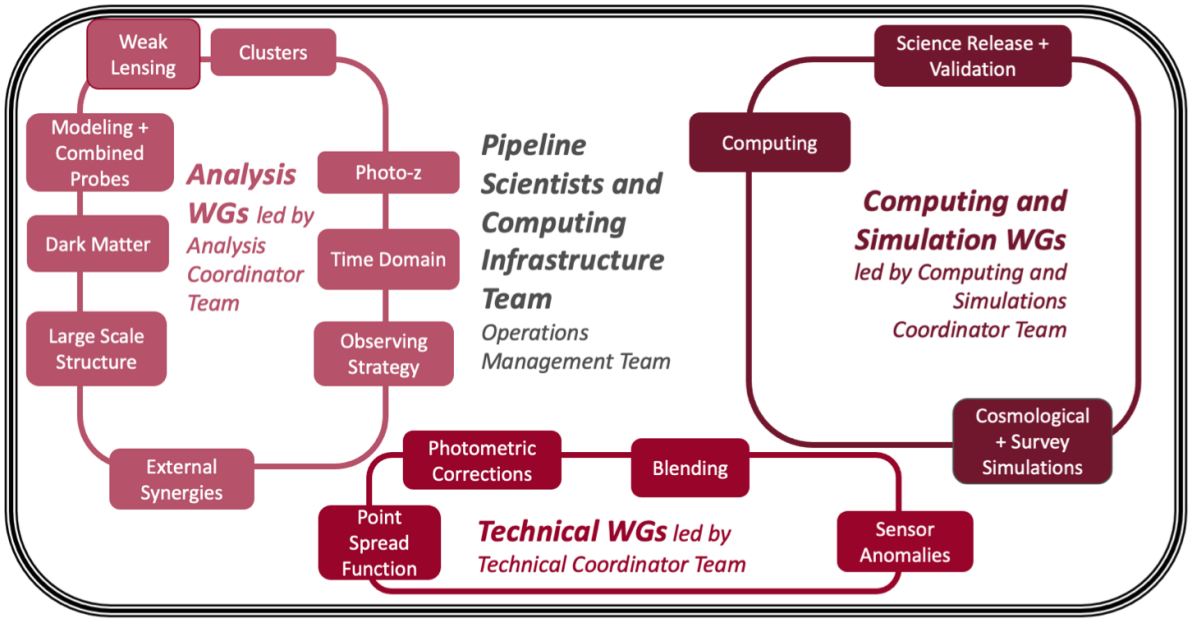

Figure 3.5: Current diagram of DESC working groups. This organization may evolve with changing management. Credits `https://lsstdesc.org`.

to tackle the several challenges arising with the upcoming LSST data (separated in analysis, computing and simulations, and technical WGs). Each of the primarily LSST probes is assigned to a dedicated WG (galaxy weak lensing, galaxy clustering, galaxy clusters and Supernovae Time Domain within the Analysis WGs). Also, the Modeling and Combined probes WG is dedicated to leveraging the complementarity between the different cosmological probes by developing models and statistical methods to combine them to extract most of the information from the LSST data. The members meet during two collaboration meetings per year, during which I presented many times, and once during the meeting plenary session. The DESC discussions also take place remotely each month (or even bi-weekly) within the dedicated working groups. My work took place mainly in the Cluster Working Group, where I presented ~ 20 times during the course of my PhD.

Among all the important challenges arising, the accurate determination of the photometric redshifts for the billions of galaxies to be observed by LSST and that will be used in all cosmological probes is one of the most critical challenges to be achieved, and the corresponding WG is actively working towards enhancing the precision and reliability of these estimates. In the next chapter, we will use catalogs of photometric redshift estimates produced by the photo-z WG. Moreover, the ability to measure precise galaxy shapes is a crucial aspect for high precision cosmology from weak lensing, and the weak lensing WG is responsible for developing and testing several shape reconstruction methods, that we will discuss in the next section.

To all the analysis WGs, the computing and simulations are crucial to the success of the LSST project by providing simulated sky to forecast the performance of the DESC pipeline. Three working groups are dedicated to these areas (we will discuss the Data Challenge 2 simulated data set in the next section). Because of my strong involvement with the DC2 simulations, I was asked to serve as a liaison between the Cluster and the Cosmology and Survey Simulations groups. In addition, two other working groups focus on technical aspects such as characterizing and correcting undesirable CCD effects inherent to LSST images that will affect the photometric reconstructions of lensing-related quantities (photometric redshifts, shapes, magnitudes, etc).

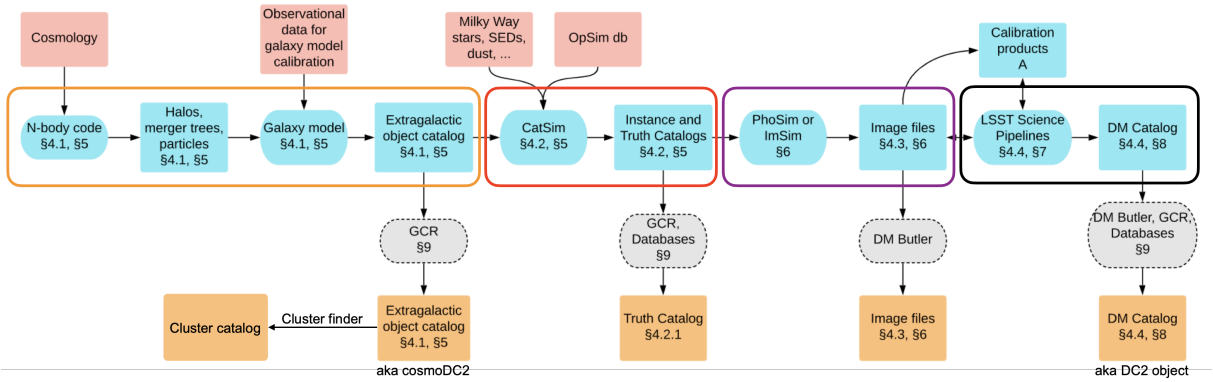

Figure 3.6: Workflow overview of the generation of the DC2 simulated sky. Adapted from Abolfathi et al. (2021). The Extragalactic object catalog will be referred to as cosmoDC2 while the DM Catalog is called DC2 object.

The DESC is also organized around topical teams which focusing on specific activities beyond the scope of a single project within or across DESC WGs (for example, I am an active member of the Cluster weak Lensing Mass Modeling team dedicated to developing the DESC code CLMM (Aguena et al., 2021b)). All contributions to the development of the DESC pipelines are pushed in a GitHub repository at `https://github.com/LSSTDESC`.

3.3.2 The DESC Data Challenge 2

The Data Challenge 2 (DC2), is a vast simulated astronomical dataset covering 440 deg^2, that is designed to help develop and test the pipeline and analysis tools of DESC for interpreting the LSST data (see full details in Abolfathi et al. (2021)). It follows DC1, which ended in 2017 and covered 40 deg^2 and aims at producing a small LSST-like dataset representative of the 10-year-long survey in all 6 bands reaching typical magnitudes $r = 27.5$ for galaxies up to redshift $z \approx 3$ and suitable for all cosmological probes (which was not the case for DC1). In this section I first present the production workflow of DC2, then I focus on two data products of the DC2, namely the cosmoDC2 and DC2 object catalogs.

3.3.2.1 Production workflow of the DESC DC2 data set

The workflow of the production in the Data Challenge 2 is shown in Fig. 3.6, going from the cosmological N-body simulations up to the processing of simulated images with the LSST Science pipelines. The production of this dataset represented a major effort of the DESC collaboration. Here, we will not enter into the details but simply give an overview of the main steps. The workflow is separated into four main components. The first element (represented by the orange box) contains the run of a cosmological N-body simulation, the modeling of the galaxy population and the generation of the cosmoDC2 extra-galactic catalog. The OuterRim N-body (gravity-only) simulation (Heitmann et al., 2019) is used as a starting point of the Data Challenge 2, where past light-cones have been created from simulation snapshots. The dark matter halos have then been identified by a friend-of-friend (FoF) halo finder[6]. Each halo has been assigned with a mass $M_{\rm FoF}$ (the sum of the individual dark matter particles associated between them)

[6]The Halo finder algorithm that was used defined groups containing all particles separated by a distance less than a given bl, where l is the average inter-particle separation and $b = 0.168$ is the linking length.

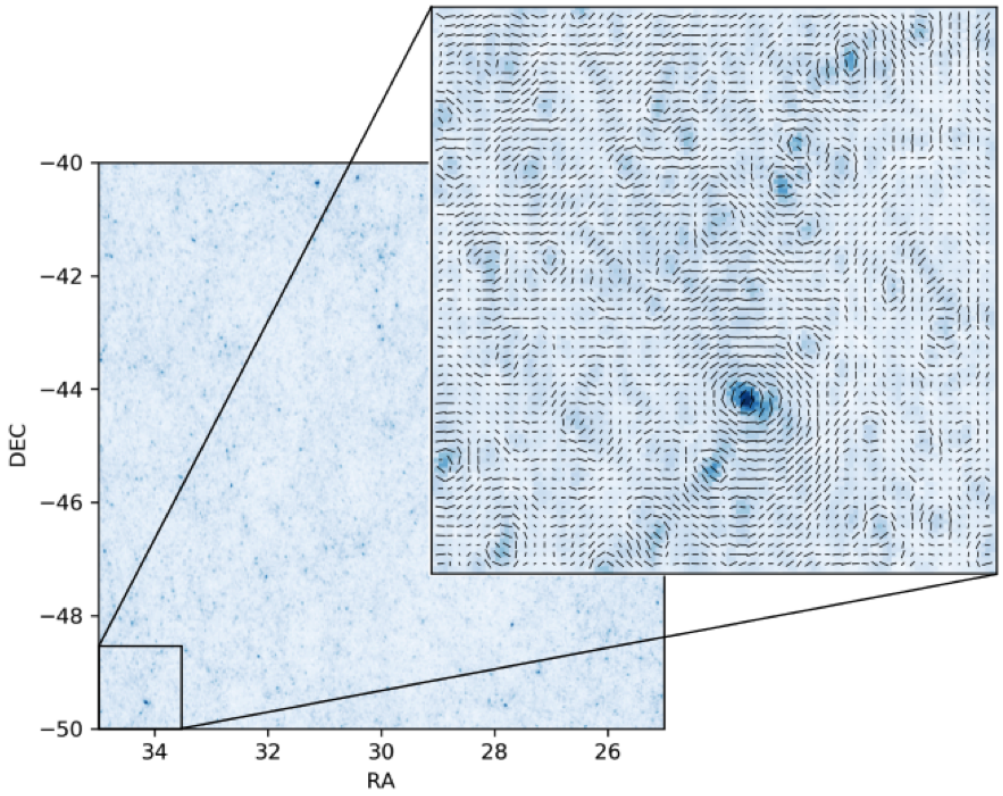

Figure 3.7: Left image: cosmoDC2 projected convergence field in a DC2 patch at $z = 1$. Right: Zoom in this region, with the two components of the shear field represented by the black lines. Figure from Korytov et al. (2019).

and a spherical overdensity mass M_{200c}[7]. To maintain a high degree of physical realism, the several key properties of the galaxies were drawn from the Galacticus semi-analytic model of galaxy formation (Benson, 2012), and were painted onto dark matter halos using GalSampler (Hearin et al., 2020). The derived galaxy properties include stellar mass, morphology, spectral energy distributions, broadband filter magnitudes and host halo information. The weak lensing shears and convergences at each galaxy position were estimated by a ray-tracing algorithm applied to the past light-cone particles in the simulation[8]. Fig. 3.7 shows the convergence and shear fields, that were reconstructed by ray-tracing for a DC2 patch in the sky. A truth catalog (red box) has been generated from cosmoDC2 and has been used for the generation of realistic images of the sky (purple box), including several realistic observational effects, such as atmospheric turbulence, telescope optics, and some detector effects (pixelization and noise of the CCDs, etc.). All of these injected effects mimic the Point Spread Function (PSF, quantifying the impact of all these sources of noise on a point source image) that is excepted with the

[7] M_{200c} masses are provided in the cosmoDC2 extension called Skysim5000 (covering 5000 deg^2, with an improved ray-tracing resolution by a factor of two in angular scale) obtained by fitting a NFW profile to each dark matter particle distribution. The correspondence between M_{FoF} and M_{200c} for cosmoDC2 halos is explored in Kovacs et al. (2022).

[8] The process involves first subdividing the particle light cone into distinct shells, and then retroactively tracking the paths of photons from an observer's grid back to a "source" shell. The deflections applied to the photon's trajectory are based on the particle density at each intermediate "lens" shell between the source and the observer, utilizing a ray-tracing algorithm. The methodology is described in detail in Korytov et al. (2019).

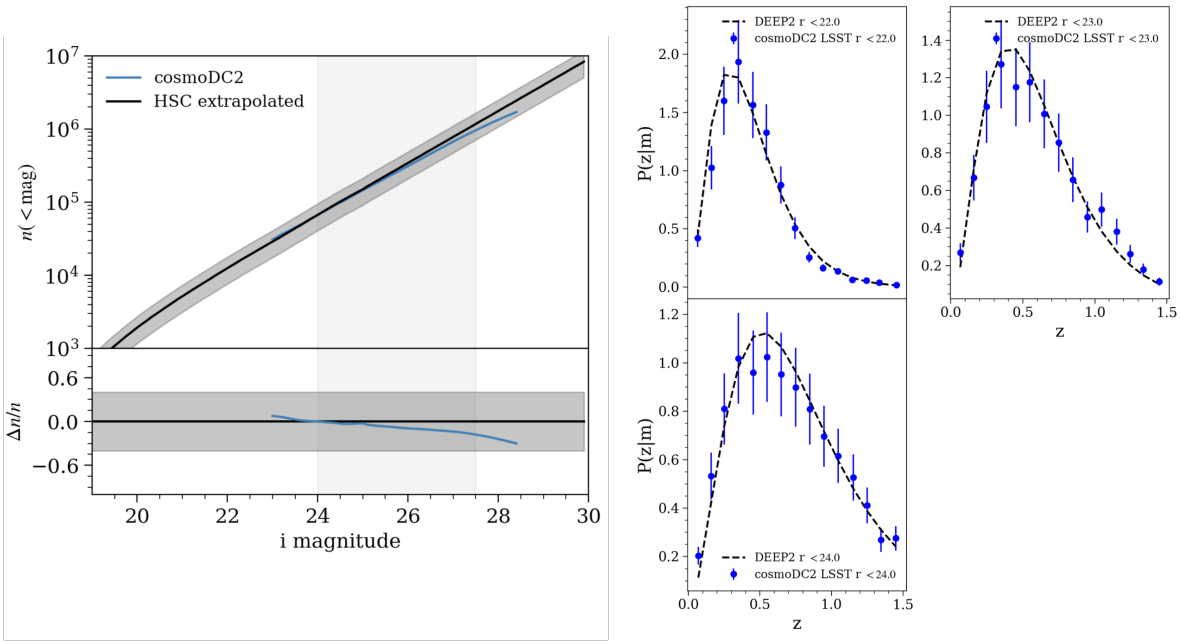

Figure 3.8: Left: Cumulative i-band number counts per square degree as a function of magnitude from cosmoDC2 (blue) and from the HSC survey (black). The bottom figure shows the relative difference between the two curves in the top panel. Right: The redshift distribution of cosmoDC2 galaxies (blue) compared to DEEP2 data (black dashed lines) for three magnitude-limited samples. Figures from Kovacs et al. (2022).

Rubin telescope, and its characterization is therefore crucial for making precise and reliable measurements in weak lensing cosmology (the PSF impacts the shape measurements, photo-z estimates, etc.). These images have been processed by the Rubin LSST science pipeline (the DM Stack) to obtain the DC2 object catalog (black box), accounting for the several observational systematic effects.

We now present the two different galaxy catalogs (cosmoDC2 and DC2 object) that differ by their level of complexity in accounting for observational effects (the DC2 object catalog) or by taking the inventory of truth galaxies that are simulated on top of the N-body simulation (cosmoDC2 extra-galactic catalog and related). All the galaxies in these catalogs will serve as sources to estimate the lensing signal around lenses as presented in Chapter 2, then we will present the lens catalogs, i.e., catalogs of galaxy clusters that were detected by different cluster finding algorithm.

3.3.2.2 cosmoDC2 extra-galactic catalog and the redshift add-on catalogs

The cosmoDC2 extra-galactic catalog

The cosmoDC2 extra-galactic catalog (detailed in Korytov et al. (2019), tested and validated in Kovacs et al. (2022) to which I contributed) contains ~ 2.26 billions galaxies and takes the inventory of ~ 550 properties of the "true" galaxies (i.e. true magnitudes in the six LSST bands, true redshift, true shapes, etc.) as well as the ray-tracing quantities per galaxy (shear and convergence) up to a magnitude depth of 28 in the r-band and to redshift $z \sim 3$. In that sense, it represents an ideal LSST dataset (there is no extinction from dust, stars, etc., only galaxies). Fig. 3.8 shows two of the various tests that have been carried out to validate the cosmoDC2

catalog. For instance, the left figure shows the great compatibility between the cosmoDC2 cumulative number density of galaxies in the i-band and real data from the Hyper Suprime Cam (Aihara et al., 2018), while the right figure compares the cosmoDC2 redshift distribution to the DEEP2 data[9] (Newman et al., 2013) for three magnitude-limited samples. The validation of cosmoDC2 was an important achievement in the collaboration, ensuring that the DC2 data can be used extensively to prepare the analysis of the LSST data. A first level of complexity is added by computing the photometric redshifts of cosmoDC2 galaxies, using existing photoz codes. Two photometric redshift codes have been run and the corresponding estimated galaxy redshifts are stored in two add-on catalogs.

FlexZBoost

The first is FlexZBoost[10] (Izbicki & Lee, 2017). FlexZBoost is an empirical technique that uses machine learning and learns the mapping between the galaxy's observed colors and the true cosmoDC2 redshift by using a training dataset (for real data, it is done by matching a spectroscopic reference sample with the photometric dataset, the former redshifts being much more precise). In cosmoDC2, FlexZBoost was trained with a complete subsample of galaxies that extended to $i < 25$. The primary source of systematic error in these photometric redshift approaches is the uncorrected selection functions in the reference sample that is used for training. In contrast, if the training set is a perfect subsample of the photometric sample, FlexZBoost will obtain very accurate predictions because the aforementioned mapping can be learned with very high accuracy. This scenario was considered in the estimation of FlexZBoost redshifts in cosmoDC2, due to limitations in the available simulation dataset that does not include a spectroscopic reference sample but true galaxy redshifts. As such, the primary source of systematic error in FlexZBoost could not be accurately modeled, and the obtained photometric redshift predictions can be expected to be unrealistically accurate. It was then found that the accuracy of photometric redshift results obtained on cosmoDC2 galaxies by using the FlexZBoost code is overestimated.

BPZ

The second is a template-based algorithm called BPZ[11] (Bayesian Photometric Redshifts, Benítez (2011)). The BPZ method is a template-fitting technique that formulates a likelihood of the galaxy's observed colors from a set of Spectral Energy Distribution (SED) models. The BPZ method relies on smooth physical likelihoods of observed photometry, while the simulation data available in DC2 contains some un-physical discontinuities in the observed galaxy photometry. Thus, the continuous physical BPZ likelihood will not be able to produce accurate predictions. The BPZ photometric redshift predictions will therefore be un-physically biased due to the DC2 simulation limitations. We finally note that this limitation does not affect FlexZBoost because empirical techniques can *learn* this un-physical mapping between photometry to redshift as it does not rely on SED models. Opposite to FlexZBoost, it was then found that the accuracy of BPZ photometric redshift in cosmoDC2 is underestimated. We illustrate the performance of the two photoz codes in Fig. 3.9 that shows the mean photometric redshift reconstructed from the photometric redshift distribution for ~ 30.000 galaxies, with true cosmoDC2 redshift $z \geq 0.5$. Redshifts obtained respectively with FlexZBoost and BPZ are represented in magenta and cyan,

[9]The DEEP2 survey covers 3 deg^2 and the corresponding galaxy catalog is complete up to $r = 24.1$ over the redshift range $[0, 1.4]$. The color completeness of this dataset was suited for validating the colors of cosmoDC2 galaxies.

[10]The code is available here `https://github.com/rizbicki/FlexCoDE`.

[11]The code is available here `https://www.stsci.edu/~dcoe/BPZ/`.

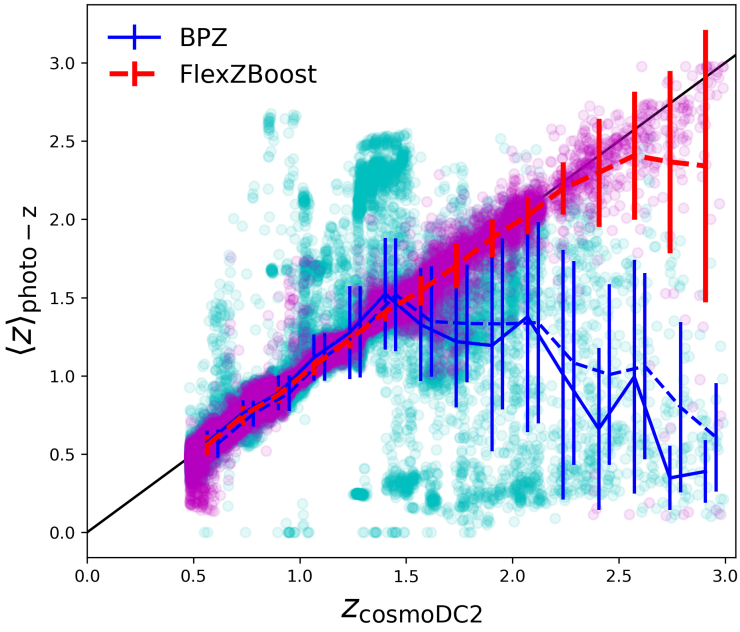

Figure 3.9: BPZ (cyan) and FlexZBoost (magenta) individual photometric redshift estimates. BPZ and FlexZBoost mean photometric redshift in bins of cosmoDC2 galaxy redshifts are respectively in blue and red. The black line corresponds to the $x = y$ direction. The dashed blue line corresponds to the average BPZ mean photometric redshifts after a quality cut on BPZ photozs.

and the average photometric redshifts in bins of cosmoDC2 galaxy redshift are respectively represented in red and blue. We see that, at high redshift galaxies, the BPZ photometric redshift estimates are significantly degraded (the distribution of catastrophic redshift estimates is more important for distant sources) while we find that for the FlexZBoost estimates, they correlate correctly with the true redshift values at all redshifts. We applied a quality cut on the BPZ catalog, defined as the integral of the photometric redshift distribution over a fixed redshift range around the maximum value of the PDF (this quality is computed and available for each galaxy in the add-ons catalog). The dashed blue line corresponds to the mean photometric distribution of the new BPZ catalog. As such, none of these estimations are realistic and updated redshifts are currently being produced by the collaboration. These were not available at the time of this work, so we proceeded with the first version of the photoz catalogs. The real LSST data will most likely lie between the pessimistic BPZ and optimistic FlexZboost run we used. Therefore, by considering both approaches, we will be able to bracket the photoz impact on the results of Chapter 4.

3.3.2.3 The DC2 object catalog

An example of the DC2 simulated sky is shown in Fig. 3.10: The left panel shows the subdivision of the sky in numerous tracts and the middle panel shows the DC2 image corresponding to the upper quadrant of a tract (the right panel shows a zoom-in region). The DC2 object catalog is a table of objects detected on the processed DC2 images, with measured positions, magnitudes and shapes (again, see full details in Abolfathi et al. (2021)). DC2 object catalog also contains stars, for which a flag allows us to filter them out. The photometric redshift estimates for the DC2 objects have been derived from their observed DC2 magnitudes with the BPZ algorithm. For each detected object on DC2 images, two shape catalogs are available, which differ by the

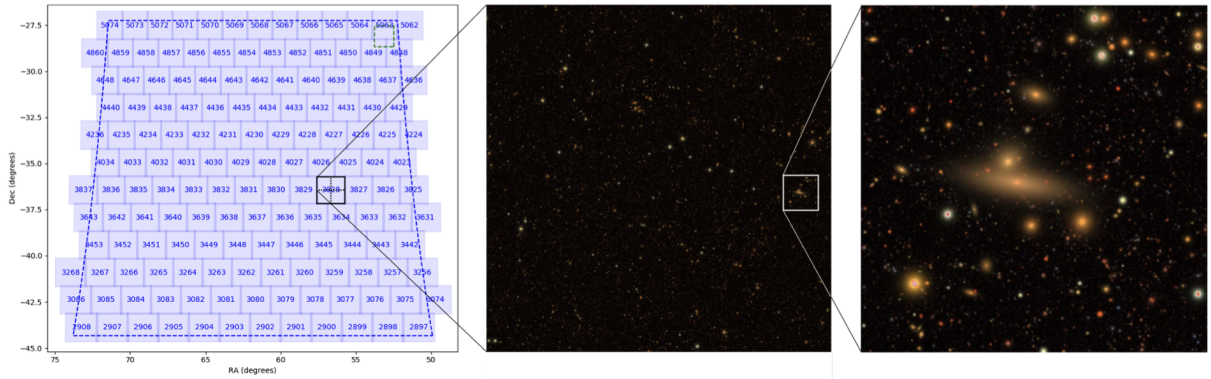

Figure 3.10: Detailed view of the simulated DC2 sky. Left: map of the tracts in the DC2 area. Middle upper quadrant of the tract 3828. Right: further zoom in the image simulation, showing a massive galaxy cluster characterized by the presence of red galaxies. Figure from Abolfathi et al. (2021).

algorithm that is used for image processing and object shape measurement.

- The first catalog is obtained by applying the HSM algorithm[12], detailed in Hirata & Seljak (2003) and tested/characterized using real data by Mandelbaum et al. (2005). HSM is based on the estimation of the object ellipticity from the second-order moments of its surface brightness (as presented in Appendix B). The ellipticity estimate for a single object usually follows the decomposition (Mandelbaum et al., 2014)

$$\widehat{\epsilon} \approx \epsilon^{\rm int} + (1+m)\gamma + c, \tag{3.1}$$

  accounting for a coherent additive bias c due to PSF uncorrected anisotropies, that introduce a preferred direction in the image plane. Other effects such as the uncorrected smoothing of the PSF and or the impact of (unobserved) galaxy properties are included in a multiplicative bias m. HSM shapes need external calibration from simulations to evaluate these coefficients (see e.g. Mandelbaum et al. (2018) that calibrated HSM bias factors on HSC data with simulations).

- The second shape catalog was obtained applying the METACALIBRATION[13] algorithm (Sheldon & Huff, 2017) enabling shear self-calibration using only the available imaging data. The code was recently used in the DES-Y3 cosmic shear analysis in Abbott et al. (2022). The METACALIBRATION procedure is significantly different from HSM. Given $\widehat{\epsilon}$ to be the biased estimator of the ellipticity, we have at first order of shear perturbation that

$$\widehat{\epsilon} \approx \epsilon^{\rm int} + \mathrm{R} \cdot \gamma, \tag{3.2}$$

  where $\mathrm{R}_{ij} = \partial\widehat{\epsilon}_i/\partial\gamma_j$ is the ij response matrix coefficient, namely the derivative of the i-th ellipticity component with respect to a change in shear amplitude in the j-th direction. At a given location in the sky and for a given selection cut on background galaxy redshift, the "true" average shear γ can be recovered by making the product of the inverse response matrix with the average measured ellipticity, namely $\langle\gamma\rangle = \langle\mathrm{R}\rangle^{-1} \cdot \langle\widehat{\epsilon}\rangle$. The matrix

[12]The code is available at `https://galsim-developers.github.io/GalSim/_build/html/hsm.html`.
[13]The code is available at `https://github.com/emhuff/MetaCalibration`.

R decomposes in two separate contributions, the shear and selection responses denoted respectively by R_γ and R_S. The R_γ for a single galaxy is derived from several shape estimates measured on artificially sheared versions of the realistic DC2 sky images, so there is no need for external calibration. The selection response denotes the impact of the sample selection based on sheared ellipticity quantities applied to un-sheared ellipticities. METACALIBRATION has evolved into METADETECT (Sheldon et al., 2023) which will be the prime method to measure shapes from LSST data.

The DC2 object catalog is built upon a series of observational effects affecting the reconstruction of galaxy properties, and is a playground to study their impact on shear calibration, photometric redshift calibration, the effect of blending (the misinterpretation of overlapping flux densities from separate sources that are aligned along the line-of-sight) on weak lensing observable.

3.3.2.4 Halo and galaxy cluster catalogs

A list of dark matter halos may be extracted from the cosmoDC2 catalog. cosmoDC2 galaxies are associated with a halo ID, identifying the halo they belong to. The halo's central galaxies have specific flags that allow us to extract a list of halos from the cosmoDC2 catalog.

Galaxy clusters are the observational counterpart of dark matter halos and can be identified by running cluster finder algorithms on the DC2 catalogs. The redMaPPer cluster finder (Rykoff et al., 2014), which identifies galaxy clusters through the presence of red-sequence galaxies, has been run on cosmoDC2. It has already been widely used on SDSS (Abdullah et al., 2020) and DES (Abbott et al., 2020) data and as such it is one of the best-studied cluster finders. The resulting cluster catalog provides the cluster positions, redshifts and richnesses (calculated as the sum of membership probabilities of galaxies around the cluster) along with the list of potential member galaxies. We will give more details about the redMaPPer cluster catalog in the Chapter 4. In addition to redMaPPer, other detection methods have been applied to find clusters of galaxies in the cosmoDC2 data (WaZP Aguena et al. (2021a), AMICO Bellagamba et al. (2018)) whose corresponding catalogs are, or will soon be, available within the collaboration, but they were not at the time I worked on the projects presented in Chapter 4.

3.4 Conclusions

The Vera C. Rubin Observatory will conduct the ten-year Legacy Survey of Space and Time providing a vast catalog of billions of galaxies, and will enable the scientific community to tackle the fundamental questions about dark energy and dark matter. In that context, the Dark Energy Science Collaboration is preparing the analysis and interpretation of the LSST data. To do so, the collaboration has initiated the Data Challenge 2, which has been designed to mimic LSST data and to develop and exercise the DESC science pipelines. In the next chapter, I present in detail my effort in contributing to the DESC effort for cluster abundance and cluster lensing science, by developing the necessary software tools and using the DC2 products.

Chapter 4

Preparing for DESC cluster cosmology analyses with the DC2 dataset

Contents

Clusters of galaxies can provide a wealth of information about the growth of structures in the Universe. The Vera C. Rubin facility will conduct the LSST that will mark a new era of precision cosmology with galaxy clusters due to an order of magnitude increase in the number of detected clusters, and with the improved precision of weak gravitational lensing information. To take full advantage of the upcoming data, the DESC has initiated the DC2 simulations to develop and exercise the several tools that are needed to take full advantage of the LSST data. The DC2 products required for this work were described in the previous chapter. In this chapter, I detail my contributions to the DESC effort in developing software and analyzing the DC2 data for cluster lensing and cluster abundance cosmological analyses. The main focus of this chapter is the inference of the weak lensing mass for stacks of clusters in the DC2 simulations, of their mass-richness relation from cluster lensing and cluster abundance. From the work I present

below, I lead and/or contributed to four refereed DESC notes[1][2][3][4](we will refer to them in the manuscript as 1: [DN1], 2: [DN2], 3: [DN3], 4: [DN4]) and I am currently writing a DESC paper (Payerne et al., in preparation). These works aim to prepare for the cluster lensing and cluster abundance analyses of LSST data, for which I developed several software tools that will be used to build the DESC cluster pipeline.

In Section 4.1, we present the measurement of the lensing signal around the position of DC2 lenses, namely the DC2 dark matter halos. The methodology developed to estimate the lensing signal will serve as a basis for the other sections, that mostly differ by modeling aspects. In Section 4.2 we detail the studies of the DC2 mass-richness relation where, this time, the lenses are redMaPPer clusters. In Section 4.3 we present the inference of the mass-richness relation from a joint analysis of cluster lensing and cluster abundance, as well as preliminary cosmological constraints using the DC2 data. While it will be mentioned through the different sections of this chapter, I recap in Section 4.4 the software I developed for the analysis of DC2 clusters, as well as my several contributions to the DESC pipeline development. We conclude this chapter in Section 4.5.

4.1 Cluster lensing in the DC2 simulations: assessment and limitations

In Chapter 2, we presented the theoretical aspects of weak gravitational lensing by clusters. Here, we put these concepts into practice using DC2 data. Because of the various levels of realism/complexity of the DC2 data, we will proceed sequentially. First in Section 4.1.1 we focus on the lensing profiles around ensembles of halos first using the true, known, redshifts and shapes of background galaxies. This is then complexified by using measured photometric redshifts and shapes in Section 4.1.2. This essentially is an assessment study to evaluate the signal around haloes in the DC2 simulations.

4.1.1 cosmoDC2 catalog: haloes, true/measured redshifts, true shapes

The DESC undertook a large campaign of validation of the cosmoDC2 dataset (published in Kovacs et al. (2022)), to which I contributed in a cluster weak lensing perspective. This is what I present in this section.

We consider the lensing signal around the center of cosmoDC2 dark matter halos (identified by the position of their central galaxy) using the cosmoDC2 extra-galactic catalog (version 1.1.4) for the background source samples. We apply magnitude cuts on the i-band so that the effective number density of source galaxies is comparable to the effective number density for weak lensing that will be used in the context of LSST after 5 years of data, $n_{gal} \approx 25$ gal.arcmin^{-2} (Chang et al., 2013). Applying these magnitude cuts, we extract the galaxies located in the background of ≈ 5000 dark matter halos up to the projected physical radius $R \sim 10$ Mpc. We consider the "true" galaxy shapes, namely accounting for their simulated intrinsic ellipticities combined with the local shear[5] at each galaxy position. We also consider the "true" redshifts of galaxies. This data set consists in the most ideal case, where shapes of galaxies are known without degradation due to observational systematic[6].

[1]C. Payerne et al., 2022, Mass-richness relation in DESC simulations I: Effect of photometric redshifts and shape measurements.
[2]C. Payerne et al., 2022, Mass-richness relation in DESC simulations II: Effect of modeling choices.
[3]M. Ricci et al., 2022, Measuring cluster masses from the magnification of background galaxies in cosmoDC2.
[4]C. Payerne et al., 2022, `https://github.com/LSSTDESC/DC2_CL_PZ/tree/compute_deltasigma`.
[5]The local shear per galaxy is obtained by ray-tracing, see Section 3.3.
[6]The shape noise per ellipticity component (i.e., considering the intrinsic ellipticity and shear) is $\sigma_\epsilon \sim 0.15$.

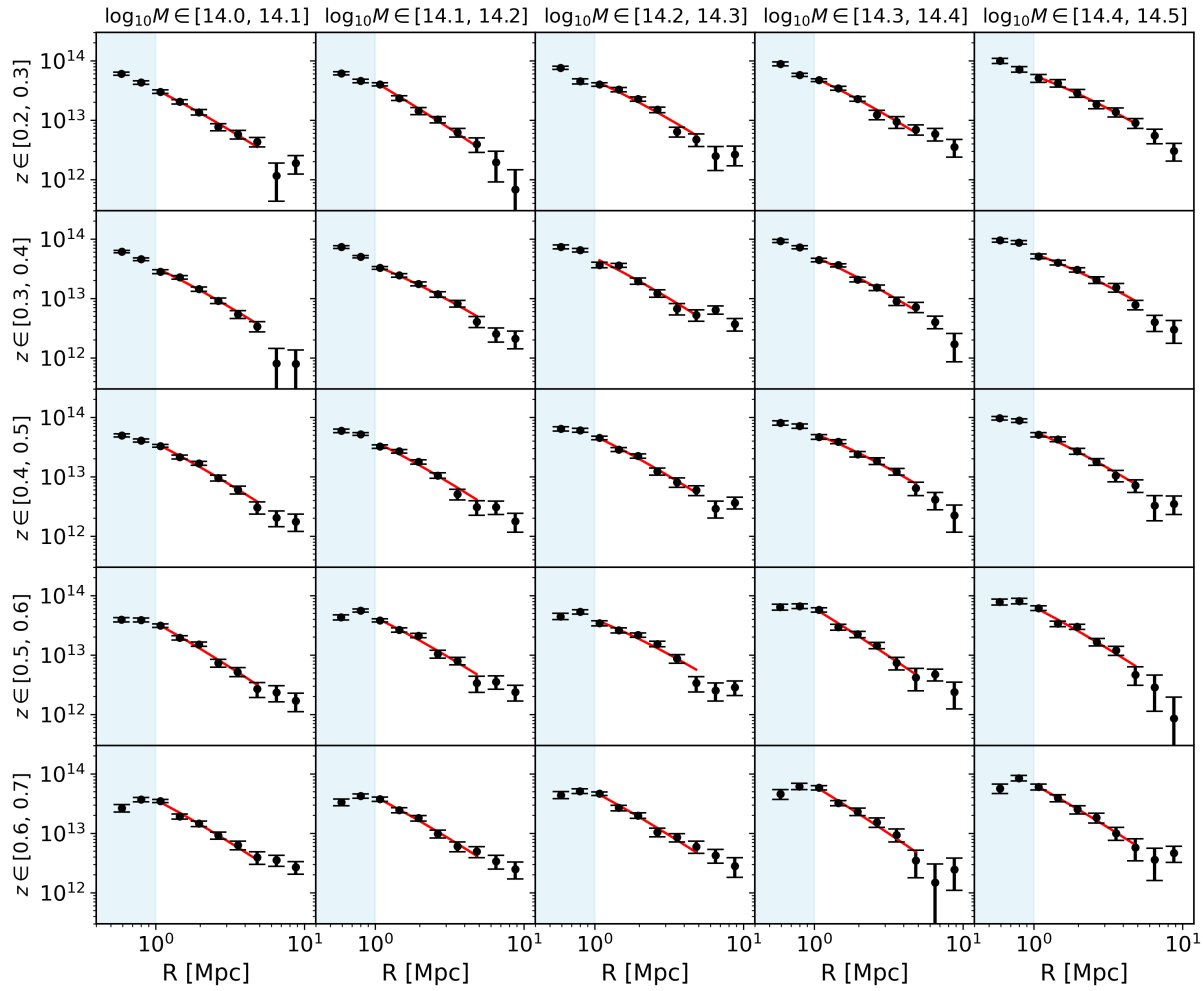

Figure 4.1: Stacked excess surface density profiles in mass and redshift bins. The blue region shows the scales that are affected by the non-physical attenuation. The red lines correspond to the best fit excess surface density profiles. See text for details.

Using the procedure described in Section 2.4, we compute the stacked lensing profiles[7] around dark matter halos. Fig. 4.1 shows an example of stacked excess surface density profiles (in black) around dark matter halos, separated in five redshift bins from 0.2 to 0.7, and in 5 mass bins from $10^{14} M_\odot$ to $10^{14.5} M_\odot$ (the mass-redshift binning of the dark matter halo catalog will change through this section since we tested variety of setups to explore multiple aspects of the data). The stacked lensing profiles are estimated in 10 radial bins from 0.5 to 10 Mpc. The redshift bins correspond to the different rows and the mass bins to the columns. The error bars correspond to the diagonal elements of the corresponding stacked covariance matrices estimated by 200 bootstrap re-samplings (for details on the stacked lensing profiles and covariance estimators, see Section 2.4.2).

From left to right, we see in Fig. 4.1 that the overall amplitude of the stacked excess surface density is correlated with the mean mass within the bins (it increases with the mean mass within the mass bin at fixed redshift). This is the first qualitative behavior that is expected. From top to bottom in Fig. 4.1 (this time the mean redshift increases at fixed mean mass), stacked lensing profiles display an increasing turn-off region (corresponding to $R \leq 1$ Mpc) with respect to the

[7] In this idealistic dataset, the weight per galaxy only accounts for the geometrical weights $\Sigma_{\rm crit}^{-2}(z_l, z_s)$.

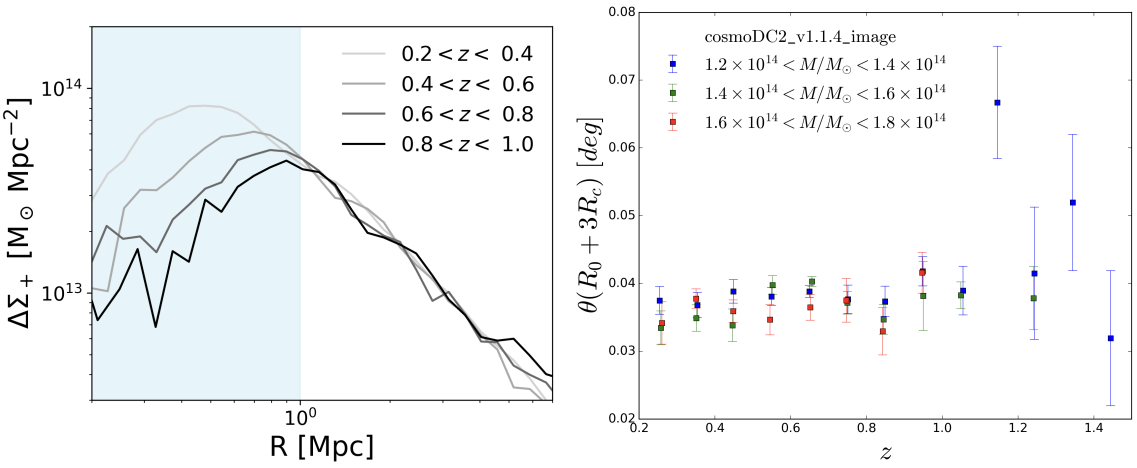

Figure 4.2: Left: excess surface density profiles in a single mass bin and in four redshift bins. The shaded region corresponds to the radial range affected by the non-physical attenuation due to the cosmoDC2 ray-tracing resolution. Right: width of the attenuation region as a function of redshift. The different colors correspond to the different mass bins.

mean redshift in the stack (blue shaded region in Fig. 4.1). This is not the expected behavior and deserves more scrutiny.

The situation is clearer in Fig. 4.2 (left), that shows the stacked lensing profile in a given mass bin and in four different redshift bins. For clarity, we do not show the error bars. We see a strong redshift dependence on the shape of the weak lensing profile at small scales. As seen in Chapter 2, this effect could be a signature of miscentering, which was our first thought.

As a sanity check, we investigated the redshift dependence of the angular size of this attenuation region. For each stack k with measured excess surface density $\widehat{\Delta\Sigma}_k$ within the radial range $[0.2, 5]$ (we estimated the signal at even smaller radii than what is done in Fig. 4.1 to measure more accurately this attenuation), we model the lensing signal by a NFW profile multiplied by a sigmoid $S(R)$ with two free parameters R_0 and R_c. This function goes from 0 to 1, with a transition occurring around R_0 and a transition width controlled by R_c. For different mass and (thinner) redshift bins, the best fit of the sigmoid width $R_S = R_0 + 3R_c$ (such that $S(R_S) = 0.99$) is represented in Fig. 4.2 (right) with respect to redshift. We see that the typical scale of the angular resolution $\theta = R_S/D_l$[8] is roughly constant with redshift (the scatter and large error bars of high redshift points are due to the low cluster statistics), which is not compatible with miscentering, but is a (known) numerical angular resolution issue[9], thus being redshift-dependent when converted to physical distance in Mpc. This shear signal attenuation is one limitation of the DC2 simulation for cluster weak lensing studies.

As a consequence, we chose to use only the $R > 1$ Mpc region for each stack in the analysis. This is a conservative choice that appears valid over the full mass and redshift range. However, this also means we cannot use the innermost region that has the largest SNR values. DC2 is an impressive achievement for the collaboration but this intrinsic limitation will limit the cluster-related forecasts of what could be achieved with the LSST data.

To estimate the error bars, we rely on the stacked covariance matrix of the cluster lensing

[8] D_l is the distance to the cluster.

[9] This attenuation has already been observed in the galaxy-galaxy lensing in DC2 (Korytov et al., 2019) and is associated to the limited resolution of the ray tracing to compute the lensing shear and convergence at each galaxy positions.

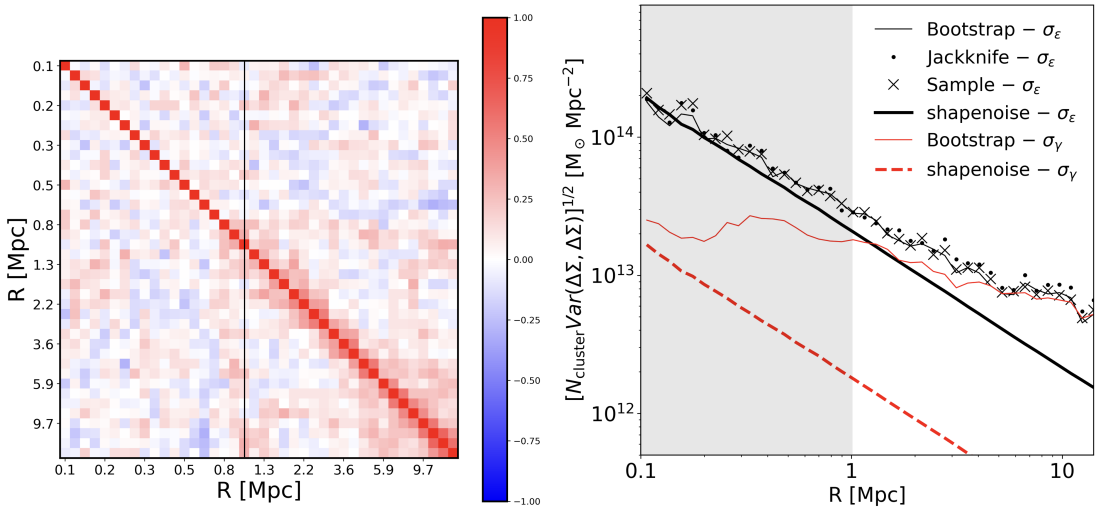

Figure 4.3: Left: Correlation matrix of the excess surface density profile using bootstrap covariance matrix with 300 re-samplings. Right: Square root of the diagonal elements of covariance matrix estimation of a stack excess surface density profile. When considering shape noise, we show the diagonal element of the covariance matrices for the bootstrap (full black line), jackknife (black dots) and sample (black crosses) covariance estimation methods. The shape noise variance considering intrinsic ellipticities is in black full line. For the noise-free case, we show the bootstrap (red full line) method. The shape noise variance using the shear intrinsic dispersion is represented with the red dashed line.

profiles, that also need to be assessed. For a thinner binning of the radial range (to see in detail the behavior of the covariances between bins), we show the correlation matrix obtained via bootstrap resampling in Fig. 4.3 (left). The correlation between bins is expected to be larger at scales below $\sim$ 1 Mpc (Wu et al., 2019). However, below this scale (denoted by the black vertical line), we find that the correlation between bins is consistent with 0. It is likely due to the shear attenuation, smoothing the underlying correlation between bins. We see that the correlation between distant bins is significantly reduced, and most of the off-diagonal terms are very noisy. We find the same correlation matrix using jackknife resampling.

The mean error on the shear measurement in a given stack is shown in Fig. 4.3 (right). We show $\sqrt{N_{\rm cluster}\mathbf{C}_{i,i}}$ where $\mathbf{C}_{i,i}$ is the covariance matrix estimated via the 3 different methods (bootstrap: thin black line, jackknife: black dots, sample covariance: black crosses). The three methods give consistent results. In thick black line, we show the corresponding shape noise variance prediction of the excess surface density measured on cosmoDC2 shapes[10]. We note that at small scale ($1 \leq R \leq 3$), the variance is dominated by shape noise, whereas above 3 Mpc the variance is dominated by the large-scale structure fluctuations (Wu et al., 2019). The interpretation of the variance for scales below 1 Mpc (grey-shaded region) is more difficult since the measurement of the cluster lensing signal is limited by the ray-tracing resolution, thus erasing physical variance and correlation from cluster physics. For the fitting of cluster masses that we will perform through this chapter, the lensing signal will be largely dominated by shape noise.

Estimation of stacked cluster masses: To infer cluster masses from the cosmoDC2 dark

[10]In red full line, we show the error on the excess surface density without considering intrinsic shapes (by considering the cosmoDC2 shear as the individual galaxy ellipticities), and the dashed red line is the corresponding shape noise variance with $\sigma^{\gamma}_{\rm shape} \approx 0.015$. This noise-free case enables to investigate the behavior of the variance at intermediate scales which corresponds also to the large-scale structure fluctuations.

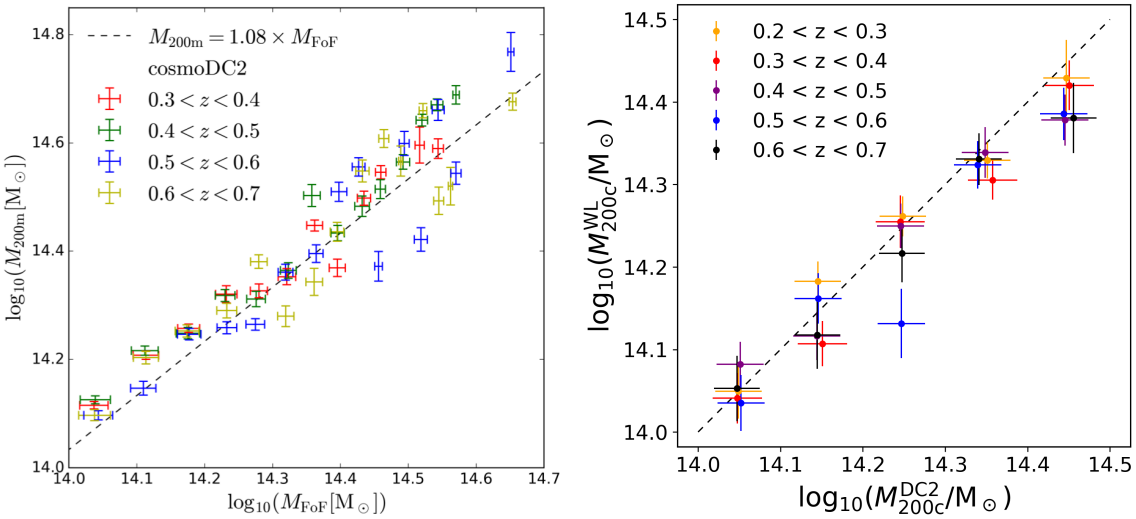

Figure 4.4: Left: lensing mass estimates M^{WL}_{200m} with respect to the average cosmoDC2 masses M_{FoF}. The different colors correspond to the different redshift bins (figure from Kovacs et al. (2022)). Right: lensing mass estimates M^{WL}_{200c} with respect to the average cosmoDC2 masses M^{DC2}_{200c}. The different colors correspond to the different redshift bins. The dashed line corresponds to $x = y$.

matter halo shear profiles in Fig. 4.1, we will use only diagonal elements of the covariance matrix, i.e. neglecting correlation between bins, due to the low statistical level of ~ 100 clusters per stack[11] (the off-diagonal terms are very noisy) compared to standard cosmological analysis with ~ 1000 clusters per stack. We consider a NFW profile, that depends on two parameters, the mass $\log_{10}(M_{200c})$ and the concentration c_{200c}. We use a mass-concentration relation from Diemer & Kravtsov (2014) so that only the mass remains a free parameter. As we neglect the inner part of the profile (corresponding to $R < 1$ Mpc), we do not account for the effect of miscentering, contamination of foreground galaxies[12] and non-weak lensing corrections (see Chapter 6) in the mass fitting procedure. We also consider 5 Mpc as an upper radial limit, so the lensing shear signal is dominated by the 1-halo term, and we can restrict the modeling of the NFW profile only.

We show in Fig. 4.4 (left) an example of validation test that we performed to validate the cosmoDC2 extragalactic-catalog. We fitted[13] an effective M^{WL}_{200m} mass corresponding to an "effective" cluster, and that produces a lensing signal that would match with the stacked one. At the time of this first analysis, published in Kovacs et al. (2022) the only "true" mass information in cosmoDC2 was the Friend-of-Friend mass, making it difficult to give strong conclusions. Later, the catalog was supplemented with M_{200c} masses and I redid the comparison using this mass definition in Fig. 4.4 (right). The inference setup is different from the results showed in Fig. 4.4 (left), since this time we fit an effective M_{200c} mass within bins of M_{200c} masses (we have applied different source selection, number of radial bins, etc.), and we have reported the best fitted lensing profiles in Fig. 4.1. The results of these two analyses show a large/good positive correlation between the mass estimates and masses from the simulation. Given these radial cuts and modeling choices, the reconstructed weak lensing masses are roughly $1 - 2\sigma$ compatible

[11] cosmoDC2 is only 440 deg^2.

[12] Also because background galaxies are perfectly identified in this idealistic dataset.

[13] We use the emcee package (Foreman-Mackey et al., 2013) to draw the posterior of the stacked masses within bins of M_{FoF} mass and redshift. We use a Gaussian likelihood with the excess surface density profile covariances. We plot the recovered mean mass and posterior dispersion as error bars.

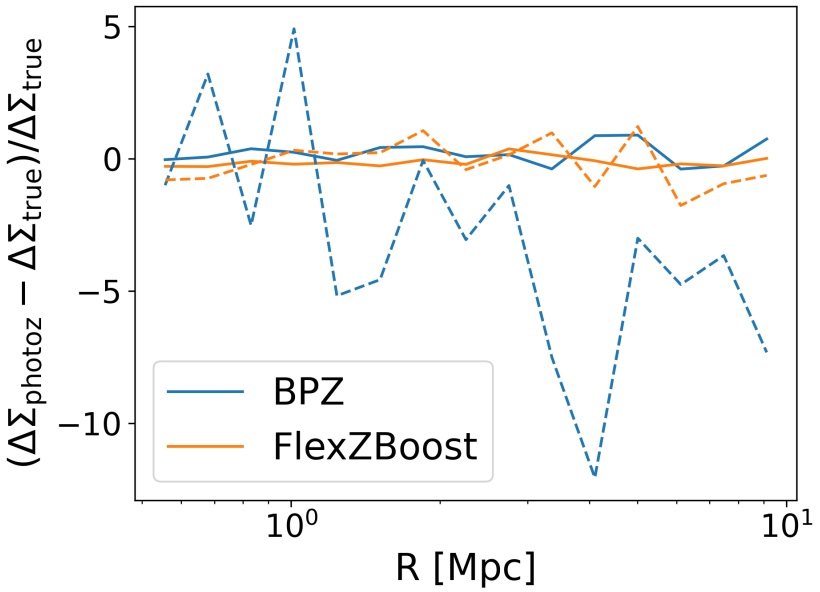

Figure 4.5: Relative difference (in %) between the stacked profiles using photoz's (BPZ: blue, FlexZBoost: orange) and using cosmoDC2 galaxy redshifts. Full lines: cluster redshift $0.2 < z_l < 0.3$, dashed lines: cluster redshift $0.7 < z_l < 0.8$.

with the true dark matter halo masses in Fig. 4.4 (right), over the redshift and cosmoDC2 mass bins[14]. Furthermore, the modeling choices we made will also impact the mass reconstruction, this will be looked at in more detail in Section 4.2.

This first test provides a verification that the masses estimated with weak lensing and the true cosmoDC2 halo masses (here M_{200c} and $M_{\rm FoF}$) are qualitatively well correlated, while some caveats exist. This work contributed to the validation of cosmoDC2 from the cluster weak lensing perspectives.

We also estimate the cluster lensing profiles accounting for the photometric redshift of cosmoDC2 galaxies, obtained with the BPZ (Benítez, 2011) and FlexZBoost (Izbicki & Lee, 2017) algorithms. In this slightly more realistic setup, the photometric redshift uncertainties can be taken into account in the stacked excess surface density profile through the individual lens-source geometric weights $w_{ls}^{\rm geo}$. The systematic effects in photometric redshift estimates will bias the measurement of cluster lensing profiles. In practice, distant galaxies are more sheared than close galaxies and closer galaxies are less sheared. Then if galaxy redshifts are biased low (respectively high), it will induce an overestimation (respectively underestimation) of weak-lensing profiles. Fig. 4.5 shows the relative difference between stacked cluster lensing profiles using the source photometric redshift information and using the true cosmoDC2 redshifts. We consider two different cluster samples within the same mass range but within two different redshift bins (full lines: cluster redshift $0.2 < z_l < 0.3$, dashed lines: cluster redshift $0.7 < z_l < 0.8$). By increasing the redshift of the cluster, we lower the contribution of "closer" background galaxies for the estimation of the lensing profile, which redshifts are usually well reconstructed. So the contribution of more distant galaxies increases, thus accounting relatively for more catastrophic photometric reconstructions that were found using BPZ (see Section 3.3.2). For the low redshift

[14]We note a significant underestimation of the lensing mass compared to the true mass for the redshift bin $[0.5; 0.6]$ and for the (log10)mass bin $[14.2, 14.3]$ (at more than 3σ). We note also a slight underestimation of the weak lensing mass at high mass, which may arise from the possible mass dependence of the signal attenuation region that is not taken into account here. Also, this attenuation effect at fixed redshift may be more important relative to noise, since SNR is higher for high-mass clusters.

cluster sample (full line), BPZ and FlexZboost performances are similar, we find no significant bias with respect to the idealistic case. However, for the high redshift cluster sample (dashed line), the BPZ cluster lensing profile (with no quality cuts) shows significant deviation from the idealistic lensing profiles. We recall that the BPZ redshifts are very pessimistic and restricting to $z_{\rm max} = 1.5$ would make the BPZ profile more compatible with the ideal profile (here, we use $z_{\rm max} = 3$, the maximum galaxy redshift of cosmoDC2).

4.1.2 DC2 object catalog: haloes, measured redshifts, measured shapes

Having worked with cosmoDC2, we now turn to the characterization of the more realistic DC2 object catalog.

In this section, we aim to characterize the lensing around dark matter halos considering galaxies that were detected on images, and for which the shapes were measured using two different algorithms, namely HSM (Hirata & Seljak, 2003) and METACALIBRATION (mCal, Sheldon & Huff (2017)).

As introduced in Section 3.3, the HSM algorithm measures the galaxy χ-ellipticity via the second moment of the surface brightness $Q_{i,j}$ such as

$$\chi^{\rm HSM} = \frac{Q_{11} - Q_{22} + iQ_{12}}{Q_{11} + Q_{22}}. \tag{4.1}$$

The $\chi^{\rm HSM}$ shapes[15] are readily available in the catalog. As mentioned in Chapter 3, HSM shapes need to be calibrated. When working with real data, this calibration is done using simulations. With DC2, we do have access to the true underlying shape, so we have performed a naive calibration that is based on the proximity matching between cosmoDC2 galaxies to DC2 objects (as part of Manon Ramel's PhD work on the DC2 data at LPSC); Galaxies from cosmoDC2 and DC2 objects are first matched using a Friends-of-Friends algorithm (Mao et al., 2021). The match is based on proximity matching of sky coordinates with a linking length of 0.4 arcseconds. After matching, we compute the mean $\epsilon_{\rm HSM}$ in bins of $\epsilon_{\rm cosmoDC2}$ and we fit a linear relation such that $\epsilon_{\rm HSM} \approx 0.85\epsilon_{\rm cosmoDC2}$. We find from this (simplified) approach that the HSM shapes are biased low compared to true cosmoDC2 shapes. We will check the impact of the calibration on stacked weak lensing profiles, by rescaling HSM shapes with the fitted relation.

For METACALIBRATION, we need to estimate the response matrix (as presented in Section 3.3.2), given by

$$\langle {\rm R} \rangle = \langle {\rm R}_\gamma \rangle + \langle {\rm R}_S \rangle. \tag{4.2}$$

To do so, mCal provides several shape catalogs with each sheared component of the ellipticity of a galaxy when applying a shear positively and negatively along the two directions to the galaxy image. For a given sample of selected galaxies, we then computed the average shear response element $({\rm R}_\gamma)_{ij}$ such as

$$({\rm R}_\gamma)_{ij} = \frac{\langle \epsilon_i^+ \rangle - \langle \epsilon_i^- \rangle}{\Delta\gamma_j}, \tag{4.3}$$

where i is the component of the ellipticities ϵ^+ and ϵ^- respectively sheared along the positive (+) and negative (-) j-direction. The shear amplitude is constant along the two sky directions (i.e., $\Delta\gamma_j = \Delta\gamma = 0.02$). The ${\rm R}_S$ matrix denotes the response of the selection based on sheared ellipticity quantities applied to un-sheared ellipticities. When applying a selection cut S_j^+ and

[15] We converted the χ shapes to the ϵ ellipticities such as $\epsilon = \chi^{\rm HSM}(1 + (1 - |\chi^{\rm HSM}|^2))^{-1/2}$.

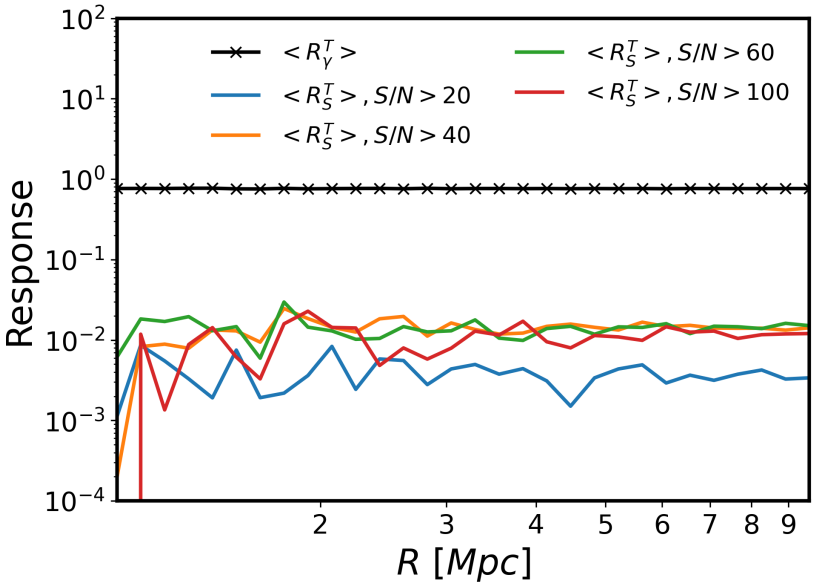

Figure 4.6: The two components of the mCal tangential response matrix $\langle \mathrm{R}_\gamma^T \rangle$ (black), and $\langle \mathrm{R}_S^T \rangle$ (colors) as measured from DC2. The different colors correspond to different S/N cuts on sheared ellipticities.

S_j^- on respectively the positive and negative sheared versions of the galaxy image, the selection response is given by

$$(\mathrm{R}_S)_{ij} = \frac{\langle \epsilon_i \rangle^{S_j^+} - \langle \epsilon_i \rangle^{S_j^-}}{\Delta \gamma_j}. \tag{4.4}$$

To estimate the stacked lensing profiles around dark matter halos, it is necessary to express the response R in the tangential reference frame relative to the cluster center. We use the method described in McClintock et al. (2019), where the individual tangential shear response R_γ^T for each background galaxy is computed via

$$\mathrm{R}_\gamma^T = \mathrm{R}_{\gamma,11} \cos(2\varphi)^2 + \mathrm{R}_{\gamma,22} \sin(2\varphi)^2 + (\mathrm{R}_{\gamma,12} + \mathrm{R}_{\gamma,21}) \cos(2\varphi) \sin(2\varphi), \tag{4.5}$$

with φ is the polar angle from the galaxy relative to the cluster center relative to the x-axis. The tangential average selection response is approximated by

$$\langle \mathrm{R}_S^T \rangle \approx \frac{1}{2} Tr(\langle \mathrm{R}_S \rangle). \tag{4.6}$$

The estimator of the excess surface density estimator using mCal shapes $\{\epsilon_{ls}^+\}$ can be re-written as

$$\Delta\Sigma(R) = \frac{1}{\sum_{l=1}^{N_l} \sum_{s=1}^{N_{ls}} w_{ls} (R_{\gamma,ls}^T + \langle R_S^T \rangle)} \sum_{l=1}^{N_l} \sum_{s=1}^{N_{ls}} w_{ls} \Sigma_{\mathrm{crit},ls} \epsilon_{ls}^+. \tag{4.7}$$

For a given redshift-mass stack, we show in figure Fig. 4.6 the average tangential shear response (black) given by

$$\langle \mathrm{R}_\gamma^T \rangle = \frac{1}{\sum_{l=1}^{N_l} \sum_{s=1}^{N_{ls}} w_{ls}} \sum_{l=1}^{N_l} \sum_{s=1}^{N_{ls}} w_{ls} \ (R_\gamma^T)_{ls} \tag{4.8}$$

and the average selection response $\langle \mathrm{R}_S^T \rangle$ (colors) when applying different signal-to-noise ratio cuts. We find that the shear response is constant with radius and is roughly two orders of mag-

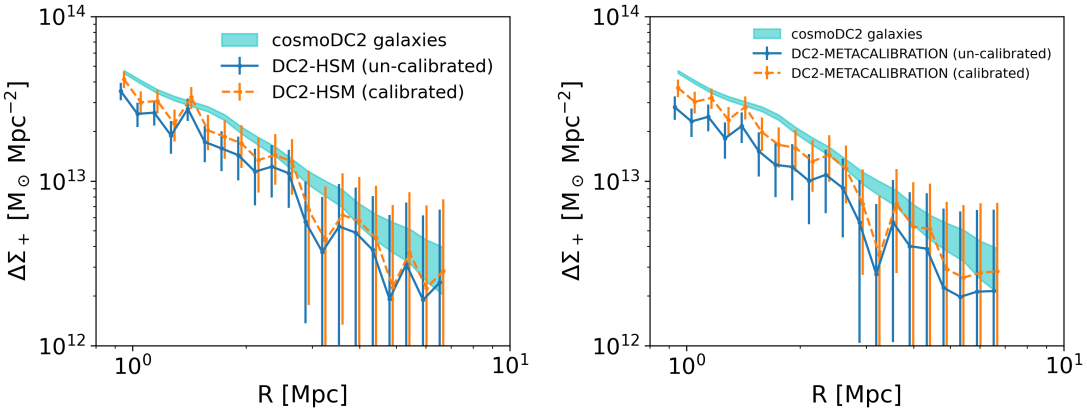

Figure 4.7: Left: stacked excess surface density profile around mass and redshift selected dark matter halos, first by using cosmoDC2 background galaxies (cyan), HSM un-calibrated (blue) and calibrated (orange) shapes. Right: stacked excess surface density profile using cosmoDC2 background galaxies (cyan), and using mCal un-calibrated shapes (blue) and by applying the shear self-calibration (orange).

nitude higher than the selection response. The shear response scatters around 0.75 (variations too small to see on the figure).

We repeated the estimation of the stacked excess surface density this time by using HSM and mCal shapes. We also consider the photometric redshifts obtained from the BPZ run on the measured DC2 magnitudes. We extract the background object catalog by applying the selection $\langle z \rangle > z_l + 0.1$, where $\langle z \rangle$ is the mean photometric redshift per galaxy. The lensing profile might be affected by the contamination from un-sheared member galaxies that have been misidentified as background galaxies, although the large inner radial cut we apply to avoid the unphysical attenuation region of the signal likely mitigates that effect (Varga et al., 2019). We show in Fig. 4.7 (left) a stacked shear profile for a given mass and redshift bin, using the different shape catalogs. As expected, we find the un-calibrated HSM ellipticities (blue) to be biased low compared to the ideal cosmoDC2 profile (cyan), because the uncalibrated HSM ellipticities are usually underestimated (we have found a multiplicative bias of 0.85). After our simple cosmoDC2-based calibration of HSM shapes, we still do not find that the calibrated profile (orange) is compatible with the ideal case. We are not completely surprised by this result, since this methodology is quite simple and does not cover all the complexity of HSM shape calibration (see Mandelbaum et al. (2018)). This result confirms that the calibration of HSM shapes is a more complicated task and our simplified method needs to be improved. Since HSM is not the default shape measurement method that will be applied to LSST data, we did not investigate this further.

The Fig. 4.7 (right) shows the results obtained with mCal shapes. More surprisingly, we also find a bias after calibration (orange) compared to using true cosmoDC2 shapes. At the time of this work, we did report the problem to the Weak Lensing WG. These results were noted in parallel in other weak lensing analyses with DC2 data. To our knowledge, the problem is yet to be fixed. This unforeseen difficulty prevented us from fully exploring the DC2 data to constrain the mass-richness relation with LSST-like catalogs.

The work presented in this section has led to the release of the DESC note [DN1] and focused on characterizing the lensing shear around cosmoDC2 dark matter halos using the DC2 shape catalogs available. Moreover, I participated to the study of lensing magnification in cosmoDC2 described in [DN3] that I will not detail in this manuscript. The lensing magnification produces

a shift in the observed galaxy position due to the deflection of light rays coming from distant sources, as well as magnifying the observed galaxy fluxes. In this note, we estimated stacked weak lensing masses from the magnification signal around cosmoDC2 dark matter halos.

4.2 $P(M|\lambda)$ scaling relation in DC2

Having characterized the DC2 datasets and their limitation from the cluster weak lensing perspectives, we now turn our attention to the mass-richness relation of the DC2 data. While we know from the above that we will not be able to run an analysis completely representative of LSST-like data, we can nonetheless gain valuable experience in modeling and statistical approaches, by pushing this further. The idea is the same as before but now the binning is performed in redshift and richness bins as redMaPPer provides a richness for each identified cluster. For that, we measure the stacked profiles around redMaPPer clusters[16]. We consider the redshift bins with edges $[0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1]$ and the richness bin edges $[20, 35, 70, 100, 200]$. The Fig. 4.8 shows the corresponding stacked weak lensing profiles. As expected, the amplitude of the shear profiles increases with richness, indicating a positive correlation between the cluster mass and richness. The work in this section aims at quantifying this correlation by using several statistical methods.

We first describe the formalism of the $P(M|\lambda)$ relation, before moving to estimating its fiducial form in Section 4.2.2. We then estimate lensing masses in Section 4.2.3 and we study the weak lensing cluster scaling relation in Section 4.2.4. Finally, we present in Section 4.2.5 a methodology to infer the cluster scaling relation based on the cluster stacked profiles directly without fitting stacked masses as an intermediate step.

4.2.1 Backward modeling of the cluster scaling relation

At fixed richness, the cluster mass is scattered due to complex physics occurring in the galaxy cluster environment, as well as observational systematics in measuring the effective member galaxies within a cluster. Generally, the relation between the individual cluster mass and richness is given by the probability function $P(M|\lambda)$[17] for which we adopt a log-normal model. Using this model, we use the parametrization

$$P(M|\lambda) = \frac{1}{M\sqrt{2\pi\sigma^2}} \exp -\frac{1}{2}\left(\frac{\ln M - \mu}{\sigma}\right)^2 . \tag{4.9}$$

We define that the mean mass at given richness and redshift follows (Simet et al. (2017); McClintock et al. (2019); Melchior et al. (2017))

$$\langle M|\lambda, z\rangle = M_0 \left(\frac{1+z}{1+z_0}\right)^{G_z} \left(\frac{\lambda}{\lambda_0}\right)^{F_\lambda} , \tag{4.10}$$

[16]Several versions of the redMaPPer catalog were published within the DESC collaboration, the version 0.7.5 was used in the DESC notes, whereas for this section we used the version 0.7.8.

[17]We point out that this section focuses on the $P(M|\lambda)$ relation instead of the $P(\lambda|M)$ modeling used in e.g. Murata et al. (2019); Sunayama et al. (2023) and that will be considered in the next Section 4.3. The $P(M|\lambda)$ distribution can be obtained from $P(\lambda|M)$ using the Bayes theorem, and this correspondence is studied in Evrard et al. (2014). It is possible to switch from one to the other by considering some approximations of the underlying halo mass function. The $P(M|\lambda)$ is refereed to as *backward* modeling by Murata et al. (2019), whereas $P(\lambda|M)$ is called *forward* modeling. We will discuss their differences in Section 4.3 and why we favor the forward modeling of the scaling relation for the cosmological inference in the DC2 simulations.

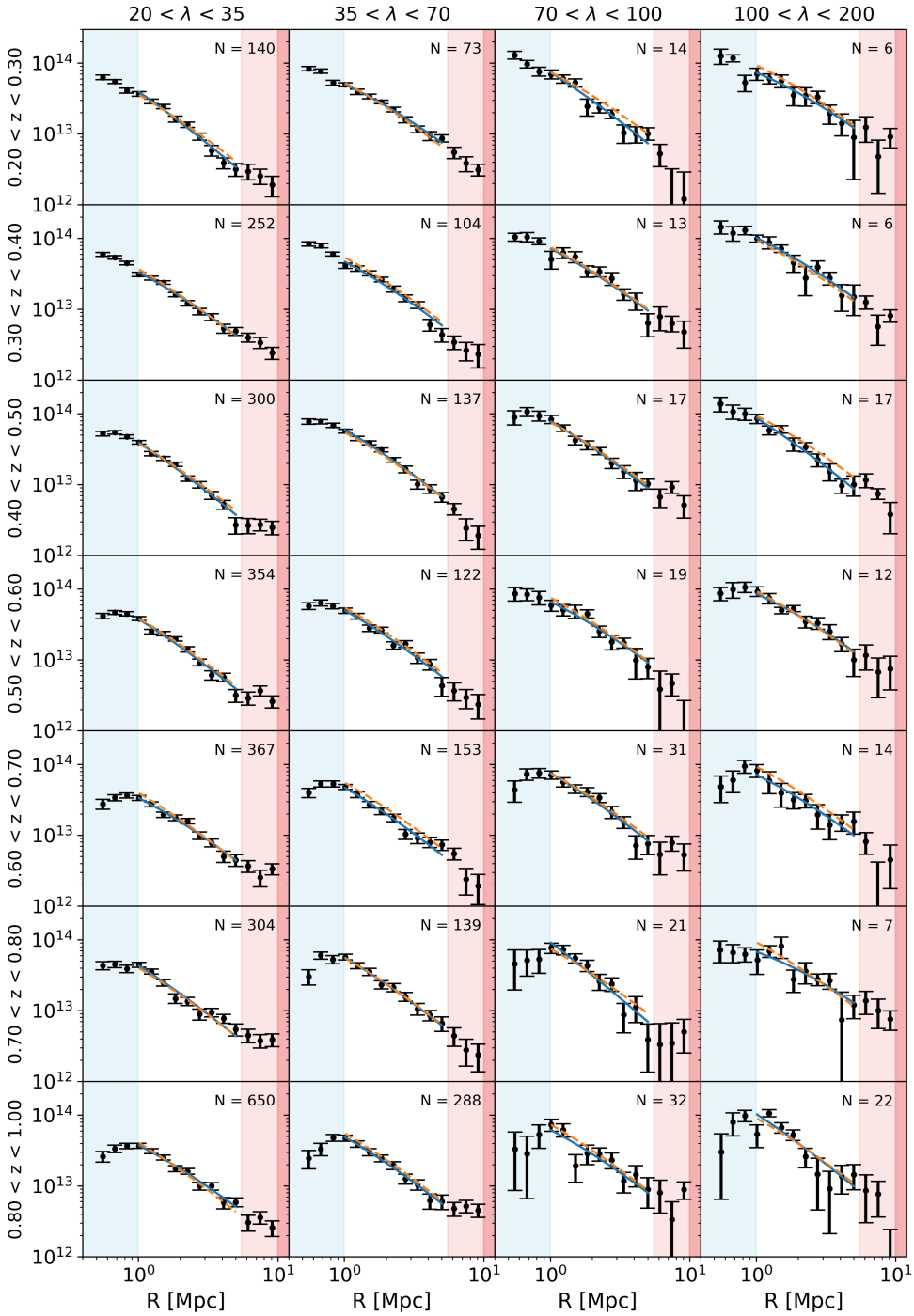

Figure 4.8: Stacked excess surface density profiles of redMaPPer detected cluster in bins of richness and redshift. The blue-shaded region corresponds to $R < 1$ Mpc, the light red region to $R > 5$ Mpc range, and the dark red region to $R > 10$ Mpc.

where M_0 is the normalization factor, G_z and F_λ quantify respectively the redshift and richness dependency. Using this formalism, we have that $\mu = \langle \ln M|\lambda, z\rangle$ given by (Rozo et al., 2009)

$$\mu = \ln\langle M|\lambda, z\rangle - \frac{1}{2}\sigma^2 \tag{4.11}$$

where the variance of the logarithmic mass[18] $\ln M$ given by $\sigma^2 = \mathrm{Var}[\ln M|\lambda]$ at fixed redshift-richness in Eq. (4.9) has two contributions: the first denotes the intrinsic variance of the distribution, the second is due to the Poisson variance of the measured richness (Evrard et al., 2014; Simet et al., 2017). It is given by

$$\sigma^2 = \sigma^2_{\ln M|\lambda} + \frac{F^2_\lambda}{\lambda}. \tag{4.12}$$

In doing so, the scatter in the mass-richness relation is dominated by Poisson shot noise at low richness, whereas it is dominated by the intrinsic scatter at high richness. Following this formalism, the log-normal probability distribution is fully determined by the combination of μ and σ, described by the 4 parameters $(\ln M_0, F_\lambda, G_z, \sigma_{\ln M|\lambda})$.

4.2.2 Fiducial cluster scaling relation

Before turning to the analysis using weak lensing to constrain the scaling relation, the latter can be inferred by geometrically matching each redMaPPer-detected galaxy cluster with its corresponding dark matter halo at the same location in the simulation (namely matching the pair of two sky coordinates and the redshift for each cluster/halo). This is of course only possible because we are using simulations that will allow us to determine the "fiducial" $P(M|\lambda)$ relation of DC2 redMaPPer clusters. Each redMaPPer cluster with richness λ_k is then assigned with a "true" spherical overdensity mass M_k.

To perform this match between the two catalogs, we used the DESC code ClEvaR[19] (Aguena et al., in prep.) that provides a user-friendly software to perform this task. Two different metrics are available to spatially match two catalogs. The first is based on the proximity of the objects in the two catalogs, i.e. the center of an object (redMaPPer cluster) within the catalog A is associated with the nearest object (cosmoDC2 dark matter halo) within the catalog B. The second method is based on membership proximity: for each pair of redMaPPer cluster and cosmoDC2 halo, we evaluate the proximity between their respective member galaxy population. The metric is not selecting pairs based on the distance between cluster and halo centers, but based on the "effective" distance between their galaxy population. We use the membership-matched catalog that was built by DESC collaborators using respectively the redMaPPer member galaxies for each cluster and cosmoDC2 galaxies flagged as "members" for each cosmoDC2 dark matter halo. Membership matching has been shown more robust than proximity matching by our DESC colleagues: it is sensitive to the spatial distribution of the redMaPPer member galaxies, it accounts for more systematic effects of redMaPPer cluster finder than proximity matching, and the output matched catalog is more robust to the stochasticity of each cluster member galaxy distribution. The catalog was obtained considering $M_{200c} > 10^{13} M_\odot$ for the dark matter halo catalog and $\lambda > 5$ for the redMaPPer cluster catalog.

The Fig. 4.9 shows the joint distribution of individual dark matter halo masses and redMaPPer richnesses. We see a strong correlation, but also a significant scatter between mass and richness[20].

[18] We can pass from a definition to another by considering $\mathrm{Var}[M|\lambda] = \langle M|\lambda\rangle^2[\mathrm{e}^{\sigma^2} - 1]$.

[19] The code is publicly available here `https://github.com/LSSTDESC/ClEvaR`.

[20] We compared this matched catalog represented in Fig. 4.9 with the proximity matching output with ClEvaR, and also by using the `astropy` Python package (Astropy Collaboration et al., 2022) (which allows to match spatially two datasets in a very efficient way, each data set is associated to a Python object `SkyCoord`, and the

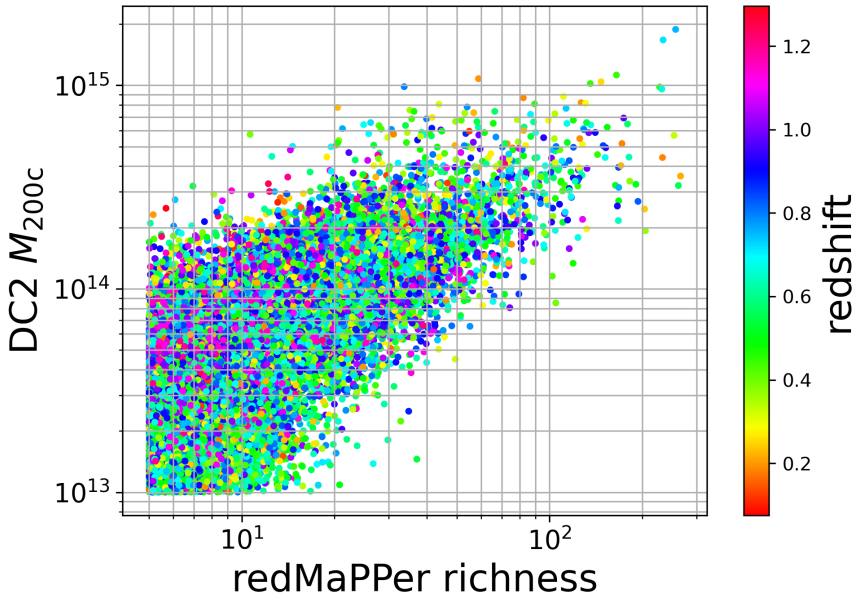

Figure 4.9: Individual cosmoDC2 dark matter halo masses with respect to redMaPPer richnesses from the matched catalogs that contain 40,000 matches down to $M = 10^{13} M_\odot$ and $\lambda = 5$. The points are color-coded with respect to cluster redshift.

From the cosmoDC2-redMaPPer matched catalog we can infer a "fiducial" scaling relation. We consider the clusters with $\lambda > 20$, $M > 10^{13} M_\odot$ and with $0.2 < z < 1$. The posterior distribution of the scaling relation parameters $\{\theta_\alpha\}$ is given by

$$\mathcal{P}(\{\theta_\alpha\}|\{M_k\}) \propto \mathcal{L}(\{M_k\}|\{\theta_\alpha\})\pi(\{\theta_\alpha\}), \tag{4.13}$$

where we consider each matched object independently, such as the likelihood for all individual masses is given by

$$\mathcal{L}(\{M_k\}|\{\theta_\alpha\}) = \prod_{k=1}^{N_{\rm tot}} P(M_k|\lambda_k, z_k). \tag{4.14}$$

In the above equation, $P(M_k|\lambda_k, z_k)$ is the probability density function for the mass M with richness λ and redshift z in Eq. (4.9)[21] for the k-th matched object in the CleVaR catalog. We use positive flat prior for $\sigma_{\ln M|\lambda}$ and flat priors $[0, 2]$ for F_λ, $[-2, 2]$ for G_z, $[13, 16]$ for $\log_{10} M_0$. We recap in the table Table 4.1 the values of the recovered parameters. These values will serve as a reference against which we will compare the results obtained using weak lensing, in the following section.

4.2.3 Stacked cluster lensing masses and impact of analysis choices

For a stack of clusters selected in a richness and redshift bin, we use the estimated mean excess surface density profiles measured in Fig. 4.8. For a given modeling choice, we reconstruct the mass posterior by considering a Gaussian likelihood with a covariance matrix estimated by

two data sets can be matched spatially with the function `match_coordinates_sky`). We found similar results, but we will use the membership-matched catalog in this section.

[21] We use truncated Gaussians (we have recomputed their normalizations) to account for the limit mass in the fitting procedure (see e.g. Abdullah et al. (2023) that proposed a methodology to find the best limit richness for a given low mass cut).

Parameters	Posterior mean
$\log_{10} M_0$	$14.147^{+0.003}_{-0.004}$
F_λ	$0.797^{0.011}_{-0.021}$
G_z	$-0.277^{0.054}_{-0.042}$
$\sigma_{\ln M\|\lambda}$	$0.467^{0.005}_{-0.005}$

Table 4.1: Posterior means for the four parameters of the scaling relation in Eq. (4.9), by using individual dark matter halo masses and redMaPPer richnesses.

Systematic effect	lists	Name
[I] - Modeling	[Ia] Halo model (1-halo)	NFW (1h - free concentration)
		Einasto (1h - free concentration)
		Hernquist (1h - free concentration)
	[Ib] $c_{\rm NFW}(M)$ (1-halo)	NFW (1h - Diemer15)
		NFW (1h - Prada12)
		NFW (1h - Bhattacharya13)
		NFW (1h - Duffy08)
	[Ic] One+two halo terms	NFW (1h - free concentration + 2h)
		NFW (1h - Diemer15 + 2h)
[II] Observational	[IIa] Photometric redshifts	NFW (1h - Diemer15) - BPZ
		NFW (1h - Diemer15) - FlexZBoost
	[IIb] Shape measurement	HSM
		METACALIBRATION

Table 4.2: Table of the different analyses in this section. The first horizontal block corresponds to the different study cases to test the impact of modeling choices (impact of the halo model, the concentration-mass relation, the two-halo regime). The second horizontal block corresponds to the impact of observational measurements (photometric redshifts, galaxy shape measurement).

bootstrap re-sampling. The stacked weak lensing masses are obtained from the mean of each mass posterior.

Such inference relies on some assumption about the underlying halo model that is chosen to predict the stacked cluster signal. The mass estimates are also affected by a series of observational systematic effects, that impact the local measurement of the weak gravitational lensing shear. All of these effects will affect the reconstruction of the stacked mass, thus the mean scaling relation parameters and, in the end, will impact the reconstruction of the cosmological parameters. In this section, we aim to use the methodology presented above to test the impact of several modeling and observational systematic effects on the mean scaling relation parameters. This work led to the DESC refereed note [DN2], and in this section, we present the different study cases that are summed up in Table 4.2. We detail below the different analyses that we considered.

We first test the impact of modeling choices of the lensing signal (first block [I] in Table 4.2). The weak lensing excess surface density is derived from the surface density, that is directly defined as the projection of the three-dimensional matter density field around the cluster. Any change in the dark matter density profile of the lens induces a modification of the weak lensing signal. Therefore, choosing a particular model of the mass distribution instead of another may impact the cluster lensing mass, and have a non-negligible impact on the mass-richness relation.

Here, we compare three different density parameterizations to test the impact of such matter distribution on the weak-lensing mass richness relation. These are the NFW (Navarro et al., 1997), Einasto (Einasto, 1965) and Hernquist (Hernquist, 1990) profiles, evoked in Chapter 2. The NFW and Einasto profiles are common parameterizations of the dark matter mass distribution around the galaxy cluster and are found to describe accurately the matter content in N-body simulations of collision-less particles, over a broad range of halo masses and redshifts. The Hernquist model was used for instance in Buote & Lewis (2004); Sanderson & Ponman (2009) to model the mass density of galaxy clusters observed in X-rays, but has not been extensively used in the literature in a weak lensing perspective.

In addition to the cluster mass M_{200c}, these three profiles have an extra concentration parameter c_{200c}. The NFW, Einasto and Hernquist concentrations cannot easily be compared since they do not rely on the same assumption of the halo matter content. However, the reconstructed weak-lensing mass M_{200c}, that is defined in terms of overdensity relative to the critical density, has no link to the concentration. Then, we can compare the masses obtained by fitting a $\Delta\Sigma$ profile (testing the impact of the halo model is referred to as [Ia] in Table 4.2). We will consider the 1-halo regime, thus restricting the radial range from 1 to 5 Mpc (by avoiding the blue shaded regions in Fig. 4.8).

Moreover, various studies based on cosmological simulations indicate that the concentration and the cluster mass are correlated, and this correlation can be quantified via a concentration-mass relation $c(M)$. The use of a $c(M)$ relation allows us to reduce the number of free parameters in the weak lensing mass reconstruction, and then get more precise constraints of the halo mass. Various parameterizations of the $c(M)$ relation exist in the literature. We will compare the results using the relations from Diemer & Kravtsov (2014); Duffy et al. (2008); Prada et al. (2012); Bhattacharya et al. (2013). We also want to compare to the simplest case where we have no assumption on the concentration, i.e., by letting the concentration as a free parameter in the fit.

Finally, as we restricted the radial range to a maximum of 5 Mpc, we did not account for the 2-halo term contribution in the modeling of the shear profile. Some valuable information on cluster properties (via the halo bias) may exist in the two-halo regime. Cluster mass and halo bias are positively correlated, meaning that more massive clusters tend to be more biased with respect to the underlying matter density field. However, this regime has a lower signal-to-noise ratio, since it is fainter and has an additive variance to the standard shape noise arising from large-scale structure fluctuations. We explore the use of the 2-halo term (referred to as [Ic] in Table 4.2) considering the radial range up to 10 Mpc[22].

After considering these modeling options, we test the impact of observational measurements on the cluster mass estimates (second block of the Table 4.2), namely accounting for the photometric redshifts of cosmoDC2 derived with BPZ (Benítez, 2011) and FlexZBoost (Izbicki & Lee, 2017) (as presented in Section 4.1), and the shape measurements by HSM (Hirata & Seljak, 2003) and METACALIBRATION (Sheldon & Huff, 2017) of DC2 objects (respectively referred to as [IIa] and [IIb] in Table 4.2). Both effects may lead to a misinterpretation of the stacked cluster lensing signal by respectively introducing a bias in the observed galaxy distances with respect to the observer (photometric redshift) and a coherent bias in galaxy shape estimates (we have addressed these two aspects in Section 4.1). We already know, from Section 4.1 that the calibration of shape measurements, be it HSM of mCal, is problematic and will yield biased

[22]The contribution of the one-halo term is still significant at large radius for the radial range we have considered. Most stacked analyses based on the one-halo and two-halo regimes consider the lensing signal up to 20-30 Mpc, to take full advantage of the information within the two-halo term and to disentangle with the one-halo term contribution. At the time of this work and the computational limitations in extracting background sources at large radii, we restricted to 10 Mpc.

results. However, it is still interesting to look at to see how the errors are impacted by the shape measurements.

The different lensing mass estimates for the cluster stack within the redshift bin $[0.2, 0.3]$ (top) and $[0.5, 0.6]$ (bottom) and within different richness bins (from left to right, represented by the different markers) are shown in Fig. 4.10. Each row corresponds to an analysis listed in Table 4.2 (red: [Ia], green: [Ib], orange: [Ic], blue: [IIa]). We will focus on the [IIb] in the next section, corresponding to using the DC2 object catalog, when addressing the inference of the scaling relation.

Whatever the richness bin, we find that the choice of the halo model ([Ia] in Table 4.2) coupled with free concentration (red points) does not impact strongly the reconstruction of the lensing mass. This is understandable as the behavior of each profile differs strongly in the innermost regions, that we avoid due to the ray-tracing attenuation, and at large radius, which we are not strongly sensitive to since we restricted the radial range up to 5 Mpc.

When imposing a concentration relation and using a NFW profile ([Ib] in Table 4.2, green points) the error on the mass is expectedly smaller since we only fit one halo parameter. Moreover, we see some small shift compared to the free concentration analysis, with no clear trend with richness, but still compatible at the 1σ level. Again, since we restricted the radial range above 1 Mpc, the lensing signal is not strongly affected by a change in concentration, giving compatible results. However, the tension between the free-concentration and the $c(M)$-relation cases would become more statistically significant if the bins were more populated, which will be the case with LSST data.

Including the two-halo regime in the mass fitting [Ic] is shown with the orange points ([by considering $R \in [1, 10]$ Mpc). Using a concentration-mass relation, we find that the lensing mass estimates are slightly lower than only using the one-halo term ($R \leq 5$ Mpc), but still compatible at $1 - 2\sigma$. The tension is reduced when letting the concentration as a free parameter.

Finally, considering stacked excess surface density profiles estimated using the photometric redshift information of background sources, we find that the lensing mass estimates are not strongly affected by the choice of the photometric reconstruction method (BPZ or FlexZBoost, last two lines). Indeed they remain compatible with the true case. Interestingly, there are no significant differences between the mass estimates using BPZ or FlexZBoost, despite the drastic differences in the algorithm performances.

When comparing the four different redshift bins, we see that the results are sensitively equivalent, with some scatter that does not affect the 1σ compatibility between the different analyses. However, the different methods may induce persistent biases that will affect the reconstruction of the scaling relation. This is what we explore in the next section.

4.2.4 Propagation to $P(M|\lambda)$ relation

We show in Fig. 4.11 the stacked lensing mass estimates in each stack, as a function of the mean richness in the bin[23] for the analysis [Ia]. The lensing mass error bar corresponds to the standard deviation of the mass posterior. For the different fitted masses, we show in Fig. 4.8 (blue) the corresponding best fit excess surface density profiles.

We have estimated the stacked weak lensing masses $\widehat{M}_m$ from the cluster lensing profiles in the previous section and we now aim to use them to constrain the cluster scaling relation. The scaling relation parameters will be obtained by considering the stacked mass likelihood

[23] We have inferred the masses using a NFW profile and letting free the concentration.

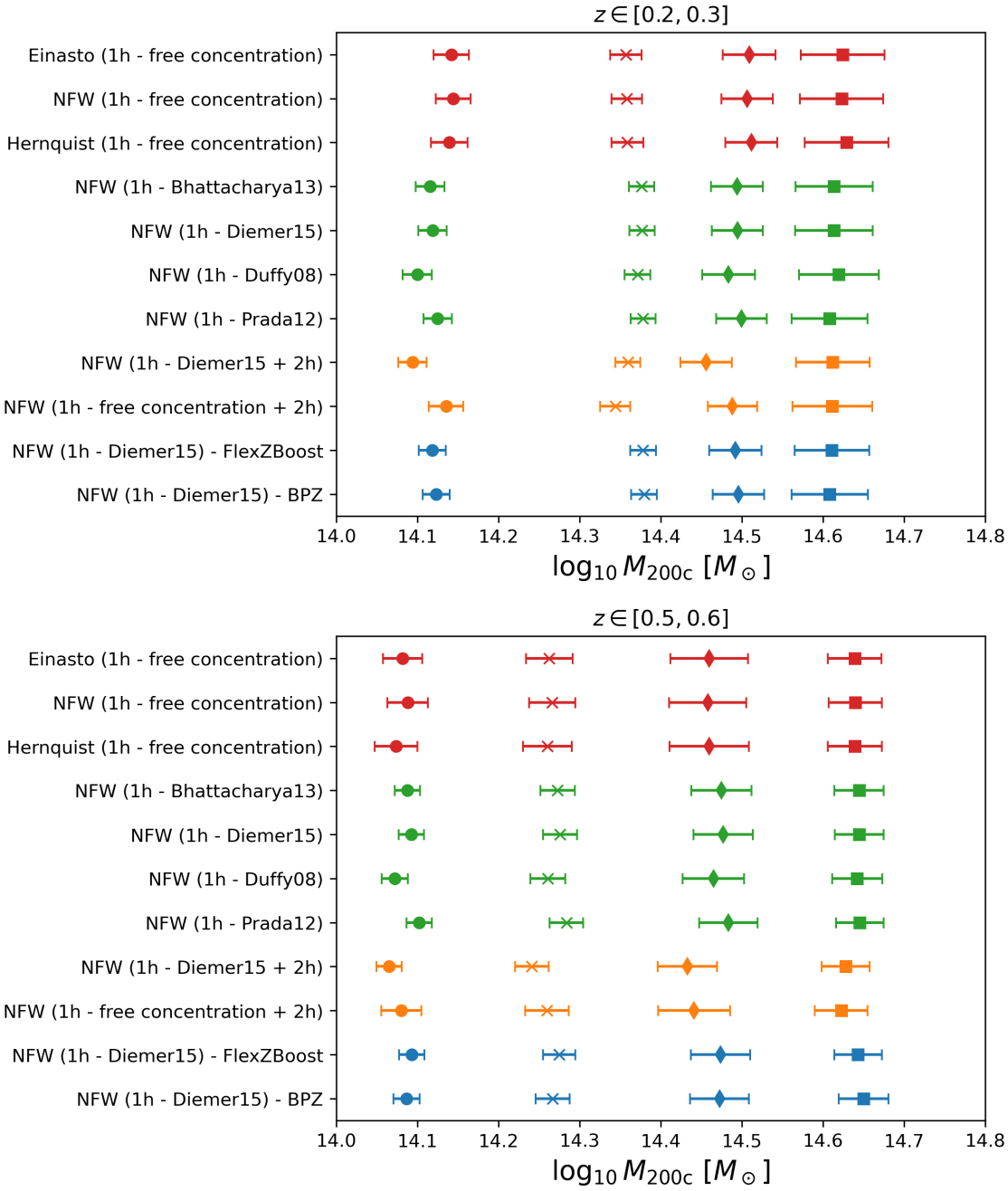

Figure 4.10: Stacked lensing mass estimates in the redshift bins $[0.2, 0.3]$ (top) and $[0.5, 0.6]$ (bottom), and in the four richness bins (corresponding to the different marker styles). The different colors correspond to the analyses listed in Table 4.2.

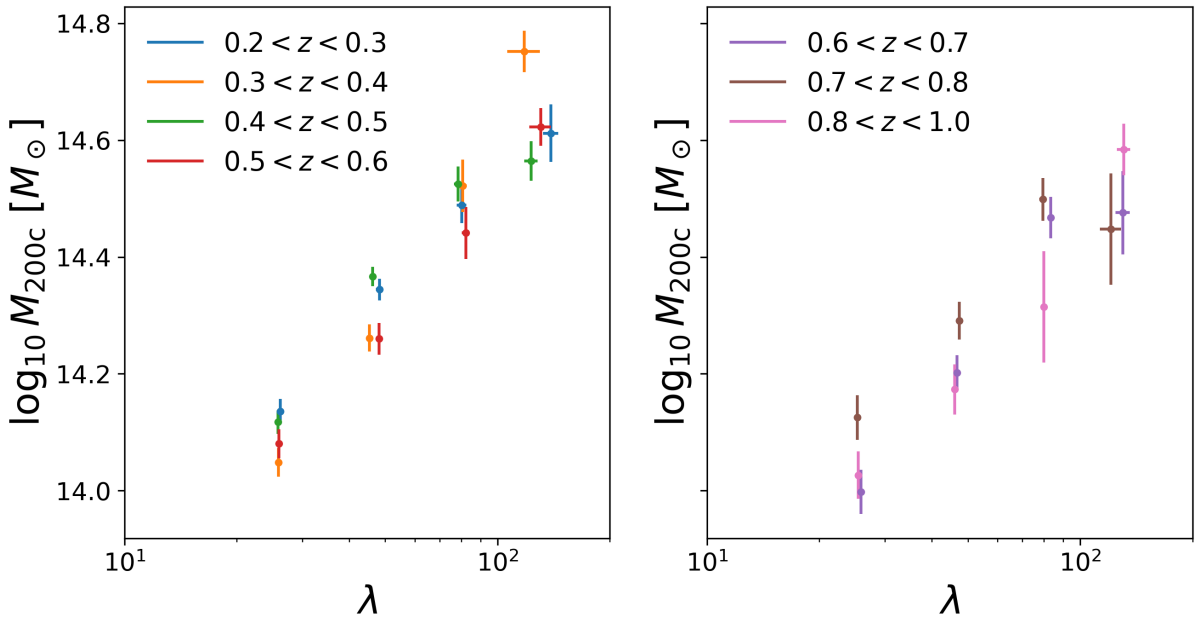

Figure 4.11: Stacked lensing mass estimates as a function of the mean richness in bins for the inference setup [Ia]. The different colors correspond to the redshift bins. The error bars on the x-axis correspond to the errors on the mean richness in each bin.

$\mathcal{L} \propto \exp(-\frac{1}{2}\chi^2)$ with (see e.g. Melchior et al. (2017); McClintock et al. (2019))

$$\chi^2 = \sum_{m,n=1}^{c} (\log_{10} \widehat{M}_m - \log_{10}\langle M_m\rangle)^T \mathbf{C}^{-1}_{mn} (\log_{10} \widehat{M}_n - \log_{10}\langle M_n\rangle), \tag{4.15}$$

where $\mathbf{C}^{-1}$ is the (inverse) covariance of the stacked log-masses (inferred by weak lensing), and $\langle M_m\rangle$ corresponds to the mean mass prediction within the stack with index m obtained in Section 4.2.3, and we consider $\langle M_m\rangle = \langle M|\lambda_m, z_m\rangle$ given in Eq. (4.10). In this methodology, it is not straightforward to take into account the mass scatter at fixed richness $\sigma_{\ln M}$ in the cluster scaling relation; and how to alleviate this difficulty. This aspect will be discussed in more detail in Section 4.2.5.

We show in Fig. 4.12 the mean scaling relation parameters $\log_{10} M_0$, G_z and F_λ of Eq. (4.10) using a MCMC approach. As in Fig. 4.11, the different rows correspond to the different analysis setups summed up in Table 4.2. We also show the constraints we obtain by fitting the scaling relation to the matched catalog between redMaPPer clusters and cosmoDC2 dark matter halos (in black) as detailed in Section 4.2.2.

We see that most of the lensing analyses give a somewhat underestimated normalization $\log_{10} M_0$ (left panel) with large variation between the modeling choices. The normalization seems more sensitive to the modeling choices than the mass reconstruction itself.

We first see that the redshift evolution of the scaling relation (G_z, middle panel) is rather robust, stable between all weak lensing analyses and compatible with the fiducial one. The weak lensing constraints on G_z are also slightly underestimated compared to the fiducial value. Similar conclusions are reached by looking at the F_λ parameter (right panel).

As we mentioned in the lensing methodology, the lensing signal is affected by the non-physical attenuation due to the ray-tracing resolution, leading to a misinterpretation of the stacked lensing signals. This effect still may have a small impact on the lensing signal for $R > 1$ Mpc, that

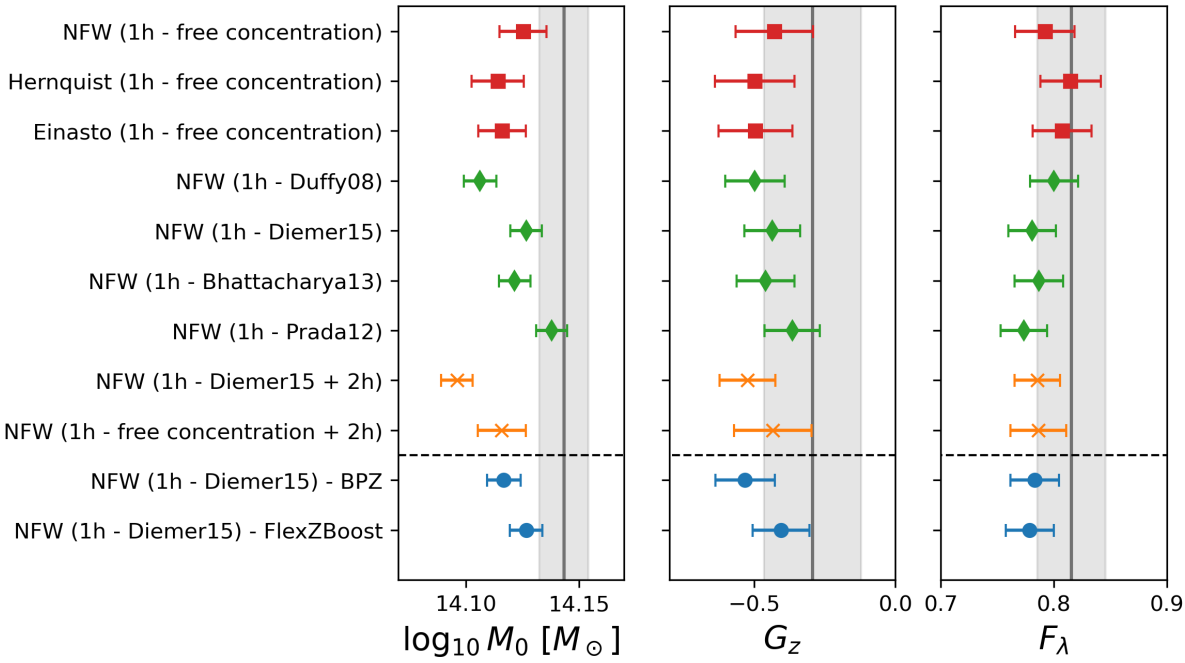

Figure 4.12: Best fit parameters of the mean mass scaling relation. The different colors correspond to the analyses listed in Table 4.2. The shaded regions correspond to the constraints of the fiducial scaling relation detailed in Section 4.2.2. Above the dashed line: Impact of modeling choices. Below the dashed line: impact of photoz codes.

we did not consider in the analysis. Moreover, the miscentering contribution was neglected in the modeling, since we considered the $R > 1$ Mpc region. But this effect may still affect the lensing signal beyond 1 Mpc to some extent (for instance depending on the signal-to-noise ratio of the lensing measurement), and may slightly lower lensing masses if not accounted for (the miscentering tends to smooth the lensing signal at small scales). Finally, we also mention that in this section we modeled the lensing signal with a single-lens approach, whereas the measured lensing signal is the result of multiple individual lensing contributions, as discussed in e.g. Simet et al. (2017); McClintock et al. (2019) (we will also discuss this aspect in Section 4.2.5).

The two bottom lines of Fig. 4.12 show the impact of using photometric redshifts instead of true ones. Using BPZ does not modify M_0 and G_z, while F_λ remains stable. We see no differences in using FlexZBoost or true redshift (but we recall that FlexZBoost is optimistic). Although we knew from Fig. 4.7 that the shear calibration we apply is not sufficient, we nonetheless perform the exercise of checking the impact on the $P(M|\lambda)$ parameters. We show in Fig. 4.13 the posterior distribution of the scaling relation parameters when considering DC2 object shape catalogs, then comparing the performance of HSM and METACALIBRATION[24]. The uncalibrated cases are represented with the unfilled contours. The uncalibrated results for both HSM and mCal are strongly biased in all parameters. The calibrated results (same colors, filled contours) improve the situation but are still far from compatible with the fiducial or "ideal" weak lensing case. It is however interesting to note the increase of the errors when using measured shapes instead of the true ones (errors are ~ 3 times larger).

The series of caveats of the DC2 dataset (lensing shear attenuation, issue with mCal run) have

[24]For this work, we used an older version (0.7.5) of the redMaPPer catalog, we see that the ideal weak lensing constraints are in statistical agreement with the fiducial constraints from the 0.7.8 version of the catalog.

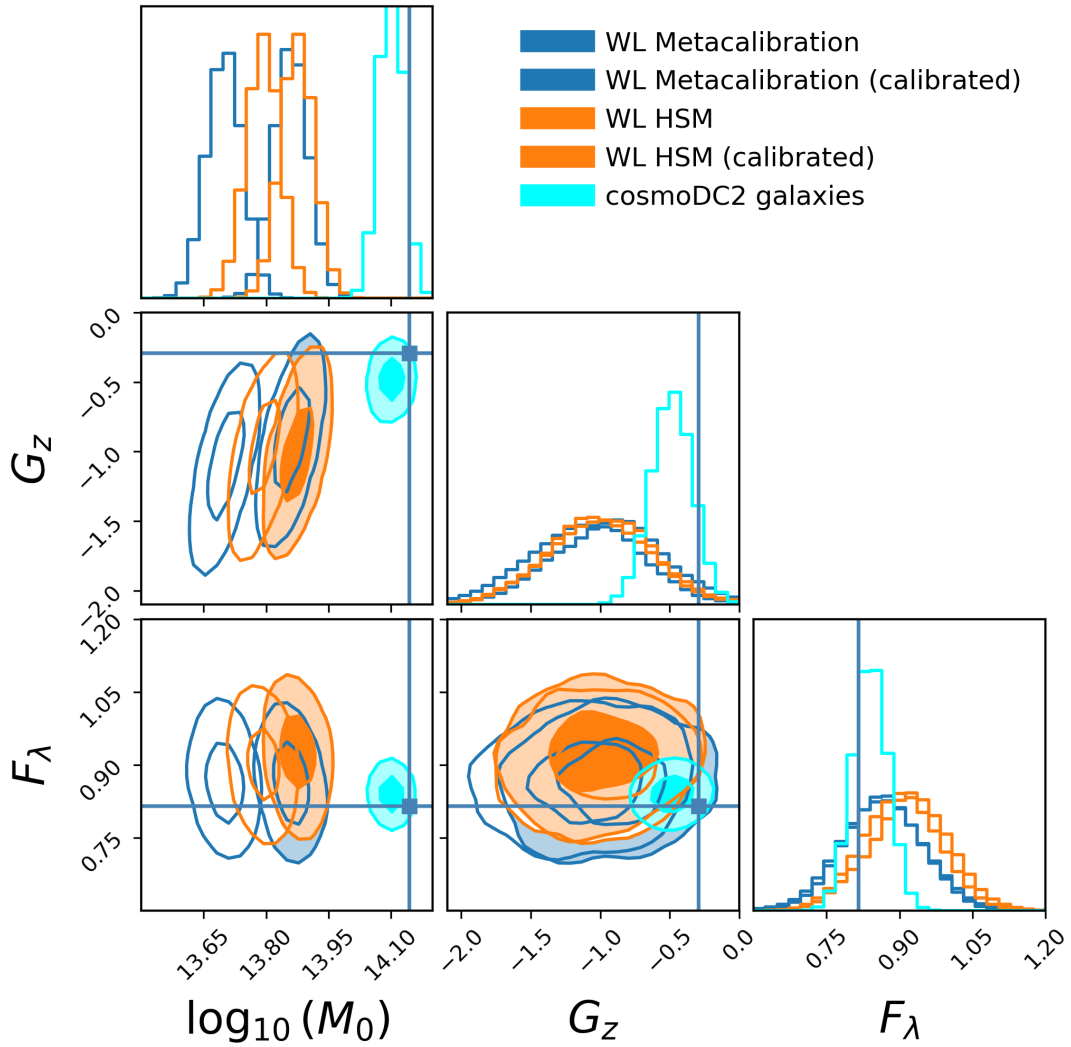

Figure 4.13: Posterior distribution of the mean scaling relation parameters for the (uncalibrated: empty contours, calibrated: filled contours) mCal and HSM shapes (respectively in blue and orange). The cosmoDC2 ideal case is represented in cyan using the 0.7.5 version of the redMaPPer catalog. The lines correspond to the fiducial constraints obtained with the 0.7.8 version.

prevented us from really evaluating the precision with which the mass-richness relation could be constrained using LSST data. However, this preliminary work has initiated the development of public software tools to interpret the cluster lensing signal with even more realistic data, i.e., detected objects with measured shapes whose estimators are affected by a series of systematic effects. The approach developed above has room for several improvements that we describe in the next section.

4.2.5 Perspectives: Improved modeling of the scaling relation

This section presents several projects that I am currently conducting to improve the constraints on the cluster scaling relation parameters. I first present how to account for the mass dependence on $\Delta\Sigma$ in the inference of scaling relation parameters. Second, I present how to skip the stacked lensing mass as an intermediate step to use directly the shear profiles to infer the mass-richness relation. Finally, I present how the DC2 stacked lensing profiles can be used to test the impact

of selection bias on the inference of scaling relation parameters.

Incorporating the mass dependence of $\Delta\Sigma$

In this section, we discuss the effect of stacking shear profiles to constrain the "effective" cluster properties. The link between the "true" mean mass of a stack and the inferred mean weak lensing mass may be not trivial. The excess surface density profile does not simply scale with respect to the underlying cluster mass. Melchior et al. (2017) has proposed that the excess surface density follows

$$\Delta\Sigma(R|M) \propto M^{\Gamma(R|M)}, \quad \text{where} \quad \Gamma(R|M) = \frac{d\ln\Delta\Sigma(R|M)}{d\ln R} \tag{4.16}$$

is the logarithmic slope of the excess surface density, and proposed that (see also McClintock et al. (2019))

$$\langle M_m\rangle^{\Gamma} = \sum_{k=1}^{N_m} \widehat{W}_k \langle M|\lambda_k, z_k\rangle^{\Gamma}, \tag{4.17}$$

where $\widehat{W}_k$ correspond to the sum of lens-source weights w_{ls} for each cluster k within the radial range considered for the mass fitting, and $\langle M|\lambda_k, z_k\rangle$ is the mean mass-richness relation in Eq. (4.10). For instance, for a NFW profile, we find a typical average $\Gamma \approx 0.75$ for $R \in [1, 5]$ Mpc, and $\Gamma \approx 0.8$ for $R \in [1, 10]$ Mpc within the typical range of masses and redshifts probed in this work (e.g., Melchior et al. (2017) found typical value of $\Gamma = 0.74$). We did preliminary studies to infer the cluster scaling relation parameters using this methodology that led to very similar results to what was obtained using the "mean mass" approach in the previous section. This work is still under development to refine the cluster scaling relation constraints, accounting for the richness and redshift scatter of clusters in the stack, as well as this non-trivial mass dependence of the excess surface density.

The lensing mass is however impacted by several other modeling systematics associated to individual cluster properties, that are very difficult to model. For instance, McClintock et al. (2019) used simulations to reconstruct the stack lensing signal around clusters with masses drawn from known $P(M|\lambda)$ relation, as well as concentrations, miscentering offsets drawn from specific distributions. They found that the bias between the true average mass and stacked weak lensing mass (see their Figure 9) has a strong richness and redshift dependence, denoting a strong impact of intrinsic scatter (in richness, concentration, misccentering, etc). In their analysis, they rescaled the fitted lensing masses accounting for this calibrated bias, before using them in the stacked mass likelihood to infer the cluster scaling relation. As a perspective of the work presented in Section 4.2.3, I aim to implement this methodology in the reconstruction of DC2 stacked lensing masses using CLMM mock clusters.

Scaling relation from a joint analysis of the stacked lensing profiles

So far we have explored the inference of the cluster scaling relation from weak lensing mass estimates. We present an alternative method to derive the cluster scaling relation directly from the cluster lensing profiles, without using the mean mass inference as an intermediate step (see e.g. Simet et al. (2017)). The main objective of this alternative is to interpret more accurately the signature of a stacked cluster lensing profile with respect to the underlying scaling relation, accounting for the intrinsic scatter in individual cluster properties (concentration, triaxiality, miscentering, etc.), as well as the scatter in individual masses in each single richness bins (that can be very large for low richness). This methodology enables to measure the scatter of the

cluster scaling relation, that the precedent approach was not "directly" sensitive to, by simply fitting the "mean" scaling relation to mean mass estimates.

As part of my work within DESC, I developed a methodology to infer the cluster scaling relation directly from cluster lensing profiles. Following what we have presented above, for a single richness-redshift bin m with selected clusters with indexes $1 \leq k \leq N_m$, we measure a stack of individual $\Delta\Sigma$ profiles at different masses. The specific mass and concentration dependency of the individual $\Delta\Sigma_k$ profiles (that is not purely linear) induces that $\langle\Delta\Sigma(R)\rangle_m \neq \Delta\Sigma(R|\langle M|\lambda_m, z_m\rangle)$. Assuming only a non-zero scatter of individual masses at fixed richness and that $\Delta\Sigma_k \propto M_k^\Gamma$ (where Γ is the logarithm slope of the excess surface density), the stacked excess surface density approximates by

$$\langle\Delta\Sigma(R)\rangle_m \approx \Delta\Sigma(R|\langle M|\lambda_m, z_m\rangle)\left(1 + \frac{1}{2}\Gamma(\Gamma-1)\sigma^2_{\ln M}\right). \tag{4.18}$$

As an example, using $\sigma^2_{\ln M} = 0.118$ (Murray et al., 2022) and $\Gamma = 0.75$, we find that the multiplying factor in Eq. (4.18) is 0.98, so the stacked profile is about 2% smaller than the "individual" profile scaled by the mean mass in the stack. With this toy model, we see how the specific scaling of the individual excess surface density profiles with respect to mass (here $\Gamma \neq 1$) can "bias" the prediction of the stacked excess surface density profile when individual masses are scattered at fixed richness ($\sigma^2_{\ln M} \neq 0$), compared to the more naive approach telling that the mean profile is scaled by the mean mass only.

More generally the individual lensing profiles are scattered in many different ways (richness, triaxiality, concentration, substructure, etc.). The estimated stacked excess surface density can be re-written in terms of a weighted mean of individual estimated lensing profiles within the stack (with weights W_k as presented in Section 2.4), the expectation of the excess surface density profile in a stack is given by the weighted mean

$$\langle\Delta\Sigma(R)\rangle_m = \sum_{k=1}^{N_m} W_k \Delta\Sigma_k(R|\Theta_k), \tag{4.19}$$

where each individual cluster with index k is described by a set of parameters $\Theta_k^{\rm ind} = (M_k, c_k, ...)$. Then, each stack m is described by a set of parameters $\Theta_m^{\rm stack} = \{\Theta_k^{\rm ind}\}_{1\leq k\leq N_m}$, and the set of cluster lensing profiles are described by a set of parameters $\Theta^{\rm full} = \{\Theta_m^{\rm stack}\}_{1\leq m\leq c}$ where c is the number of richness-redshift bins. Each prediction (in each bin) would require to fit from a few hundred to a few thousand degrees of freedom. In absolute, we should draw the joint parameter posterior of all of them from the data. In practice, as it is performed in Simet et al. (2017), the individual masses M_k within the stack m can be drawn according to the cluster scaling relation in Eq. (4.9) (as well as e.g. for the individual concentrations that can be drawn from a concentration-mass relation), thus reducing significantly the number of degrees of freedom from few thousand (in Θ) to the scaling relation parameters only.

We can estimate the joint posterior distribution of the scaling relation parameters considering the likelihood

$$\ln\mathcal{L}(\widehat{\Delta\Sigma}|\theta) = \sum_{m=1}^{c} \ln\mathcal{L}_m(\widehat{\Delta\Sigma}_m|\theta), \tag{4.20}$$

where $\mathcal{L}_m \propto \exp(-\frac{1}{2}\chi^2_m)$ with χ^2_m given by

$$\chi^2_m = (\langle\Delta\Sigma\rangle_m - \widehat{\Delta\Sigma}_m)^T \mathbf{C}_m^{-1} (\langle\Delta\Sigma\rangle_m - \widehat{\Delta\Sigma}_m). \tag{4.21}$$

In the above equation, $\mathbf{C}_m$ is the data-driven covariance matrix between radial bins of the stacked excess surface density profile in the richness-redshift bin m, thus accounting for the

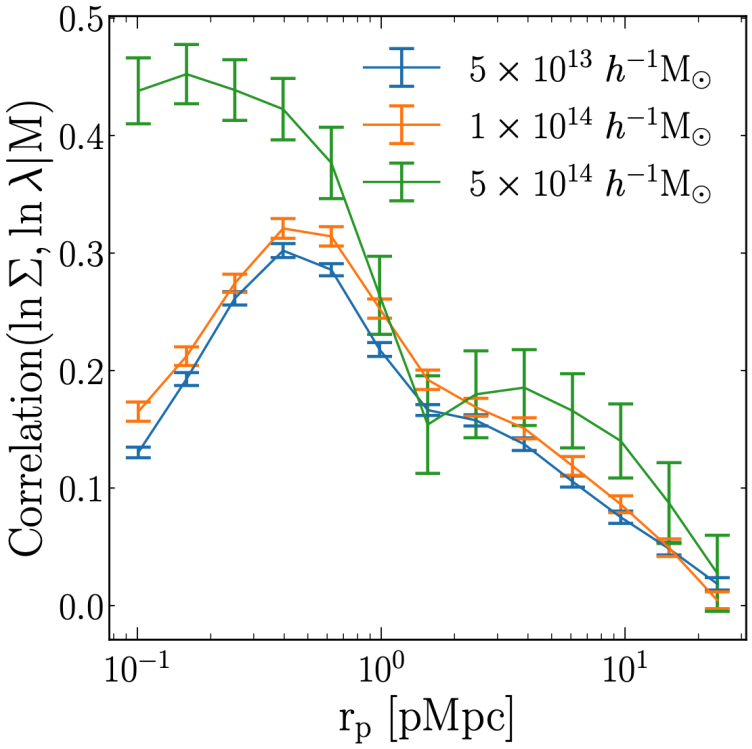

Figure 4.14: Correlation between the projected mass density (obtained from dark matter particles) and richness estimates at fixed mass for redMaPPer-detected clusters in the Buzzard simulations (DeRose et al., 2022). The colors correspond to the narrow mass bins used for the analysis. Figure from Wu et al. (2022).

intrinsic scatter of the lensing profiles with respect to cluster mass and richness but also to the other cluster properties[25]. $\langle \Delta\Sigma \rangle_m$ is the predicted excess surface density in the redshift-richness bin m in Eq. (4.19), where masses were drawn following the $P(M|\lambda, z)$ relation. Then, the modeling of the excess surface density in Eq. (4.19) takes into account both scattering in richness and concentration, as well as the lensing efficiency in over/down-weighting clusters in the stack. I am currently repeating the analysis checking for the impact of modeling choices in that framework.

Selection bias of redMaPPer cluster finder in the DC2 simulations - Impact on lensing

The advantage of working directly with the lensing profile is that we can model and correct them from the selection bias, that is introduced by the cluster finder (Wu et al., 2022); It is typically assumed that the richness-detected clusters exhibit an unbiased weak lensing signal when sharing the same mass. However, selection bias may introduce a correlation between the cluster shear and measured richness. For instance, if prolate clusters with major axes aligned along the line of sight are preferentially detected by the cluster finder, then their lensing signals would be boosted compared to spherical halos with same masses (see Chapter 6). This effect of projection would also change the cluster richness estimate, that is measured in a fixed aperture around the cluster center. The recent DES (Abbott et al., 2020) analysis from cluster abundance

[25] McClintock et al. (2019) show that the data-driven covariance matrix accounts for every intrinsic scatter of cluster properties.

and lensing suggests that uncontrolled systematics may have occurred from the optical selection of clusters by the redMaPPer algorithm (Rykoff et al., 2014), thus impacting the interpretation of the stacked cluster lensing signal, and in the end, the weak lensing mass. Fig. 4.14 shows the positive correlation between the projected mass density and richness estimates of redMaPPer detected clusters at fixed mass that was measured by Wu et al. (2022) in the Buzzard simulations (DeRose et al., 2022). From this effect, the estimated stacked lensing shear and the derived lensing mean mass would be biased high. The lensing-richness residuals can be modeled through the covariance

$$\mathrm{Cov}(\{\ln\Delta\Sigma, \ln\lambda\}|m,z) = \begin{pmatrix} \sigma^2_{\ln\lambda} & r\sigma_{\ln\lambda}\sigma_{\ln\Delta\Sigma} \\ r\sigma_{\ln\lambda}\sigma_{\ln\Delta\Sigma} & \sigma^2_{\ln\Delta\Sigma} \end{pmatrix}, \tag{4.22}$$

where $r = \mathrm{Corr}(\{\ln\Delta\Sigma, \ln\lambda\}|m,z)$ ($\neq \mathrm{Corr}(\{\ln\Sigma, \ln\lambda\}|m,z)$ represented in Fig. 4.14), and $\sigma_{\ln\lambda}$ and $\sigma_{\ln\Delta\Sigma}$ are respectively the richness and the excess surface density variance. The lensing signal $\Delta\Sigma$ and the richness λ follow the joint distribution $P(\Delta\Sigma, \lambda|m,z)$ with covariance given in Eq. (4.22). When accounting for this effect in the prediction of the stacked excess surface density profile (in the redshift bin $[z_1, z_2]$ and richness bin $[\lambda_1, \lambda_2]$), we get that (Wu et al., 2022; Evrard et al., 2014)

$$\langle\Delta\Sigma\rangle \propto \int_{z_1}^{z_2} dz \int_{m_{\rm lim}}^{+\infty} dm \frac{dn(m,z)}{dm} \frac{d^2V(z)}{dz d\Omega} \int_{\lambda_1}^{\lambda_2} d\lambda P(\lambda|m,z)\langle\Delta\Sigma|\lambda,m,z\rangle, \tag{4.23}$$

where $P(\lambda|m,z)$ encodes the relation between the cluster mass and richness, and

$$\langle\Delta\Sigma|\lambda,m,z\rangle = \langle\Delta\Sigma|m,z\rangle \exp\left(r\frac{\sigma_{\ln\Delta\Sigma}}{\sigma_{\ln\lambda}}[\ln\lambda - \langle\ln\lambda|m,z\rangle] - \frac{1}{2}\sigma^2_{\ln\Delta\Sigma}r^2\right). \tag{4.24}$$

The multiplying factor of $\langle\Delta\Sigma|m,z\rangle$ in the above equation denotes the effect of selection bias. Similar to the modeling and observational systematics we have studied in Section 4.2.4, this selection effect may bias the inference of the cluster scaling relation when inferred from stacked shear profiles if not correctly modeled. I am currently co-leading a project with a DESC colleague, Zhuowen Zhang, who aims at characterizing the selection bias (the shear-richness correlation, represented by r in Eq. (4.22)) of the redMaPPer cluster sample in the DC2 simulations. Our goal is to test how the optical selection of redMaPPer clusters impacts the measurement of stacked lensing profiles and then the scaling relation parameters. This work is a dedicated DESC project that will lead to a DESC publication.

4.3 Cosmological analysis with galaxy clusters in DC2

We have looked in detail at how modeling choices may affect the mass-richness relation, albeit in a simple approach. But this relation is not the end of the story and we also want to use the DC2 dataset all the way to cluster cosmology constraints. In this section, we perform a joint analysis of cluster count and cluster lensing to simultaneously constrain some cosmological parameters and the redMaPPer cluster mass-richness relation. We first introduce the redMaPPer selection function (i.e., purity and completeness) of the cluster catalog that is mandatory to constrain cosmology with cluster abundance. We then present the results of our joint analysis. The motivation behind this analysis was raised during the discussions within the DESC cluster WG for the design of the cosmological pipeline using galaxy clusters.

From Chapter 2, we recall that the redshift-richness cluster number density is given by

$$\frac{d^2N(\lambda,z)}{dzd\lambda} = \int_{m_{\rm min}}^{+\infty} dm \ \frac{d^2N(m,z)}{dzdm}\Phi(\lambda,m,z)P(\lambda|m,z), \tag{4.25}$$

where $\Phi(\lambda, m, z)$ is the redMaPPer selection function and where, this time, $P(\lambda|m, z)$ corresponds to the forward modeling of the mass-richness relation.

In Section 4.3.1 we present the performance of the cluster detection algorithm, redMaPPer, encoded in the selection function. In Section 4.3.2 we cover the forward modeling of the scaling relation. The method we apply to infer the cluster scaling relation and cosmological parameters is detailed in Section 4.3.3. Finally, Section 4.3.4 presents the results of the cosmological analyses with galaxy clusters in the DC2 simulations.

4.3.1 Cluster selection function

One of the key ingredients for constraining cosmology with cluster abundance is the selection function Φ associated to the cluster finder algorithm and survey strategy (see Eq. (2.14) in Chapter 2). It reflects the fact that the cluster finder algorithm may miss a non-negligible fraction of true clusters of galaxies that are in reality associated to halos, as well as detect "false" clusters that are not related to underlying collapsed dark matter structures. In this work, we follow the selection function parametrization from Aguena & Lima (2018), where the local number of detected clusters $N_{\rm obs}$ decomposes as

$$N_{\rm obs} = \underbrace{\frac{N_{\rm obs}}{N_{\rm true}}}_{\text{(inverse) purity}} \times \underbrace{\frac{N_{\rm true}}{N^{\rm tot}_{\rm true}}}_{\text{completeness}} \times \underbrace{N^{\rm tot}_{\rm true}}_{\text{Number of DM halos}}, \tag{4.26}$$

where $N_{\rm true}$ is the number of true dark matter halos that are effectively detected (i.e., each one of them is truly associated to one redMaPPer-detected cluster), and $N^{\rm tot}_{\rm true}$ is the true underlying number of dark matter halos. Following this decomposition, the selection function reduces to

$$\Phi(\lambda, m, z, z_{\rm obs}) = \frac{c(m, z)}{p(\lambda, z_{\rm obs})}, \tag{4.27}$$

where $c(m, z)$ is the completeness and $p(\lambda, z_{\rm obs})$ is the purity. The selection function usually accounts for the true halo redshift z and the observed redshift $z_{\rm obs}$ of the detected galaxy cluster, that may differ from the true halo redshift. Here, we consider $z = z_{\rm obs}$ for simplicity[26]. We see that completeness and purity answer two different questions about cluster detection: (i) does the cluster finder miss some halos, and if yes how many? (ii) does the cluster finder detect some "false" clusters, and if yes how many? By assuming a cosmological model that predicts $N^{\rm tot}_{\rm true}$, the quantity $N_{\rm true}$ is not directly observable. It can be accessed by running the cluster finder on simulations, then the geometrical match between the cluster and the halo catalog allows to disentangle between the effects of purity and completeness and to measure $N_{\rm true}$. Here, because we are already using simulations, we can directly measure these quantities.

Applying this methodology, the redMaPPer purity and completeness were measured with ClEvaR in narrow richness, mass and redshift bins. The completeness $\widehat{c}_{ij}$ in the mass and redshift bins $[m_i, m_{i+1}]$ and $[z_j, z_{j+1}]$ and the purity $\widehat{p}_{kl}$ in the richness and observed redshift bins $[\lambda_k, \lambda_{k+1}]$ and $[z_{{\rm obs},l}, z_{{\rm obs},l+1}]$ given by

$$\widehat{c}_{ij} = \frac{N^{\rm halos}_{{\rm match},ij}}{N^{\rm halos}_{{\rm tot},ij}} \quad \text{and} \quad \widehat{p}_{kl} = \frac{N^{\rm clusters}_{{\rm obs},kl}}{N^{\rm halos}_{{\rm match},kl}}, \tag{4.28}$$

where $N^{\rm halos}_{{\rm tot},ij}$ is the total number of halos in the ij mass-redshift bin, $N^{\rm clusters}_{{\rm obs},kl}$ is the number of redMaPPer clusters in the kl richness-redshift bin. The quantity $N^{\rm halos}_{\rm match,ij}$ (respectively $N^{\rm halos}_{{\rm match},kl}$) is the number of halos that have been matched to redMaPPer clusters in the ij mass-redshift

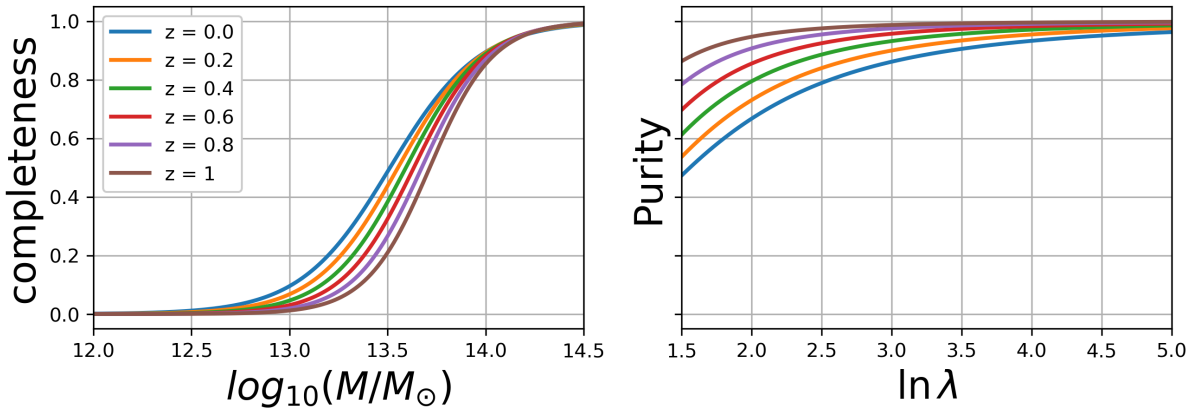

Figure 4.15: Left: Best fit redMaPPer completeness with respect to dark matter halo mass. The different colors correspond to the different redshift bins. Left: redMaPPer purity with respect to redMaPPer richness. Figures by Thibault Guillemin, Cluster WG.

bin (respectively in the kl richness-redshift bin). We consider that the purity and completeness are well described by a smoothed step function, that is given by

$$f(x) = x^\beta/(1+x^\beta), \tag{4.29}$$

where we have, for purity

$$x = \log\lambda/\log\lambda_p \quad \text{with} \quad \log\lambda_p = a_\lambda^p + b_\lambda^p(1+z), \tag{4.30}$$
$$\beta_p(z) = a_n^p + b_n^p(1+z), \tag{4.31}$$

and for the completeness,

$$x = \log(m/m_c) \quad \text{with} \quad \log_{10} m_c = a_m^c + b_m^c(1+z), \tag{4.32}$$
$$\beta_c(z) = a_n^c + b_n^c(1+z). \tag{4.33}$$

My colleague Thibault Guillemin fitted these parameters on the measured completeness and purity from the simulations[27]. The best fit parameters of the purity and completeness functions are shown in Table 4.3. The Fig. 4.15 shows the best-fit redMaPPer completeness and purity. From the figure, we see that the redMaPPer cluster catalog is complete at the 80% level for $M_{200c} > 10^{14} M_\odot$, with a small redshift dependency, and is pure at the $> 90\%$ level for $\lambda > 12$. At respectively high mass and high richness, the completeness and purity roughly equal one, meaning that each halo and each cluster have been paired respectively with a cluster and a halo. At low mass, the effect of completeness is important, since more than half of the halos are not detected below $M_{200c} = 10^{13.6}\ M_\odot$ (the completeness is 1/2).

4.3.2 Forward modeling of the cluster scaling relation

The $P(\lambda|m,z)$ quantity is referred to as *forward* modeling by Murata et al. (2019) and was also used by Saro et al. (2015). It is different from the *backward* modeling $P(m|\lambda,z)$ that we explored in the previous section.

[26]The observed density of clusters in Eq. (2.14) (see Chapter 2) considers that $z = z_{\rm obs}$ for simplicity, otherwise it should account for the conditional distribution $P(z_{\rm obs}|z)$.

[27]The figures with measured completeness and purity and best fits by T. Guillemin are available here `https://me.lsst.eu/tguillem/clusters/cluster_challenge/selection_function/`.

	$p \pm \Delta p$
a_n^c	1.1321 ± 0.6974
b_n^c	0.7751 ± 0.3885
a_m^c	13.31 ± 0.0043
b_m^c	0.2025 ± 0.0024
a_n^p	0.8612 ± 0.3966
b_n^p	0.3527 ± 0.2209
a_λ^p	2.2183 ± 0.0401
b_λ^p	-0.6592 ± -0.0223

Table 4.3: Best fit parameters and 1σ error bars for the completeness (upper block) and purity (lower block) functions.

	$p \pm \Delta p$
$\ln \lambda_0$	3.01 ± 0.01
μ_m	2.29 ± 0.02
μ_z	-0.01 ± 0.04
$\sigma_{\ln \lambda_0}$	0.556 ± 0.006
σ_m	-0.039 ± 0.006
σ_z	-0.12 ± 0.02
m_0	$10^{14.2}\ M_\odot$
z_0	0.4

Table 4.4: Best fit parameters and 1σ error bars for the mass-richness relation parameters (upper block: log-mean parameters, middle block: log-dispersion parameters, lower block: pivot values).

Compared to backward modeling, forward modeling allows us to account for the scaling relation more naturally in the abundance formalism. Although conceptually simple (using the Bayes theorem) the transformation between $P(m|\lambda, z)$ and $P(\lambda|m, z)$ is difficult to obtain in practice, since it implies knowing the richness function $P(\lambda)$ over the whole λ range, which depends on the underlying cosmological model in a complicated way (see Murata et al. (2019); Sunayama et al. (2023)) and depends on the richness range of the detected cluster catalog[28].

We consider the log-normal scaling relation with 6 free parameters, that is given by

$$P(\lambda|m, z) \propto \frac{1}{\lambda} \exp\left\{ -\frac{[\ln \lambda - \langle \ln \lambda|m, z\rangle]^2}{\sigma^2_{\ln \lambda|m,z}} \right\} \tag{4.34}$$

where the mean richness has both mass and redshift dependencies, such as

$$\langle \ln \lambda|m, z\rangle = \ln \lambda_0 + \mu_z \log\left(\frac{1+z}{1+z_0}\right) + \mu_m \log_{10}\left(\frac{m}{m_0}\right), \tag{4.35}$$

and the dispersion of the log-normal scaling relation is given by

$$\sigma_{\ln \lambda|m,z} = \sigma_{\ln \lambda_0} + \sigma_z \log\left(\frac{1+z}{1+z_0}\right) + \sigma_m \log_{10}\left(\frac{m}{m_0}\right). \tag{4.36}$$

[28] Evrard et al. (2014) have studied this transformation accounting for some strong assumption on the underlying cosmological models (especially the mass variation of the halo mass function), to switch from a quantity to another.

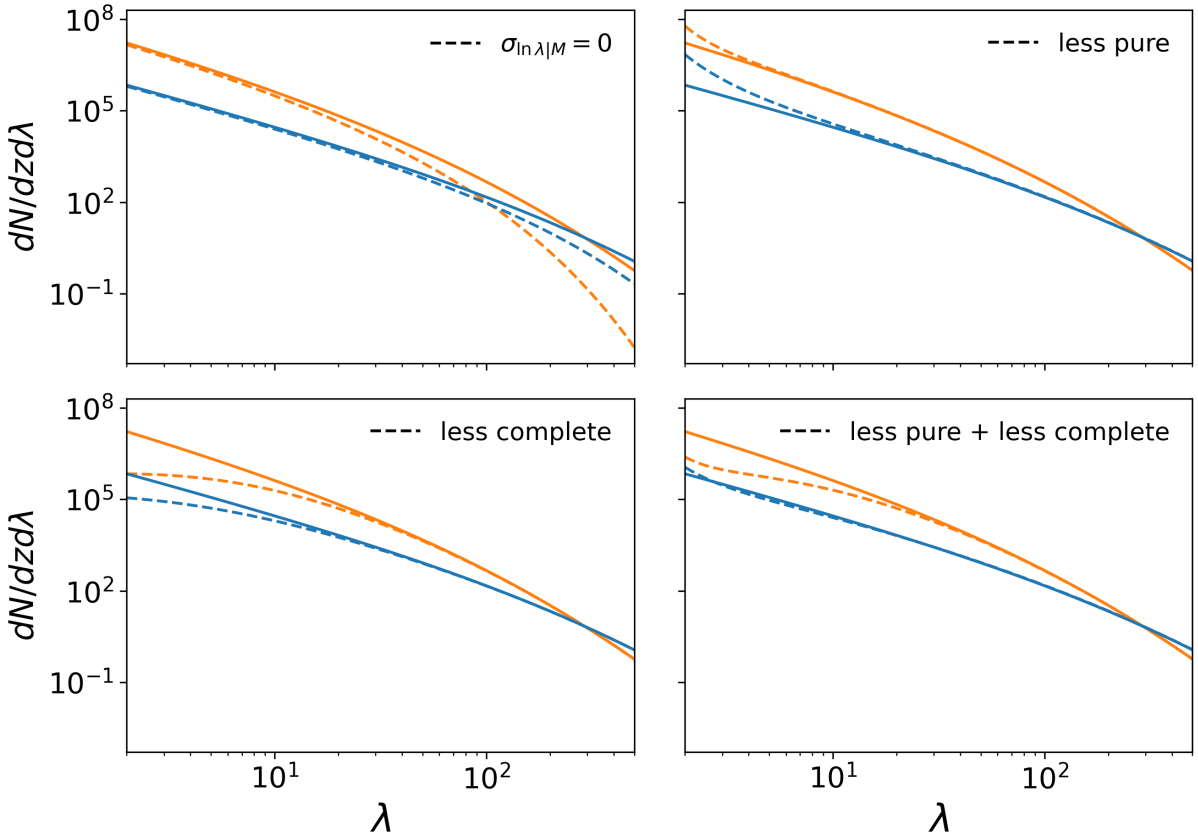

Figure 4.16: Observed number density of clusters with respect to richness. The different colors correspond to the different redshifts used (orange: $z = 0.6$, blue: $z = 0.1$). In each subplot, the full line corresponds to the predicted number density for a pure and complete catalog, with a scattered scaling relation. Upper left: the dashed lines correspond to the number density for a richness-mass relation with no dispersion (i.e., $\sigma_{\ln\lambda|m,z} = 0$). Upper right: $p(\lambda, z) \neq 1$, $c(m, z) = 1$, Lower left: $p(\lambda, z) = 1$, $c(m, z) \neq 1$. Lower right: $p(\lambda, z) \neq 1$, $c(m, z) \neq 1$.

In our case, the methodology to derive the fiducial relation $P(\lambda|m, z)$ is similar to that of Section 4.2.2 for the $P(m|\lambda, z)$ relation, by considering the cosmoDC2 masses matched to redMaPPer richnesses (this time fitting $P(\lambda|m, z)$ instead of $P(m|\lambda, z)$). The fiducial values of the 6 parameters (and pivot values) of the $P(\lambda|m, z)$ relation are shown in Table 4.4.

Before performing a cosmological analysis, we can test how the different quantities (selection function and mass-richness relation) impact the observed number density of clusters. Fig. 4.16 shows the fiducial redMaPPer number density as a function of the redMaPPer richness, accounting for the best fit parameters for the completeness, the purity as well as the scaling relation $P(\lambda|m, z)$ (the fiducial constraints). The figure shows how the different ingredients impact the observed number density of objects in a real data set. First (upper left panel), the possibly large scatter of the scaling relation affects the cluster abundance especially at large richness, compared to the zero-scatter case. Second (upper right panel) a less pure cluster catalog has a non-zero fraction of "false" detection, thus boosting the detected abundance (in this case, especially at low richness). Reversely (lower left panel), an incomplete catalog means that the cluster finder misses some dark matter halos, so the observed abundance is lower. Both effects combine in the lower right panel, where the observed abundance is boosted/lowered depending on the richness range.

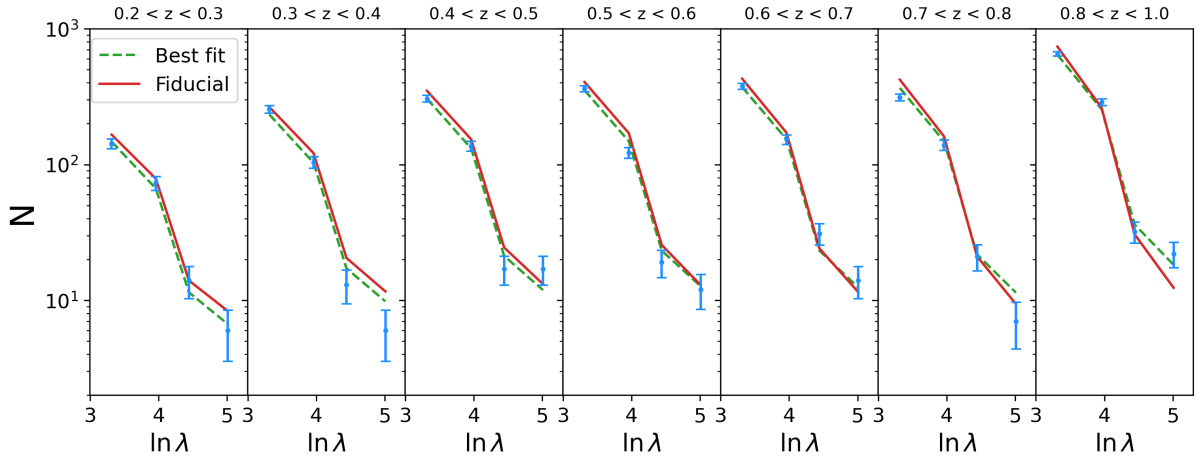

Figure 4.17: Measured redMaPPer cluster abundances (blue dots and error bars) in richness bins. The seven panels correspond to the seven redshift bins.

4.3.3 Methods for cluster abundance and weak lensing analyses

Having understood the qualitative impact of these ingredients of the analysis, let us turn to the actual cosmological analysis.

Cluster abundance model and likelihood

We show in Fig. 4.17 the measured abundance of redMaPPer clusters with respect to the different richness bins and for different redshift bins (blue points). The error bars correspond to the Poisson standard deviation, namely $\sigma_i^2 = \widehat{N}_i$. In our analysis, we will consider that the predicted cluster count in a i-th richness and j-th redshift bin is given by

$$N_{ij} = \int_{z_i}^{z_{i+1}} dz \int_{\lambda_j}^{\lambda_{j+1}} d\lambda \int_{m_{\rm min}}^{+\infty} dm \, \frac{d^2N(m,z)}{dzdm} \frac{c(m,z)}{p(\lambda,z)} P(\lambda|m,z), \tag{4.37}$$

where we used the Bocquet et al. (2015) halo mass function (dark matter only). We consider a Poisson likelihood, i.e., the number of galaxy clusters in each richness and redshift bin is the result of a Poisson realization of the underlying cosmological abundance[29].

Cluster stacked weak lensing mass model and likelihood

To constrain cosmological parameters, we can use the weak lensing masses estimates obtained in Section 4.2.3. In that case, the model is given by

$$M_{ij} = \frac{1}{N_{ij}} \int_{z_i}^{z_{i+1}} dz \int_{\lambda_j}^{\lambda_{j+1}} d\lambda \int_{m_{\rm min}}^{+\infty} dm \; m \frac{d^2N(m,z)}{dzdm} c(m,z) P(\lambda|m,z). \tag{4.38}$$

We will consider the stacked masses obtained by considering a NFW profile by letting free the cluster concentration ([Ia] in Table 4.2). For each redshift-richness bin, we use a Gaussian like-

[29]This choice may not be the most appropriate, since the contribution for the Super Sample Covariance that correlate bins and add variance may not be negligible for very high populated bins (see the Section 2.3.1). This will impact the inference with cluster abundance by underestimating errors (we explore this effect in more detail in Chapter 5. However, for this first analysis using DC2 data, we opted for the more straightforward Poisson likelihood implementation. Inclusion of the SSC is more demanding (see Chapter 5) and needs to be addressed; this is work I have started in the public DESC CLCosmoSim repository.

lihood with a diagonal covariance matrix, accounting for the variance of the recovered lensing mass[30].

Cluster weak lensing profile model and likelihood

We will also consider the inference of cosmological and scaling relation parameters by analyzing directly the stacked redMaPPer lensing profiles. We adopt a methodology inspired from to that presented in Park et al. (2023); Murata et al. (2019) and more recently in Sunayama et al. (2023). Compared to the two-step approach (first fitting the lensing masses and then constraining the cosmology), the "forward modeling" of cluster lensing observables enables more flexibility to incorporate several systematic effects associated to the lensing measurement, such as the mis-centering, the projection of correlated and uncorrelated structures around the clusters, cluster member contamination and the shape of non-spherical clusters. Lensing profiles may also be directly cosmology-independent[31], compared to the lensing masses that are determined assuming a fiducial cosmology[32]. However, the forward modeling of the cluster lensing signals still represents a significant computing challenge, since it requires to marginalize over many theoretical and observational effects (Aguena et al., 2023).

We will compare each measured lensing profile to the prediction

$$\Delta\Sigma_{ij} = \frac{1}{N_{ij}} \int_{z_i}^{z_{i+1}} dz \int_{\lambda_j}^{\lambda_{j+1}} d\lambda \int_{m_{\rm min}}^{+\infty} dm \, \frac{d^2N(m,z)}{dzdm} c(m,z) P(\lambda|m,z) \Delta\Sigma(R|m,z). \tag{4.39}$$

For the $\Delta\Sigma(R|m,z)$ terms, we use a NFW profile with a $c(M)$ relation from Diemer & Kravtsov (2014). In the above equation, we do not account for purity in the integrand (i.e., $p = 1$), since we account for the impact of false detection that will dilute the lensing signal (false detection may not be associated to any underlying collapsed dark matter structure). The "false" detections may also be related to the pairing of more than one dark matter halo to one galaxy cluster due to projection effects, then still revealing projected matter, and that will contribute to a coherent lensing signal. We do not account for this effect in this analysis, which is meant to be preliminary. We also do not account for selection bias on cluster lensing that we discussed in Section 4.2.5, but we plan to include it in the lensing formalism in the future. For each richness-redshift bin, we use a Gaussian likelihood with data-driven covariance[33].

4.3.4 Results

Now, everything is in place to perform cluster cosmology analyses with DC2 data. The likelihood of cluster count and cluster lensing should account for the correlation between the two datasets since they both rely on the same matter density field and cluster properties. For simplicity, we consider in this work that the count and lensing data are un-correlated, such that the full likelihood "count+lensing" is the product of the two likelihoods, i.e.,

$$\mathcal{L}_{\rm tot} = \mathcal{L}_{\rm count} \times \mathcal{L}_{\rm WL}. \tag{4.40}$$

Each likelihood is also considered independently to provide separate "count" and "lensing" constraints. Here, the lensing constraints can come from considering either the weak lensing profiles or the weak lensing mass estimates.

[30] The error on the lensing mass obtained in Section 4.2.3.

[31] It is not the case in this work, since we compute the excess surface density in (physical) radial bins.

[32] For instance, Abbott et al. (2020) incorporated the cosmological dependence of the lensing masses in a second time, using mocks at different cosmologies.

[33] The one used in Section 4.2.3 when we inferred stacked masses from lensing profiles.

We first discuss the results obtained on the cluster scaling relation by fixing the cosmological parameters to their fiducial values. Second, we constrain the cosmological parameters jointly with the scaling relation parameters.

We consider flat priors for the cosmological parameters Ω_m and σ_8 and for the six parameters of the scaling relation presented in Section 4.3.1, respectively within the ranges Ω_m : $[0.1, 0.5]$, σ_8: [0.5, 1], λ_0: $[0, +\infty]$, μ_z: [-2, 2], μ_m: $[0, +\infty]$, $\sigma_{\ln\lambda 0}$: $[0, +\infty]$, σ_z: [-2,2] and σ_m: [-2,2] (priors on the scaling relation parameters are inspired from Murata et al. (2019)). We also consider the prior $\sigma_{\ln\lambda|m,z} > 0$ for all mass and redshift ranges we consider in this work.

Fixed cosmology, constrain $P(\lambda|M)$ only

Here, we fit the scaling relation parameters at fixed cosmology. We show in Fig. 4.18 the joint posterior distribution of the scaling relation parameters by considering the cluster abundance (red), the stacked lensing masses (blue, unfilled contours) and the stacked lensing profiles (green, unfilled contours). We find that the lensing-based results (cluster lensing masses or profiles) show different correlations between the recovered parameters, and are compatible at the 1σ level. We obtain the joint constraints by combining the cluster abundance and lensing masses in blue (filled contours), and by combining the cluster abundance and lensing profiles in green (filled contours). The joint constraints are significantly more precise, and we find that the two joint analyses give similar constraints. It is interesting as the 2-step approach, first computing the weak lensing masses and then using them for the cosmology, is less computationally demanding than using the weak lensing profiles directly. The "fiducial" constraints are obtained by fitting the scaling relation on the redMaPPer-cosmoDC2 matched catalog (the best-fit parameters are given in Table 4.4 and are represented in black dashed lines). We see that the inferred parameters from the abundance and lensing analyses alone are compatible at 1σ with the fiducial constraints and that the joint analysis shows $1-2\sigma$ compatibility.

Constraining all parameters at once

The Fig. 4.19 shows the joint posterior distribution of the scaling relation parameters, as well as of the cosmological parameters Ω_m and σ_8. The best fit cluster counts are represented in dashed green lines in Fig. 4.17 (when fitting the cosmological parameters and the scaling relation together). The fiducial count predictions (derived using the fiducial scaling relation in Table 4.1 and cosmoDC2 cosmology) are represented in red. The best fit lensing profiles are represented in dashed lines in Fig. 4.8.

For the scaling relation parameters, we see that we now recover at the $1-2\sigma$ level the fiducial constraints. The cosmological parameters display a non-zero correlation with some scaling relation parameters (e.g., Ω_m and μ_m are positively correlated, and σ_8 and μ_m are negatively correlated). The abundance-alone cosmological constraints are not very strong (the posterior are prior dominated), but the abundance provides better constraints when combined with cluster lensing. Combining cluster lensing (masses or profiles) and cluster count allows to recover the input cosmology at $1-2\sigma$. We also see that the two separate lensing constraints are similar.

4.3.5 Conclusions and perspectives

With this work, we conducted the first full cluster abundance and weak lensing cosmological analyses using DESC simulations, that are satisfactory in recovering the "input" parameters despite the limitation of the DC2 cluster lensing. Of course, there is plenty of room for improvement and further checks. We have a series of perspectives for this analysis, that are currently

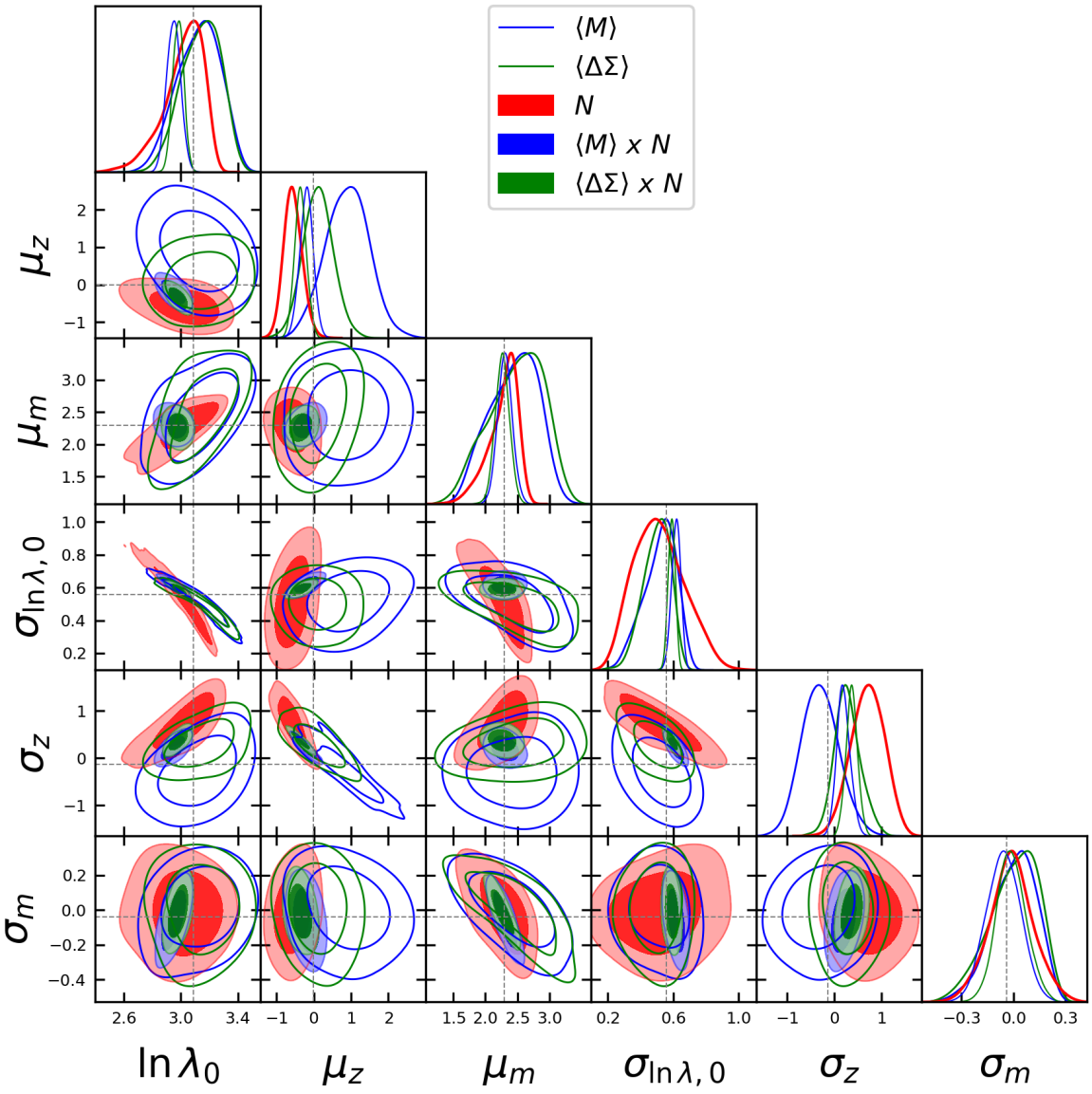

Figure 4.18: Joint posterior distribution of scaling relation parameters from cluster abundance (red), stacked lensing masses (blue, unfilled contours), stacked cluster lensing profiles (green, unfilled contours). The joint analyses are shown in blue (filled contours) for cluster abundance combined with stacked masses, and green (filled contours) for cluster abundance with stacked lensing profiles. The black lines show the fiducial constraints.

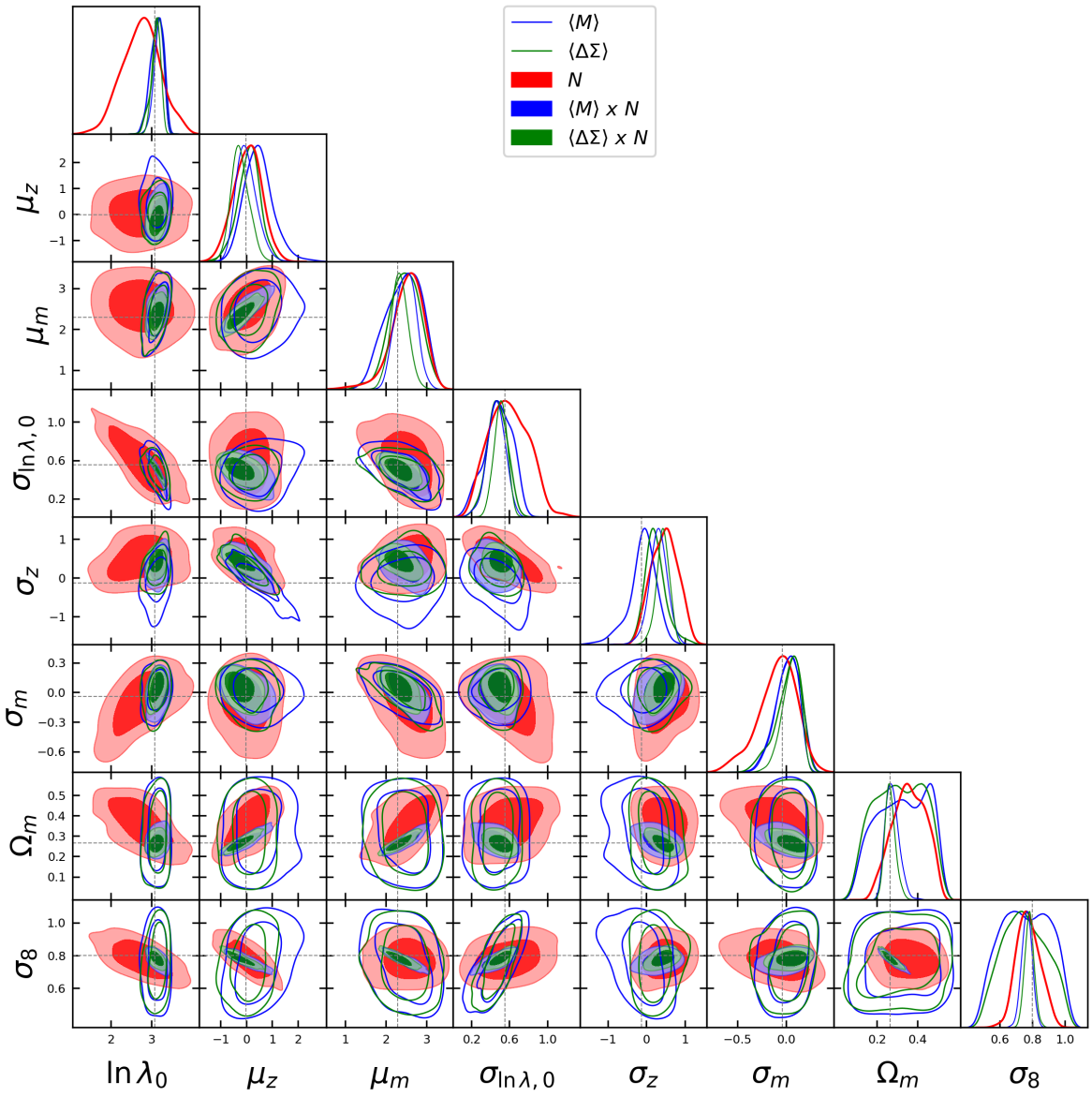

Figure 4.19: Same label as for the Fig. 4.18, by considering the cosmological parameters Ω_m and σ_8 (two last lines of the corner plot).

discussed within the DESC.

The cosmoDC2 ray-tracing is affected in the innermost regions of cluster fields, so preventing us from extracting the cosmological information from the cluster lensing profiles below $R \sim 1$ Mpc. However, the lensing profile in this region is highly sensitive to the intrinsic scatter in halo properties and benefits from a high signal-to-noise ratio. The SkySim5000 catalog is an extension of cosmoDC2 over 5000 deg^2 and has received an improved resolution of the ray-tracing, impacting the reported shears listed in the catalog. As a result, at a given redshift, the unphysical attenuation region using SkySim5000 happens at smaller scales than for cosmoDC2. On top of the high statistics thanks to the larger simulated sky, using SkySim5000 would allow us to use high SNR regions of the shear profile, which in turn would change our conclusions on the mass reconstruction (by changing the sensitivity to miscentering, scatter in $c(M)$ relation, etc.).

We want to test the robustness of the cosmological constraints to the scaling relation modeling (for instance considering log-squared dependencies, not only log-log relation as explored in Murata et al. (2019)). Since we considered the Bocquet et al. (2015) halo mass function, we want to test the robustness of our results by considering other halo mass functions, and the study cases we explored in Section 4.2.4 (different halo modeling, etc.). Here, we also mention that we did not account for the cosmological dependence of the fitted lensing masses and of the estimated stacked cluster profiles. The measured lensing masses depend on cosmology since we have to fix the distance relative to the observer, and also due to the critical surface mass density coefficients in the lens-source pairs. The estimated stacked cluster profiles also depend on cosmology in that sense[34]. Neglecting this effect may bias the recovered cosmological parameters (Abbott et al., 2020), and need to be corrected for.

4.4 Software tools

This work on DESC DC2 data and these several DESC projects have presented a major part of my PhD work. I review below the different codes that I have developed and/or used extensively to achieve the analyses I described above.

- To access the different DC2 source and lens catalogs (cosmoDC2, DC2 object, photometric add-on catalogs, redMaPPer galaxy cluster catalog), I used the DESC python package GCRCatalogs[35] (see e.g. Mao et al. (2018)) that provides a unified access interface to the DC2 catalogs that are stored in different formats. The DC2 data are stored both at CC-IN2P3 (that we presented in Chapter 3) and at the NERSC computing center. For the different aspects of the work in this chapter, I had to optimize the data extraction using the different options available in GCRCatalogs, namely by applying source extraction on reduced portions of the DC2 sky, by applying healpix/tract selection and on iterative source redshift slices. It was supplemented by parallel jobs at CC-IN2P3 to speed up the data extraction and formatting. Data extraction was a very heavy part of the whole process and it took several trials before achieving a procedure efficient enough for repeated use.

- The methodology, the software and data vectors for the inference of the scaling relation presented in Section 4.2.3 are publicly available in the DESC CLMassDC2[36] Github repository that I created to develop the analysis for the weak lensing mass-richness relation

[34]To remove this cosmological dependence of the data vector, we can measure the average tangential shear in angular bins instead of the excess surface density in radial bins.

[35]The Python package is available here `github.com/LSSTDESC/gcr-catalogs`.

[36]The repository is publicly available here `https://github.com/LSSTDESC/CLMassDC2/tree/DESC_note`.

inference. I also created a Jupyter Notebook tutorial[37] labeled as a DESC refereed note to show how to use the cosmoDC2 catalog and the photometric redshift add-on catalogs to estimate the lensing profile around clusters (i.e., to produce the data vector needed for WL mass estimation).

- I also contributed to the development of the DESC Cluster Lensing Mass Modeling code which is publicly available[38] and I am a co-author of Aguena et al. (2021b), the paper that presents the v1.0 release of CLMM. Since the first version, I implemented the functionalities to perform the stacked weak lensing shear analysis[39] and I co-authored several notebook tutorials.

- The Cluster WG is now focusing on developing the cluster cosmology pipeline and several of the codes I have developed[40] are serving as a solid basis for this work. Fig. 4.20 shows the overall pipeline architecture currently being developed. The idea is to have a suite of fully tested, documented and user-friendly public tools to perform the DESC cluster analysis. Without entering the details, the red boxes correspond to the "official" DESC software tools and several modules I developed during this work can be repurposed in this framework.

4.5 Conclusions

The abundance and lensing signal of galaxy clusters contains a wealth of information on cosmology. As part of DESC, I presented in this chapter the several projects I conducted to prepare the cosmological analyses with the galaxy clusters that will be detected on LSST data. I contributed to the development of DESC tools for analyzing the cluster lensing profiles, as well as to the development of the DESC cluster cosmology pipeline. I also contribute to characterizing the cluster lensing and cluster abundance science in the DESC DC2 simulations, from the cluster catalogs and lensed galaxy catalogs up to the cosmological parameters. In the last section of this chapter, we used the Poisson likelihood to describe cluster abundance. The choice of cluster abundance likelihood is crucial for inferring accurate and robust cosmological parameters. We address in the next chapter several aspects of cluster count likelihoods.

[37] The notebook is publicly available here `https://github.com/LSSTDESC/DC2_CL_PZ/tree/compute_deltasigma`.

[38] `https://github.com/LSSTDESC/CLMM`.

[39] The estimation of the stacked lensing profiles and associated covariance matrices, the calculation of lens-source weights accounting for photometric redshift information and shape measurement errors.

[40] Within the DESC CLCosmo_Sim repository, available here: `https://github.com/LSSTDESC/CLCosmo_Sim/tree/issue/15/binned_cosmological_analysis`.

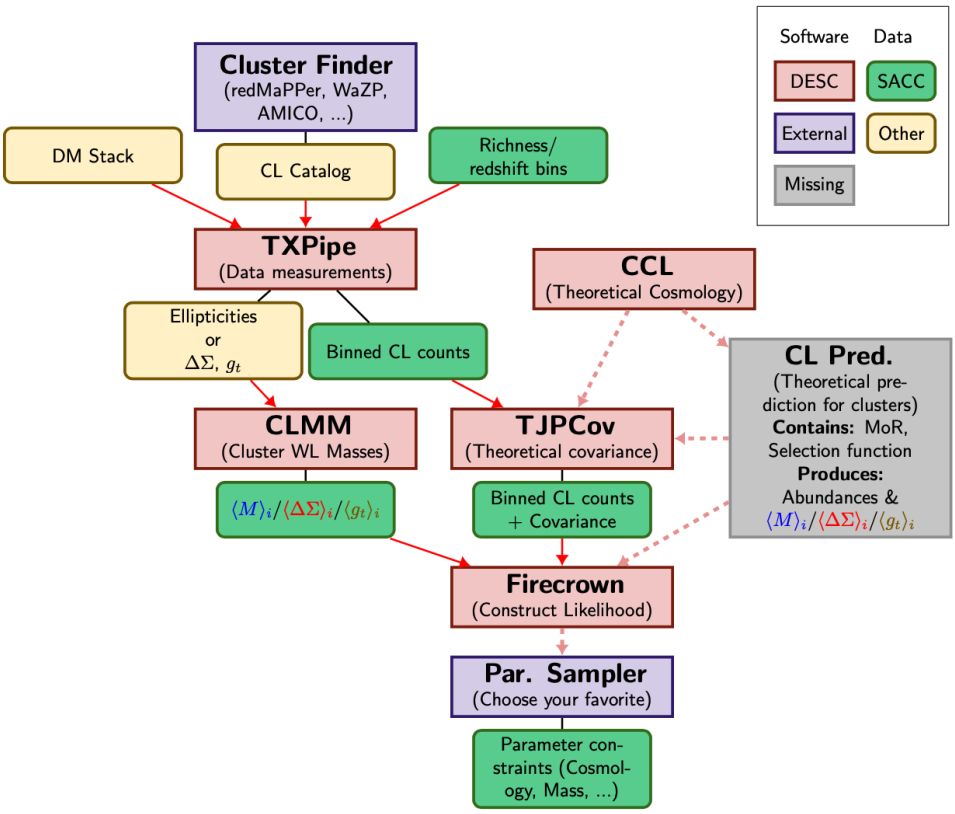

Figure 4.20: Pipeline architecture of cluster count and cluster lensing analysis in DESC (still under construction). Credits: Michel Aguena, Cluster WG.

Chapter 5

Likelihoods for cluster abundance cosmology

Contents

The inference of cosmological parameters from the observed abundance of clusters of galaxies is affected by several systematic effects. The uncertainty related to cosmological models (calibration of the halo mass function, mass-observable relation, etc.) is part of the main limiting factors for cluster-based cosmological analyses in addition to observational effects. The uncertainty in the count modeling will be even more problematic in the near future, because of the high statistical power of future large surveys. The likelihood function $\mathcal{L}$ is used at the very last step of any cosmological analysis to draw the posterior distribution $\mathcal{P}$ of the cosmological parameters $\vec{\theta}$ given by

$$\mathcal{P}(\vec{\theta}|\vec{N}) \propto \mathcal{L}(\vec{N}|\vec{\theta})\pi(\vec{\theta}). \tag{5.1}$$

The likelihood describes the statistical properties of the observed counts in bins and we used a Poisson likelihood in Chapter 4 to extract cosmological information from the abundance of DC2 clusters. In this chapter, we address in detail many aspects of cluster abundance likelihoods. Section 5.1 is based on the methodology and results we obtained in the recent paper Payerne et al. (2023) that I published during my thesis, but with added details about the methodology and the motivations. It aims to answer a very general question: is our statistical description of the abundance of clusters sufficiently accurate for future surveys? Within the cosmology community at large, this has been assessed in the context of misestimation of the covariance matrices

for cosmological analyses (Dodelson & Schneider, 2013; Taylor & Joachimi, 2014; Sellentin & Heavens, 2017).

In addition, we seek to answer multiple related questions: (i) which likelihoods can sufficiently describe the abundance of galaxy clusters and provide robust constraints, (ii) among those, does a given likelihood provide stronger cosmological constraints than the others? (iii) Is there an optimal binning regime that allows us to extract most of the cosmological information?

In Section 5.2, we present an alternative method called the unbinned cluster abundance analysis that allows us to account for individual properties of clusters, and to use the full mass and redshift distribution of the sample. Contrary to most binned cluster abundance likelihoods, the standard unbinned cluster abundance likelihood does not account for the effect of super-sample covariance. In this section, we first show analytically how to account for this effect in the unbinned formalism, and we analyze the impact on the parameter posteriors compared to the standard unbinned approach. We conclude the chapter in Section 5.3.

5.1 The accuracy of binned cluster count likelihoods

Three cluster likelihoods have been used to perform cluster count cosmology in the literature: the Poisson likelihood (e.g., Mantz et al. (2015); Ade et al. (2016); Bocquet et al. (2019)), the Gaussian likelihood (e.g., Costanzi et al. (2019); Abbott et al. (2020)) and, more recently a combination of the two (Lesci et al., 2022). However, the latent likelihood is not a Poisson process, Gaussian, or a combination of the two. By the latent likelihood here we refer to the unknown underlying likelihood which would perfectly describe the distribution of galaxy clusters in the Universe given a cosmological model. The halo model of structures in the Universe (in which each of these likelihoods is to some extent built upon) is a model, which will break down. The distribution of galaxy clusters in a universe predicted by a cosmological model most likely differs from these simplified descriptions. Therefore it is an assumption that the cluster abundance distribution is well described by these models. If not this may impact the constraint of cosmological parameters, possibly leading to misestimation of parameters errors and even bias the results[1].

In the next section, we introduce the cluster likelihoods that we have considered in this work. In Section 5.1.2 we present a statistical framework to test the accuracy and robustness of the cluster count likelihoods using a set of cosmological simulations. Then, we present the PINOCCHIO simulated dark matter halo catalogs, that serve as a reference for this work, in Section 5.1.3. To infer cosmological parameters and their errors, we use a theoretical prediction of the cluster count covariance matrix, that we present in Section 5.1.4. We present the methodology and the different cosmological inference setups we explore in this work in Section 5.1.5. Finally, we present our results in Section 5.1.6 and we conclude in Section 5.1.7. The different notations that will be used are summarized in Table 5.1.

5.1.1 Likelihoods for cluster abundance cosmology

We describe below the three different likelihoods that are evaluated in this work; the Gauss-Poisson Compound likelihood, the Poisson likelihood and the Gaussian likelihood.

[1]For instance, the insufficiency of Gaussian likelihoods in the case of the CFHTLenS cosmic shear survey was considered by Sellentin & Heavens (2018), where they note that since one-point statistic of the shear and convergence fields are skewed, the Gaussian approximation is not necessarily accurate.

Notation	Definition
$(i, j, k, ...)$	Latin symbols referring to data vector components
c_m	Number of mass bins
c_z	Number of redshift bins
$c = c_m \times c_z$	Total number of mass-redshift bins
$\vec{N} \equiv \{N_{ij}\}$	Abundance prediction vector ($1 \leq i \leq c_m$ and $1 \leq z \leq c_z$)
$\{\widehat{N}_{ij}\}$	Observed abundance vector
$\mathcal{L}_X(\{\widehat{N}_{ij}\})$	Latent likelihood
$\mathcal{L}_Y(\{\widehat{N}_{ij}\})$	Analysis likelihood
$[\Sigma_X]_{ij}$	Latent data covariance matrix element
$[\Sigma_Y]_{ij}$	Analysis data covariance matrix element
$\vec{N}_{,\alpha} \equiv \{\partial N_{ij}/\partial\theta_\alpha\}$	Derivative of the abundance prediction vector wrt θ_α
$(\alpha, \beta, \gamma, ...)$	Greek symbols referring to cosmological parameters
c_p	Total number of cosmological parameters
$\vec{\theta} = \{\theta_\alpha\}$	Cosmological parameter vector element ($1 \leq \alpha \leq c_p$)
$\mathcal{C}^{\rm ind}_{\alpha\beta}$	Individual parameter covariance matrix element
$\mathcal{C}^{\rm ens}_{\alpha\beta}$	Ensemble parameter covariance matrix element
$F_{\alpha\beta}$	Fisher matrix element

Table 5.1: Table of notations and associated definitions used in this work, for the data side (i.e., data vector, data covariances, likelihoods) and the cosmological parameters (i.e. parameter space, parameter covariances).

5.1.1.1 The Gauss-Poisson Compound likelihood (GPC)

Galaxy clusters trace the matter density field in the universe. The number density of galaxy clusters follows spatial fluctuations of the underlying matter density field. This relation can be parameterized by taking clusters to be biased tracers of the underlying density field (see Cooray & Sheth (2002) for a review of the halo model). The overdensity of galaxy clusters, δ_h, in a region centered at $\vec{x}$, is

$$\delta_h(\vec{x}) = b(m, z)\delta(\vec{x}), \tag{5.2}$$

where $b(m, z)$ is the linear bias parameter which is a function of both cluster mass and redshift and $\delta(\vec{x}, z)$ is the underlying matter density field. Using Eq. (5.2), the local halo number density, n, can be expressed as,

$$n(m, z, \vec{x}) = \bar{n}(m, z)(1 + b(m, z)\delta(\vec{x}, z)), \tag{5.3}$$

where $\bar{n}$ is the predicted mean number density of clusters at a given mass and redshift for a given cosmological model. Considering survey sizes of a few thousand square degrees, the matter density field $\delta(\vec{x}, z)$ is well described by a Gaussian random field; the same would not be true on smaller scales due to non-linear structure formation.

The abundance of galaxy clusters is the count of clusters in a specific spatial region, such as the observed number of galaxy clusters will be a Poisson realization of the *local* halo abundance, which in turn is a Gaussian realization depending on the underlying density field within the survey region. Hu & Kravtsov (2003) proposed a Gauss-Poisson Compound likelihood that combines these two statistical effects; the average number of clusters N_{ij} within a given i-th mass bin and j-th redshift bin is calculated from the halo mass function $\frac{dn(m,z)}{dm}$ predicting the average number density of halos per mass range and the comoving partial volume $\frac{d^2V(z)}{dzd\Omega}$ as

follows,

$$N_{ij} = \Omega_S \int_{z_j}^{z_{j+1}} dz \int_{m_i}^{m_{i+1}} dm \frac{dn(m,z)}{dm} \frac{d^2V(z)}{dz d\Omega}, \tag{5.4}$$

where Ω_S is the survey solid angle. In this work, we use letters of the *latin* alphabet $(i, j, k...)$ to refer to the data vector indices. Given that the statistics of the underlying matter density field $\delta(\vec{x}, z)$ is Gaussian, $n(m, z, \vec{x})$ is a random Gaussian realization of $\bar{n}(m, z)$. The observed abundance of clusters $\widehat{N}_{ij}$ in the i-th mass bin and j-th redshift bin is then a Poisson realization of the Gaussian random variable

$$N_{ij}^{\rm loc} = N_{ij} + \delta N_{ij}, \tag{5.5}$$

where $N_{ij}^{\rm loc}$ is the *local* predicted abundance with mean N_{ij} given in Eq. (5.4), where the abundance perturbation δN_{ij} is expressed in terms of the halo bias b and the matter overdensity field δ such as

$$\delta N_{ij} = \int_{\Omega_S} d^2\vec{\theta} \int_{z_j}^{z_{j+1}} dz \int_{m_i}^{m_{i+1}} dm \frac{dn(m,z)}{dm} \frac{d^2V(z)}{dz d\Omega} \times b(m,z)\delta(f_K(w)\vec{\theta}, w). \tag{5.6}$$

In this expression, $\delta(f_K(w)\vec{\theta}, w)$ is the three-dimensional matter overdensity at the comoving position $(x, y, z) = (f_K(w)\vec{\theta}, w)$, where $\vec{\theta}$ denotes the two-dimensional position on the sky, and $f_K(w)$ and w are respectively the transverse and line-of-sight comoving distances. The line-of-sight comoving distance depends on the input redshift $w \equiv w(z)$.

For a set of observed number counts $\{\widehat{N}_{ij}\}$, the Gauss-Poisson Compound likelihood is given by

$$\mathcal{L}(\{\widehat{N}_{ij}\}) = \int d\vec{N}^{\rm loc}\, \mathcal{N}(\vec{N}^{\rm loc}|\vec{N}, \Sigma_{\rm SSC}) \times \prod_{i,j=1}^{c} \mathcal{P}(\widehat{N}_{ij}|\widetilde{N}_{ij}), \tag{5.7}$$

where $c = N_z \times N_m$ is the total number of mass-redshift bins (Aitchison & Ho (1989) also introduced this formalism to analyze multivariate counts arising in many disciplines, not especially for clusters). In the above equation, $\mathcal{N}(\cdot|\vec{N}, \Sigma_{\rm SSC})$ is the multivariate Gaussian distribution with mean $\vec{N} = \{N_{ij}\}$ and covariance $\Sigma_{\rm SSC}$. The distribution $\mathcal{P}$ is the single variate Poisson distribution given by

$$\mathcal{P}(\widehat{N}_{ij}|N_{ij}^{\rm loc}) = \frac{(N_{ij}^{\rm loc})^{\widehat{N}_{ij}}}{\widehat{N}_{ij}!} e^{-N_{ij}^{\rm loc}}. \tag{5.8}$$

The quantity $\Sigma_{\rm SSC}$ is the super-sample covariance (SSC[2]), i.e., the covariance of the δN_{ij} such as $\Sigma_{\rm SSC}[ij][kl] = \langle \delta N_{ij} \delta N_{kl} \rangle$.

For a given binning scheme, if we neglect correlations between two different bins of redshift (which in general are small, see Fig. 5.2), the GPC likelihood is approximated as the product of N_z single variate integrals (instead of computing $N_z \times N_m$ multivariate integrals) and can be written as (Takada & Spergel, 2014)

$$\mathcal{L}(\{\widehat{N}_{ij}\}) = \prod_{j=1}^{N_z} \mathcal{L}_j(\{\widehat{N}_{ij}\}_{1 \leq i \leq N_m}), \tag{5.9}$$

[2]This is also termed as Sample Variance (SV) in the literature, e.g., Hu & Kravtsov (2003); Fumagalli et al. (2021). We use the SV terminology in Payerne et al. (2023) but prefer SSC in this manuscript.

where each term $\mathcal{L}_j$ in Eq. (5.9) is given by the integral form

$$\mathcal{L}_j(\{\widehat{N}_{ij}\}) = \int d\delta\, \mathcal{N}(\delta|0, S_{jj}) \prod_{i=1}^{N_m} \mathcal{P}(\{\widehat{N}_{ij}\}|N_{ij}(1 + b_{ij}\delta)). \tag{5.10}$$

In Eq. (5.10), N_{ij} is the cluster abundance prediction, b_{ij} is the average halo bias in the i-th mass bin and the j-th redshift bin, and $S_{jj} = \langle\delta_j^2\rangle$ is fluctuation amplitude of the matter density perturbation in the j-th redshift bin (we give more details about the halo bias and the super-sample covariance in Section 5.1.4).

The GPC likelihood satisfies the two limit cases depending on the relative contribution between average cluster abundance and super-sample covariance $\Sigma_{\rm SSC}$, where the GPC likelihood takes the form of a Poisson distribution or a Gaussian distribution. The two next sections briefly present these two limiting cases, which are the most widely used in the literature.

5.1.1.2 The Poisson likelihood

When the cluster abundance is dominant relative to the super-sample variance, i.e. when $N_{ij} \gg \Sigma_{\rm SSC}[ij][ij]$, the GPC likelihood in Eq. (5.7) simplifies to the form (Lima & Hu, 2004)

$$\mathcal{L}(\{\widehat{N}_{ij}\}) \approx \prod_{i,j=1}^{c} \mathcal{P}(\widehat{N}_{ij}|N_{ij}), \tag{5.11}$$

where $\mathcal{P}$ is the Poisson distribution in Eq. (5.8). In this approximation, cluster counts in different mass-redshift bins are independent samplings of Poisson laws with means $\langle\widehat{N}_{ij}\rangle = N_{ij}$ and each one has a variance $(\sigma_{ij})^2 = N_{ij}$. This variance N_{ij} is referred to as Poisson shot noise. It denotes the intrinsic scatter of an uncorrelated count experiment.

5.1.1.3 The Gaussian likelihood

When the shot noise N_{ij} is no longer dominant relative to the super-sample variance and for highly populated bins, i.e. $1 \ll N_{ij} \sim \Sigma_{\rm SSC}[ij][ij]$, the GPC likelihood takes the form of a multivariate Gaussian likelihood (Lima & Hu, 2004)

$$\mathcal{L}(\{\widehat{N}_{ij}\}) \approx \frac{1}{\sqrt{(2\pi)^c|\Sigma|}} \exp\left(-\frac{1}{2}\sum_{ijkl=1}^{c} [\widehat{N}_{ij} - N_{ij}]^T \Sigma^{-1}[ij][kl][\widehat{N}_{kl} - N_{kl}]\right). \tag{5.12}$$

In the above, the full covariance matrix Σ for cluster abundance is given by

$$\Sigma = \Sigma_{\rm SN} + \Sigma_{\rm SSC}, \tag{5.13}$$

where $\Sigma_{\rm SN} = {\rm diag}(\{N_{ij}\})$ is the Poisson shot noise covariance (which is diagonal). $\Sigma_{\rm SSC}$ is the super-sample covariance. In this approximation, the cluster counts are correlated Gaussian variables with extra-variance given by the super-sample covariance.

5.1.2 What metric to evaluate the accuracy of likelihoods?

Let us consider multiple realizations of our Universe which have emerged from the same cosmological model. For each realization, we use a cluster abundance likelihood to estimate a set of cosmological parameters according to the Bayes theorem. If we use the *true* model describing the statistics of cluster abundance in the Universe (the true *likelihood*), the parameter posterior

obtained in each Universe will be consistent, after accounting for the expected statistical fluctuations. If the cluster abundance likelihood is lacking, for example, the Poisson likelihood missing the super-sample covariance contribution, the estimated parameters may be biased and/or the parameter errors will be mis-estimated.

Therefore we can imagine a test for the likelihood model by comparing the dispersion of the posterior means over many realizations of the universe to the error obtained from an individual realization. More precisely, we can compare the average posterior covariance from an ensemble of the simulations to the covariance of an ensemble of posterior means. I will first present the individual and ensemble covariances before describing how this comparison is a valid metric to test the likelihood accuracy.

5.1.2.1 Individual parameter covariance

In the following, *greek* symbols $(\alpha, \beta, \gamma, \ ...)$ refer to parameter indices in the cosmological parameter space. The posterior distribution of a set of parameters $\vec{\theta} = \{\theta_\alpha\}_{1\leq\alpha\leq c_p}$ (c_p is the number of cosmological parameters, different from the data vector dimension c) is given in terms of the likelihood $\mathcal{L}_Y$ and the prior $\pi(\vec{\theta})$ by the Bayes theorem

$$\mathcal{P}_Y(\vec{\theta}|\{\widehat{N}_{ij}\}) \propto \mathcal{L}_Y(\{\widehat{N}_{ij}\}|\vec{\theta})\pi(\vec{\theta}). \tag{5.14}$$

where $\mathcal{L}_Y$ is the likelihood used for cosmological inference, i.e., the one *chosen* for the analysis, that may or may not match with the true underlying distribution. For an *individual* dataset $\{\widehat{N}_{ij}\}$, we define $\widehat{\theta}_\alpha$ as the average parameter given the posterior $\mathcal{P}_Y$

$$\widehat{\theta}_\alpha = \mathbb{E}_{\mathcal{P}_Y}[\theta_\alpha]. \tag{5.15}$$

From $\mathcal{P}_Y$, we can derive the covariance for a set of parameters $\vec{\theta}$ (that we will refer to the *parameter covariance* in this section), defined as

$$\mathcal{C}^{\rm ind}_{\alpha\beta} = \mathbb{E}_{\mathcal{P}_Y}[(\theta_\alpha - \widehat{\theta}_\alpha)(\theta_\beta - \widehat{\theta}_\beta)]. \tag{5.16}$$

When the posterior $\mathcal{P}_Y$ is close to Gaussian we can write

$$\mathcal{P}_Y(\vec{\theta}|\{\widehat{N}_{ij}\}) \propto \exp -\frac{1}{2}(\theta_\alpha - \widehat{\theta}_\alpha)([\mathcal{C}^{\rm ind}]^{-1})_{\alpha\beta}(\theta_\beta - \widehat{\theta}_\beta) + o[(\theta_\alpha - \widehat{\theta}_\alpha)^3]. \tag{5.17}$$

By inverting Eq. (5.17), the individual parameter covariance $\mathcal{C}^{\rm ind}$ is then given by

$$([\mathcal{C}^{\rm ind}]^{-1})_{\alpha\beta} = -\frac{\partial^2 \ln \mathcal{P}_Y(\vec{\theta} = \widehat{\theta})}{\partial\theta_\alpha\theta_\beta}. \tag{5.18}$$

Evaluating the above equation for the true underlying cosmology, i.e. $(\widehat{\theta})_\alpha = (\theta_0)_\alpha$, we get (see Sellentin et al. (2014))

$$([\mathcal{C}^{\rm ind}]^{-1})_{\alpha\beta} \approx \vec{N}^T_{,\alpha}\Sigma_Y^{-1}\vec{N}_{,\beta}. \tag{5.19}$$

Σ_Y is the data covariance matrix assumed to be independent of the cosmological parameters. In Eq. (5.19) we use the notation (see Table 5.1)

$$\vec{N}_{,\alpha} = \left\{\frac{\partial N_{ij}}{\partial\theta_\alpha}\right\} \tag{5.20}$$

for the derivative of the model vector $\vec{N} = \{N_{ij}\}$ with respect to the parameter θ_α (again see Table 5.1). The above expression gives the first order of the posterior covariance if the posterior

is not purely Gaussian over the parameter space. In this case, higher order derivatives of the posterior need to be considered (Sellentin et al., 2014; Wolz et al., 2012). In our case, we will see that posteriors are almost Gaussian so that Eq. (5.19) is accurate enough. We discussed the level of uncertainty provided by an experiment that is encoded in the individual parameter covariance $\mathcal{C}^{\rm ind}$. This covariance only depends on the properties of the likelihood $\mathcal{L}_Y$ that is chosen to do the cosmological analysis.

5.1.2.2 Ensemble parameter covariance

The posterior mean $\widehat{\theta}_\alpha$ is also a random variable. When re-sampling the data, it fluctuates with some scatter around its mean. For an unbiased estimator, the average of individual posterior means over many realizations gives the true underlying parameter θ_0. As for its mean, its scatter is not *directly* measurable in practice for cosmological applications, since it requires multiple realizations of the data, while our Universe can only be observed once. However, it can be estimated with simulations; we will consider this method in this work. More specifically, we focus on the measurement of the ensemble covariance of posterior means that is denoted by $\mathcal{C}^{\rm ens}$ and is given by

$$\mathcal{C}^{\rm ens}_{\alpha\beta} = \mathbb{E}_{\mathcal{L}_X}\left[\left(\left\langle\widehat{\theta}_\alpha\right\rangle - \widehat{\theta}_\alpha\right)\left(\left\langle\widehat{\theta}_\beta\right\rangle - \widehat{\theta}_\beta\right)\right], \tag{5.21}$$

where $\widehat{\theta}_\alpha$ is the individual posterior mean defined in Eq. (5.15), derived from the individual posterior $\mathcal{P}_Y$ in Eq. (5.14). It is sometimes called the Frequentist covariance of Bayesian estimators (Efron, 2015). Contrary to the pure Bayesian analysis mentioned before (to derive $\mathcal{C}^{\rm ind}$ from parameter posterior), the ensemble covariance depends on the likelihood $\mathcal{L}_X$, which is the true underlying likelihood describing the true statistical properties of the data (namely, the true likelihood of the observable). By definition, the parameter covariance $\mathcal{C}^{\rm ens}$ depends on both the analysis and underlying likelihoods $\mathcal{L}_Y$ and $\mathcal{L}_X$ (respectively from $\widehat{\theta}$ in Eq. (5.15) and the average $\mathbb{E}_{\mathcal{L}_X}[\cdot]$) that may be different.

We note that $\mathcal{C}^{\rm ind}$ and $\mathcal{C}^{\rm ens}$ give two different answers to two different questions. On the one hand, $\mathcal{C}^{\rm ind}$ quantifies the dispersion of the parameter scaling the cosmological models, that is considered as a random variable. The parameters fluctuate according to the conditional probability density function $\mathcal{P}_Y$, giving the distribution of all *possible* values that the parameters can have. Then, the individual (co)variance can be used to define some *credible* intervals, usually denoted by $[\widehat{\theta}_\alpha - (\mathcal{C}^{\rm ind}_{\alpha\alpha})^{1/2}, \widehat{\theta}_\alpha + (\mathcal{C}^{\rm ind}_{\alpha\alpha})^{1/2}]$. This interval gives the space range where the probability to find each cosmological parameter is $p = 0.68$ (after having marginalized over all the other parameters), accounting for the prior knowledge and likelihood response. On the other hand, $\mathcal{C}^{\rm ens}$ gives complementary and independent information on parameter inference. Formally, the ensemble covariance provides the parameter *confidence* intervals, usually denoted by $[(\theta_0)_\alpha - (\mathcal{C}^{\rm ens}_{\alpha\alpha})^{1/2}, (\theta_0)_\alpha + (\mathcal{C}^{\rm ens}_{\alpha\alpha})^{1/2}]$ (if the parameter estimator is not biased). The posterior mean $\widehat{\theta}_\alpha$ is assumed to be found in this interval ~ 68 times over 100 independent re-sampling of the data $\{\widehat{N}_{ij}\}$, that follow the latent likelihood $\mathcal{L}_X$ at fixed cosmology $\vec{\theta}_0$. So according to the Frequentist approach, the true parameter θ_0 is assumed to be fixed, and data fluctuate according to the latent likelihood $\mathcal{L}_X(\cdot|\theta_0)$, contrary to the Bayesian approach where θ is considered as a random variable.

The "ensemble" method we consider in this work is not a pure Frequentist approach, since the estimator $\widehat{\theta}_\alpha$ is Bayesian, i.e., it relies on the full posterior $\mathcal{P}_Y$ (that includes prior knowledge) and not only on the analysis likelihood $\mathcal{L}_Y$. That would be the case if we had considered the maximum likelihood estimator (MLE), that is widely used in Frequentist parameter inference in particle physics and is given by

$$\widehat{\theta}_{\rm MLE}(\{\widehat{N}_{ij}\}) = \max_{\vec{\theta}\in\mathrm{R}^d}[\mathcal{L}_Y(\{\widehat{N}_{ij}\}|\vec{\theta})]. \tag{5.22}$$

We will see that our parameter posteriors are very close to Gaussian, so that $\widehat{\theta}_{\rm MLE} \approx \widehat{\theta}$, i.e., the maximum likelihood estimator and the posterior mean will coincide. Intuitively, we can consider that the analysis likelihood that we use for the cosmological analysis corresponds to the latent one, i.e., our modeling of the observable's statistical properties is "correct". In this case, we have that $X = Y$ and using the formalism developed by Fisher (1935) (see also Tegmark et al. (1997); Heavens et al. (2014)) and the Cramer-Rao inequality (Cramer, 1952) we have that $\mathcal{C}^{\rm ens}$ verifies

$$\mathcal{C}^{\rm ens}_{\alpha\beta} \geq (\mathrm{F}^{-1})_{\alpha\beta}, \tag{5.23}$$

where F is the Fisher matrix such as

$$\mathrm{F}_{\alpha\beta} = \mathbb{E}_{\mathcal{L}_Y}\left[\left(\frac{\partial}{\partial\theta_\alpha}\ln\mathcal{L}_Y\right)\left(\frac{\partial}{\partial\theta_\beta}\ln\mathcal{L}_Y\right)\right], \tag{5.24}$$

where the expectation runs over the realizations of the data that follow the likelihood distribution. For Gaussian and Poisson multivariate likelihoods, the Fisher matrix is given by (see Appendix D)

$$\mathrm{F}_{\alpha\beta} = \vec{N}^T_{,\alpha}\Sigma_Y^{-1}\vec{N}_{,\beta}, \tag{5.25}$$

where $\vec{N}_{,\alpha}$ is given in Eq. (5.20). So the Fisher matrix forecasts both the ensemble covariance of parameters (from a Frequentist point of view) and the individual parameter covariance in a Bayesian analysis in Eq. (5.19) as they are both equivalent, i.e., $\mathcal{C}^{\rm ens} = \mathcal{C}^{\rm ind}$ when the latent and analysis likelihoods are equal. Formally, this result tells that the *credible* intervals that are given by the parameter posterior coincide with the Frequentist *confidence* intervals, obtained by looking at the distribution of posterior means when re-sampling the data.

5.1.2.3 Robustness of cluster likelihoods

So far we discussed the specific case where the analysis likelihood and the latent likelihood are equivalent (i.e., $X = Y$). However, in most cases, the analysis likelihood approximates the latent statistical processes driving the data, that is more complex. When the analysis likelihood $\mathcal{L}_Y \neq \mathcal{L}_X$, the individual and ensemble parameter covariances are no longer equal and the standard Fisher formalism is not sufficient. In particular (e.g., see the Appendix D), we can show that if Σ_X denotes the data covariance matrix derived from the latent distribution $\mathcal{L}_X$ (for example in the case of a Gaussian latent likelihood), the ensemble covariance $\mathcal{C}^{\rm ens}$ can be forecast such as

$$\mathcal{C}^{\rm ens}_{\alpha\beta} = \vec{u}^T_\alpha\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{u}_\beta. \tag{5.26}$$

where $\vec{u}_\alpha$ has the dimension of a data vector and its $i-$th component is given by

$$[u_\alpha]_i = \sum_{\gamma=1}^{c_p}\mathcal{C}^{\rm ind}_{\alpha\gamma}N_{i,\gamma}. \tag{5.27}$$

Also, using the Cauchy-Schwarz inequality, the Cramer-Rao lower bound in Eq. (5.23) is modified when the analysis and the latent likelihoods have different count covariance matrices (see Appendix D)

$$\mathcal{C}^{\rm ens}_{\alpha\alpha} \geq \frac{[u^T_\alpha\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{N}_{,\alpha}]^2}{\vec{N}^T_{,\alpha}\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{N}_{,\alpha}}. \tag{5.28}$$

As a sanity check, if we consider the two data covariance matrices to be equal[3], i.e. , $\Sigma_Y = \Sigma_X$ and plugging Eq. (5.19) into Eq. (5.26) then

$$\mathcal{C}^{\rm ind} = \mathcal{C}^{\rm ens}. \tag{5.29}$$

[3]This is the case when the latent likelihood and the likelihood used for the posterior estimation parameters are the same.

We can therefore justify the comparison of these two different parameter covariances as a metric to test the robustness of a given likelihood. Additionally in the case $\Sigma_Y \neq \Sigma_X$, we have,

$$\mathcal{C}^{\rm ind} \neq \mathcal{C}^{\rm ens}. \tag{5.30}$$

In the following we assume that this result also holds for the Gauss-Poisson likelihood, although not formally demonstrated. But this result holds for a Gaussian and a Poisson likelihood. Since we find the latent likelihood to be close to Gaussian, this is likely true in our specific case.

This equivalence is discussed in more general terms in Efron (2015), and also in Penna-Lima et al. (2014) for an unbinned Poisson likelihood, and in the Bernstein-von Mises theorem (see e.g. Johnstone (2015)). Similar calculations are made in Percival et al. (2022) in order to obtain a prior which gives equivalent results for Frequentist and Bayesian analyses. For reference, we also mention that the methodology we consider is valid for flat priors, as well as when the likelihood variations upon the parameter space are well contained within the prior ranges (i.e., containing roughly up to the third σ region of the full posterior). For the specific case of non-flat prior (which is not the case here) the equality $\mathcal{C}^{\rm ind} = \mathcal{C}^{\rm ens}$ is no longer true even if the two likelihoods are equal (we refer the reader to Appendix D where we detail the calculations for the non-flat prior case).

To use the metric described above, a large set of simulations is required, that we present in the next section.

5.1.3 The 1,000 PINOCCHIO mocks

The formation of dark matter halos is a highly non-linear dynamical process. In most cases (in the non-linear regime at cluster scales), there are no analytical approximations to derive the gravitational evolution of density perturbations. To overcome this problem, N-body numerical simulations are usually performed to build realistic dark matter halo catalogs; from a set of discrete tracers (or "particles") of the initial matter distribution, the evolution of the six-dimensional phase space (position and velocity) is obtained by solving differential equations accounting for the mutual gravity force between particles. The halos are then identified by using a halo finder that associates neighboring particles in groups according to some maximum distance criteria. Although allowing to study in great detail the history of the hierarchical formation process of large-scale structures (see the review by Bagla (2005)), they are computationally expensive.

Therefore, in order to have many simulated realizations at fixed cosmology, we use a set of halo catalogs generated by an approximate method called PINOCCHIO (PINpointing Orbit-Crossing Collapsed HIerarchical Objects, Monaco et al. (2002); Munari et al. (2017)). PINOCCHIO is an algorithm which generates dark matter halo catalogs using Lagrangian Perturbation Theory (Moutarde et al., 1991; Buchert, 1992; Bouchet et al., 1995) and ellipsoidal collapse (Bond & Myers (1996); Eisenstein & Loeb (1995), and Monaco (1997)). Relative to N-body simulations, the generation of dark matter halo catalogs is significantly faster (see Monaco (2016), showing that PINOCCHIO is quicker by a factor of $1,000$). Moreover, the accuracy of the PINOCCHIO algorithm is within $\sim 5-10\%$ for reproducing the halo mass function, halo two-point statistics, and the halo bias at all relevant redshifts used in this study (Munari et al., 2017). The PINOCCHIO algorithm is particularly relevant to our purpose, enabling fast generation of thousands of dark matter halo catalogs.

We were provided by the PINOCCHIO developers with a set of 1,000 light-cone catalogs which cover $10,313$ deg^2 ($\sim$ a quarter of the sky), which is relevant for wide cosmological surveys such as Rubin LSST and *Euclid*. The dark matter catalogs cover a redshift range from $z = 0$ to $z = 2.5$ and halos have virial masses above $2.45 \times 10^{13}\ h^{-1}{\rm M}_\odot$. The simulated cosmology

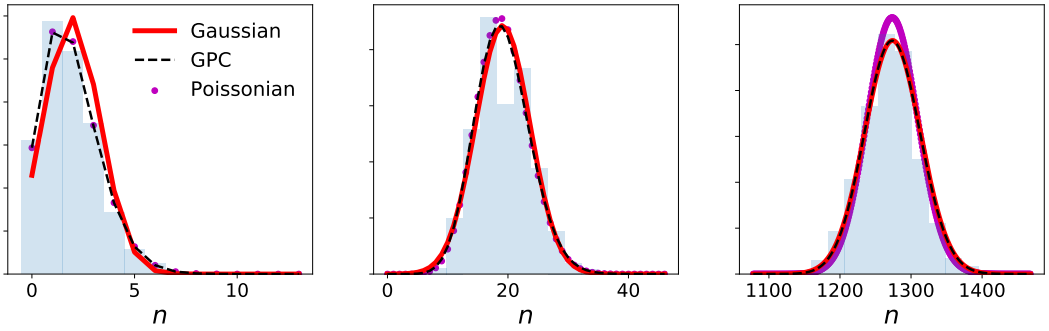

Figure 5.1: Blue histograms: distribution of the cluster counts over the 1000 PINOCCHIO simulations for a low (left), intermediate (middle) and highly (right) populated bin. Magenta symbols: Single-variate Poisson distribution from Eq. (5.11). Red line: single variate Gaussian distribution from Eq. (5.12). Black dashed line: GPC likelihood from Eq. (5.7).

is the best-fit cosmology from Planck Collaboration et al. (2014), i.e., $\Omega_m = 0.30711$, $\Omega_b = 0.048254$, h $= 0.6777$, $n_s = 0.96$, $\sigma_8 = 0.8288$. These same catalogs were used in the analysis of Fumagalli et al. (2021). To bypass the small inaccuracy of the PINOCCHIO halo mass function, PINOCCHIO halo masses have been re-scaled to match the mass function averaged over 1,000 realizations with the Despali et al. (2015) analytic fit of the halo mass function. This procedure allows us to preserve the super-sample variance of each individual realization while forcing the average to reproduce a known mass function; this is important as we want to isolate the effect of the likelihood from all other modeling systematic effects.

In Fig. 5.1, we consider three different mass-redshift bins, corresponding respectively to a low, medium and high average number of dark matter halos per bin (from left to right, $\langle N \rangle \sim 2$, $\langle N \rangle \sim 20$ and $\langle N \rangle \sim 1,300$). For each of these bins, the shaded histogram corresponds to cluster counts over the 1,000 PINOCCHIO dark matter catalogs. Doing so, we have estimates of the cluster count likelihood in each of those three bins. For each binning, we plot the 3 different likelihoods discussed in Section 5.1.1, the Poisson[4] (purple dots), the Gaussian (red line) and the Gauss-Poisson Compound[5] (black dashed line) likelihoods. We used the Core Cosmology Library (Chisari et al., 2019) to predict the halo mass function and the halo bias from Tinker et al. (2010). We see for the small population (left panel) bins that the distribution of the cluster abundance is not Gaussian; additionally, we see that for highly populated bins (right panel) the distribution of counts is not Poissonian; this is due to the SSC contribution that the Poisson distribution does not account for.

5.1.4 Cluster count covariance matrix

For the Gaussian and GPC likelihoods, the count covariance Σ needs to be calculated. The abundance covariance can be predicted from theoretical considerations and some simplifying approximations, for a given cosmology, mass function and halo bias prescriptions. As mentioned in Section 5.1.1, the count covariance has two contributions $\Sigma^{\rm th} = \Sigma^{\rm th}_{\rm SN} + \Sigma^{\rm th}_{\rm SSC}$, where $\Sigma^{\rm th}_{\rm SN}$ is the shot noise variance and $\Sigma^{\rm th}_{\rm SSC}$ is the super-sample covariance.

The shot noise contribution in the i-th mass bin and j-th redshift bin is equal to the predicted

[4]The Poisson distribution is discrete, so it is represented by points and not curve.
[5]The GPC distribution is discrete too, but here it is represented in dashed lines to differentiate with the Poisson likelihood.

count N_{ij}, as given by equation Eq. (5.4), i.e.,

$$\Sigma^{\rm th}_{\rm SN}[ij][kl] = \delta^K_{ik}\delta^K_{jl}N_{ij}, \tag{5.31}$$

where δ^K_{ij} is the Kronecker Delta function. As such, it depends on the cosmology and the halo mass function. The super-sample covariance $\Sigma_{\rm SSC}$ between two counts N_{ij} and N_{kl} is given by (Hu & Kravtsov, 2003)

$$\Sigma_{\rm SSC}[ij][kl] = \int_{z_j}^{z_{j+1}} dV_1 \int_{z_l}^{z_{l+1}} dV_2 \frac{\partial\mathcal{O}_i(z_1)}{\partial\delta}\frac{\partial\mathcal{O}_k(z_2)}{\partial\delta}\sigma^2(z_1,z_2), \tag{5.32}$$

where the quantity

$$\frac{\partial\mathcal{O}_i(z)}{\partial\delta} = \int_{m_i}^{m_{i+1}} dm b(m,z)\frac{dn(m,z)}{dm} \tag{5.33}$$

is the *response* of the probe $\mathcal{O}_i$ defined as

$$\mathcal{O}_i(z) = \int_{m_i}^{m_{i+1}} dm \frac{dn(m,z)}{dm}(1+b(m,z)\delta) \tag{5.34}$$

relative to the background density field. The quantity $\sigma^2(z_1,z_2)$ is the amplitude of the fluctuations between two different redshifts and depends on the survey window function. Considering the full sky area, $\sigma^2(z_1,z_2)$ is defined as

$$\sigma^2(z_1,z_2) = \frac{1}{2\pi^2}\int k^2 dk j_0(kw(z_1))j_0(kw(z_2))P_m(k|z_1,z_2), \tag{5.35}$$

where w is the line-of-sight comoving distance, and j_0 is the spherical Bessel function of order 0, $P_m(k|z_1,z_2)$ is the matter power spectrum at redshift z_1 and z_2. Assuming that $\frac{\partial\mathcal{O}_i(z)}{\partial\delta}$ varies slowly with redshift compared to the matter fluctuation amplitude, we get the approximation (Lacasa et al., 2018)

$$\Sigma^{\rm th}_{\rm SSC}[ij][kl] \approx \langle bN\rangle_{ij}\langle bN\rangle_{kl}S_{jl}, \tag{5.36}$$

where the quantity $\langle bN\rangle_{ij}$ is the average halo bias in the i-th mass and j-th redshift bin and S_{jl} is the matter density fluctuation covariance between the two redshift bins j and l. The average halo bias is expressed as

$$\langle bN\rangle_{ij} = \Omega_S \int_{z_j}^{z_{j+1}} dz \int_{m_i}^{m_{i+1}} dm \frac{dn(m,z)}{dm}\frac{d^2V(z)}{dzd\Omega}b(m,z), \tag{5.37}$$

where $b(m,z)$ is the halo bias at fixed mass and redshift. For a full-sky survey, the matrix $S^{\rm fullsky}_{jl}$ is the average fluctuation amplitude over the two separate volumes and is defined by

$$S^{\rm fullsky}_{jl} = \frac{1}{V_jV_l}\int_{z_j}^{z_{j+1}} dV_1 \int_{z_l}^{z_{l+1}} dV_2\ \sigma^2(z_1,z_2). \tag{5.38}$$

For partial sky survey, we rescale Eq. (5.38) by the sky fraction and use $S_{jl} \approx S^{\rm fullsky}_{jl}/f_{\rm sky} = S^{\rm fullsky}_{jl} \times 4\pi/\Omega_S$ (as proposed in Lacasa et al. (2018)). While this is only an approximation, Gouyou Beauchamps et al. (2022) shows that this performs well for sufficiently large sky areas, such as the one covered by the PINOCCHIO simulations. In practice, we use the `PySSC` package[6] to compute S_{jl}.

[6] `https://github.com/fabienlacasa/PySSC`.

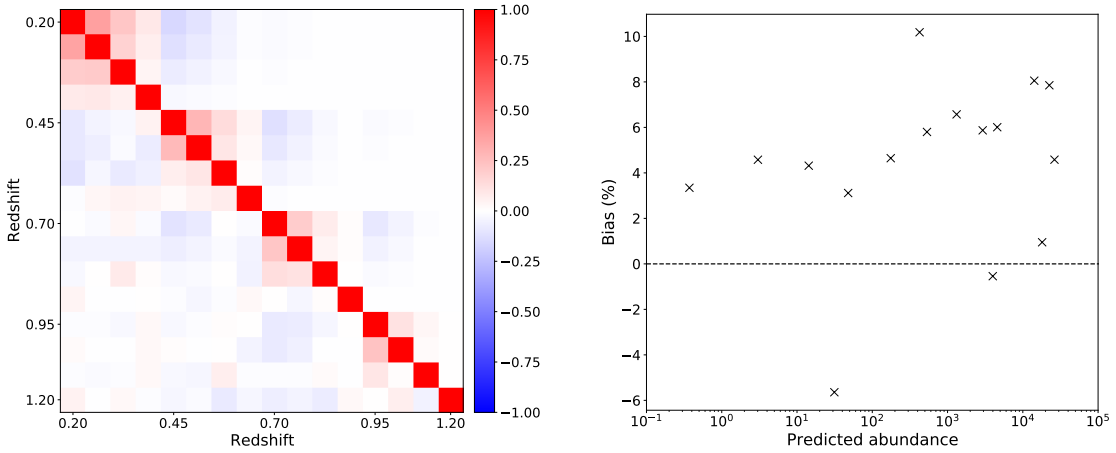

Figure 5.2: *Left*: Correlation matrix of cluster counts for the `coarse` binning scheme. The matrix is organized in four major redshift bins increasing from left to right and top to bottom (associated to the four diagonal blocks) and each of them includes four mass bins. The upper triangle corresponds to the predicted correlation matrix and the lower triangle to the estimated one from Eq. (5.39). *Right*: Relative bias of the predicted variance of cluster abundance relative to the estimated variance computed in Eq. (5.41) as a function of predicted abundance.

To check our use of `PySSC` and the PINOCCHIO simulations, we compare the theoretical prediction of the abundance covariance, given by $\Sigma^{\rm th} = \Sigma^{\rm th}_{\rm SSC} + \Sigma^{\rm th}_{\rm SN}$ ($\Sigma^{\rm th}_{\rm SN}$ is given in Eq. (5.31)), to the covariance directly estimated from the data. The estimated full covariance $\widehat{\Sigma}^{\rm data}[ij][kl]$ between mass-redshift bins ij and kl is calculated from the 1000 simulations, such as

$$\widehat{\Sigma}^{\rm data}[ij][kl] = \frac{1}{N_{\rm sim}-1}\sum_{n=1}^{N_{\rm sim}}\left([\widehat{N}_{ij}]_n - \bar{N}_{ij}\right)\left([\widehat{N}_{kl}]_n - \bar{N}_{kl}\right). \tag{5.39}$$

The left panel in Fig. 5.2 shows the correlation matrix, defined as,

$$R[ij][kl] = \frac{\Sigma[ij][kl]}{\sqrt{\Sigma[ij][ij]\Sigma[kl][kl]}}. \tag{5.40}$$

for the data-derived covariance (lower triangle) and the theoretical covariance prediction (upper triangle). This is shown for the coarsest binning scheme (see Fig. 5.3) with the correlation matrix organized in four redshift bins, appearing as block diagonal, inside which the four mass bins are added.

We find the correlation between bins to be well-represented by the theoretical prediction for the block diagonal terms and the first off-diagonal terms. However, for more distant correlations, statistical noise dominates the data-derived covariance matrix (lower triangle) while the correlation drops to zero in the theoretical estimation (upper triangle).

To further compare the two estimations, we compute the relative difference of the variance (diagonal elements) between the two approaches, defined as

$$b_{ij} = \frac{\Sigma^{\rm data}[ij][ij] - \Sigma^{\rm th}[ij][ij]}{\Sigma^{\rm th}[ij][ij]}. \tag{5.41}$$

The results are represented as a function of the mean abundance, for the 16 diagonal elements of the 4×4 `coarse` binning scheme, in the right panel in Fig. 5.2. We find a bias of about $\sim 4-8\%$, which is expected according to the precision of PINOCCHIO two-point statistics.

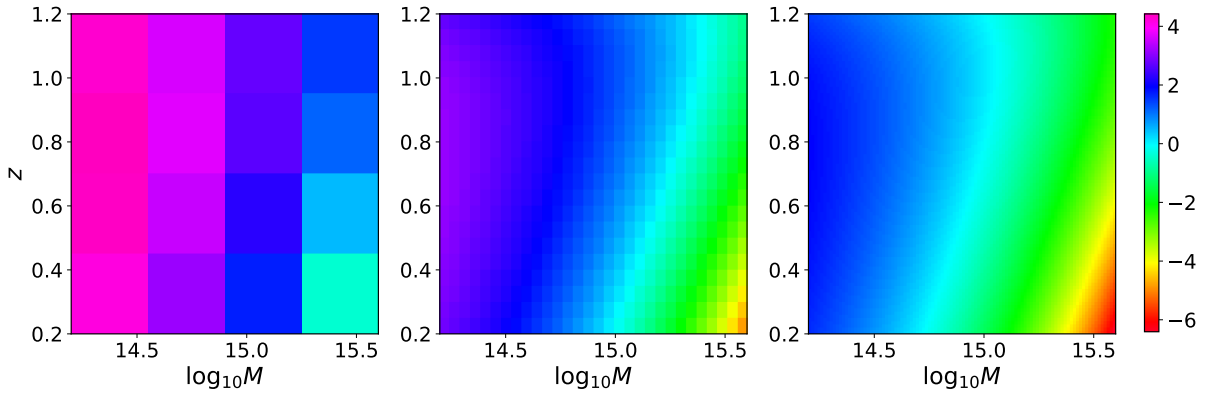

Figure 5.3: 2D histograms of $\log_{10} N_{ij}$ abundance predictions (i-th mass and j-th redshift bin) for the `coarse` (left), `medium` (center), and `fine` (right) binning schemes of Table 5.2.

This check validates our computation of the theoretical covariance, which we favor in the analysis as the data-derived covariance is affected by statistical noise due to the limited number of samples.

5.1.5 Methodology

With all these ingredients in place, the method we use to test each likelihood is the following:

- for each one of the 1,000 mock catalogs we obtain a posterior (using importance sampling presented in Section 5.1.5.2);
- for each of these posteriors we calculate the posterior mean and posterior covariance;
- from the ensemble of individual posterior means we calculate the "average" of the posterior means and the covariance of the posterior means over the 1,000 samples;
- we then compare the results of the individual posteriors (individual covariances) to the ensemble of the posteriors (covariance of posterior means). If they match, the chosen likelihood satisfactorily represents the data;
- to validate the approach and confirm the interpretation of the results from the simulations, we compute the Fisher forecasts from Eq. (5.25) and ensemble forecasts from Eq. (5.26) for comparison.

This test is performed for each of the likelihoods for three different binning schemes. We detail below the binning schemes in the mass-redshift plane, the importance sampling method that is used to estimate parameter posteriors, and how to compute the ensemble covariance.

5.1.5.1 Mass and redshift ranges, and binning setups

For this analysis, we adopt a binning scheme of the ($\log_{10}$)mass-redshift plane within the redshift range $[z_{\rm min}, z_{\rm max}] = [0.2, 1.2]$ (redshift window typically spanned efficiently by the redMaPPer cluster finder, (see Chapter 4) and the mass range $[\log_{10}(m_{\rm min}/M_\odot), \log_{10}(m_{\rm max}/M_\odot)] = [14.2, 15.6]$. We consider a `coarse`, a `medium` and a `fine` binning which are detailed in Table 5.2. Both redshift and ($\log_{10}$)mass bins are equally spaced.

	`coarse`	`medium`	`fine`
Number of redshift bins N_z	4	20	100
Number of mass bins N_m	4	30	100
Total number of bins $N_{\rm tot}$	16	600	10,000

Table 5.2: Characteristics of the `coarse`, `medium` and `fine` two-dimensional binnings of the mass-redshift plane chosen in this analysis.

Fig. 5.3 shows the predicted abundance at fiducial cosmology and for the three binning schemes. For the `coarse` case, the number of expected halos per bin never drops below one and the bins are generally well-populated (there is an average of 5,000 halos per mass and redshift over all the binning scheme, ranging from 1 to 25,000). For the `medium` and `fine` grids a significant portion of the mass-redshift plane is expected to have only a few or no halos. From this, we expect the Gaussian approximation to be valid for the `coarse` binning scheme, while this may be a less accurate statistical description for the two other cases. The different binning schemes allow us to explore a variety of regimes, from the shot noise-dominated regime to the non-negligible super-sample variance regime, and to test the performance of the likelihoods described in Section 5.1.1 in each regime. It is also the opportunity to test whether one of these binning schemes provides more precise, yet robust, constraints for Rubin LSST/*Euclid* surveys.

5.1.5.2 Posterior estimation using importance sampling

To undertake this analysis, from the cluster abundance data vector $\widehat{N}$, we perform a cosmological analysis for the 1,000 simulations, three binning schemes and three likelihoods. Therefore we wish to obtain 9,000 posteriors in the parameter space $\vec{\theta} = (\Omega_m, \sigma_8)$, that are the two main cosmological parameters to describe the growth of structure in the Universe and what cluster abundance is most sensitive to. This methodology could be extended to more cosmological parameters, which we explore in the next Section 5.2. Getting 9,000 constraints would be computationally expensive for a standard Markov Chain Monte Carlo method. Instead, we use importance sampling (IS) to efficiently sample the joint posterior distribution of the parameters, that allows us to calculate the theoretical prediction of abundances once for the 9,000 analyses. We briefly recap the framework of importance sampling below (see Elvira & Martino (2021) for a recent review).

Importance sampling can be used to estimate the statistical properties of the posterior distribution $\mathcal{P}_Y(\vec{\theta}|\vec{N})$ in Eq. (5.14) from a random sample of points that follows a proposal distribution q. Considering a random variable $X \sim p$, the expectation of X is given by

$$\mathrm{E}_p[X] = \int_{-\infty}^{+\infty} dx p(x) x. \tag{5.42}$$

The above average can be estimated from the sampling of another random variable Y following a proposal distribution q. We call $\widehat{X}$ the unbiased estimator of $\mathrm{E}_p[X]$,

$$\widehat{X} = \frac{1}{N_q} \sum_{k=1}^{N_q} w_k y_k, \tag{5.43}$$

where N_q is the number of samples and the weights, w_k are defined as,

$$w_k = \frac{p(y_k)}{q(y_k)}. \tag{5.44}$$

The posterior p is directly estimated from the weighted histogram of y_k random samples with weights w_k. For the proposal q, we use a bivariate Gaussian distribution centered on the (Ω_m, σ_8) input values of the PINOCCHIO simulations. The parameter covariance matrix of the proposal distribution q is chosen as the Fisher forecast (given in Eq. (5.25)) for the `coarse` binning scheme, multiplied by a factor 36. This ensures that the proposal distribution encompasses a large enough region of the parameter space covered by the ensemble of σ contours of each individual posterior. The mean of X is therefore expressed as a weighted mean over the sample Y. In this work, we focus on the recovered cosmological parameters and associated errors per simulation, i.e., the posterior mean and the individual posterior covariance. The estimated individual average parameter $\widehat{\theta}_\alpha^{\rm ind}$ is given by,

$$\widehat{\theta}_\alpha^{\rm ind} = \sum_{k=1}^{N_q} w_k [\theta_\alpha]_k, \tag{5.45}$$

where $[\theta_\alpha]_k$ is the sampled parameter value θ_α at position k. From this, one may also estimate the unbiased weighted covariance matrix of $\vec{\theta} = (\Omega_m, \sigma_8)$ as

$$\widehat{\mathrm{C}_{\alpha\beta}^{\rm ind}} = \sum_{k=1}^{N_q} \widetilde{w}_k ([\theta_\alpha]_k - \widehat{\theta}_\alpha^{\rm ind})([\theta_\beta]_k - \widehat{\theta}_\beta^{\rm ind}), \tag{5.46}$$

with

$$\widetilde{w}_k = w_k \frac{\sum_{k=1}^{N_q} w_k}{\left(\sum_{k=1}^{N_q} w_k\right)^2 - \sum_{k=1}^{N_q} w_k^2}. \tag{5.47}$$

To compute the posterior we use the different likelihoods $\mathcal{L}$ presented in Section 5.1.1, and we use flat priors for the parameters, namely $\Omega_m \in [0.293, 0.321]$ and $\sigma_8 \in [0.820, 0.836]$. These ranges were chosen as 12 times the square root of the Fisher variances on each parameter given in Eq. (5.25). The proposal has to be *wide* enough to contain 2σ contours for each posterior, and the proposal sample has to be large enough to reduce the possible artificial sources of scatter and degeneracy in the posterior estimation.

Compared to the more widely used MCMC, importance sampling yields by construction the same results. Fig. 5.4 shows the marginal distribution on Ω_m and σ_8 obtained with the importance sampling method (orange), sampling from the proposal distribution q (green) for the cluster count observed in a single PINOCCHIO mock and considering the Poisson likelihood. The corresponding MCMC posteriors are shown in blue and are in good agreement with the IS ones. They differ due to statistical noise associated to both methods, for instance, if we re-draw samples from the q distribution for the IS method or re-draw the initial walkers for the MCMC (or else consider one more iteration of Markov chains). This statistical noise will also increase/decrease depending on the number of samples in both cases. Importance sampling is however much faster to run since the count prediction is evaluated only once at each location of the q-sample. In particular, it is much faster for the GPC likelihood where the function is computationally taxing (for example 15h for Importance Sampling and for the fine binning, and 20 times longer for MCMC). For the Gaussian and Poisson likelihoods, that are easier to compute compared to GPC, the computation of IS posteriors lasts a few seconds to a few minutes, compared to a few hours for MCMC on a single CPU at CC-IN2P3.

The PINOCCHIO mock used in Fig. 5.4 gives a posterior that is *by statistical chance* compatible at the $\leq 1\sigma$ level to the input cosmology (see the Section 5.1.2). The level of agreement should vary from one simulation to another. As an illustration, the four colored contours in Fig. 5.5 correspond to the 1 and 2σ confidence intervals obtained with four of the PINOCCHIO simulations presented in Section 5.1.3 and using the binned Poisson likelihood. These

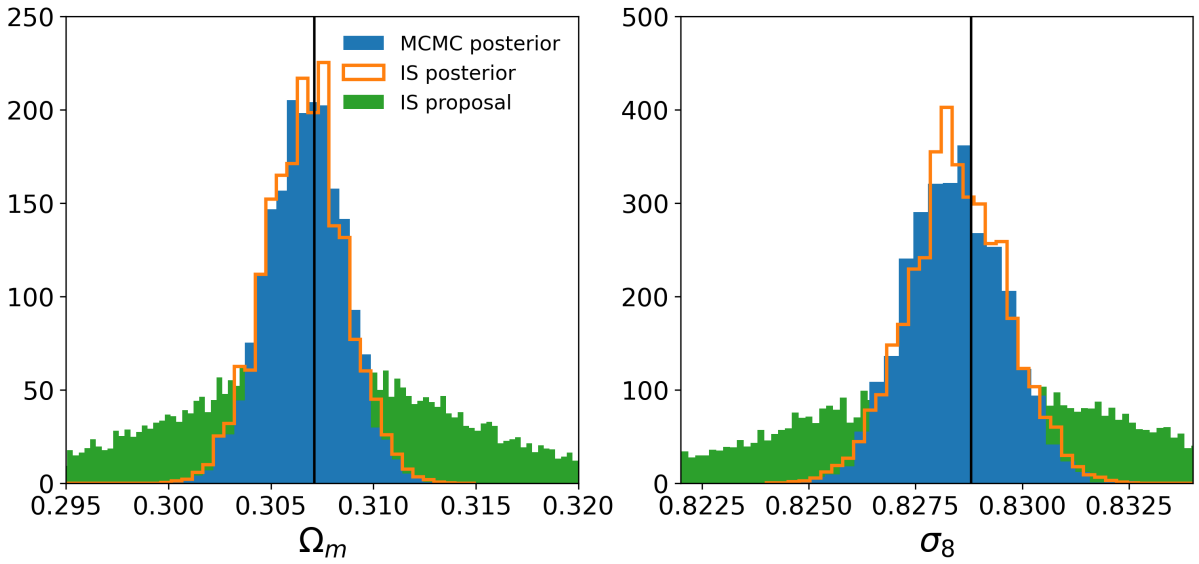

Figure 5.4: Marginal posterior distributions for Ω_m and σ_8, considering the importance sampling method (orange) based on drawn samples (green) from the proposal distribution. The MCMC posteriors are represented in blue. The input PINOCCHIO cosmology corresponds to the vertical lines. The MCMC run took several hours, while only a few minutes for IS.

are obtained as the weighted histograms of the q-sample (blue points). The contours have been obtained directly from the 10^4 samples, we see that the black posterior is 1σ compatible with input cosmology when the yellow is compatible at more than 2σ.

5.1.5.3 Estimating the ensemble posterior covariance

After the 1,000 posteriors have been estimated using importance sampling, we can estimate the ensemble posterior covariance. The estimator $\widehat{\theta}_\alpha$ in Eq. (5.15) is noted $[\widehat{\theta}^{\rm ind}_\alpha]_k$ for the k-th simulation and is calculated as in Eq. (5.45). The mean cosmological parameter $\langle\widehat{\theta}_\alpha\rangle$ over the 1,000 simulations that appears in Eq. (5.21) is now given by

$$\widehat{\theta}^{\rm ens}_\alpha = \frac{1}{N_{\rm sim}}\sum_{k=1}^{N_{\rm sim}}[\widehat{\theta}^{\rm ind}_\alpha]_k. \tag{5.48}$$

The corresponding ensemble covariance matrix is estimated as,

$$\widehat{\mathcal{C}^{\rm ens}_{\alpha\beta}} = \frac{1}{N_{\rm sim}-1}\sum_{k=1}^{N_{\rm sim}}([\widehat{\theta}^{\rm ind}_\alpha]_k - \widehat{\theta}^{\rm ens}_\alpha)([\widehat{\theta}^{\rm ind}_\beta]_k - \widehat{\theta}^{\rm ens}_\beta). \tag{5.49}$$

From this covariance matrix, we can extract the ensemble variance and correlation coefficient, and compare to those obtained from the individual posteriors calculated in Eq. (5.46).

5.1.5.4 Software

Before describing the results of this analysis, I detail below some technical and computational aspects I developed for this work. To set up the methodology, we analyzed a large cosmological dataset provided by the PINOCCHIO developers (Pierluigi Monaco, Tom Theuns and Giuliano Taffoni), and for that, we required the computation and data-processing resources provided by the CNRS/IN2P3 Computing Center (Lyon - France).

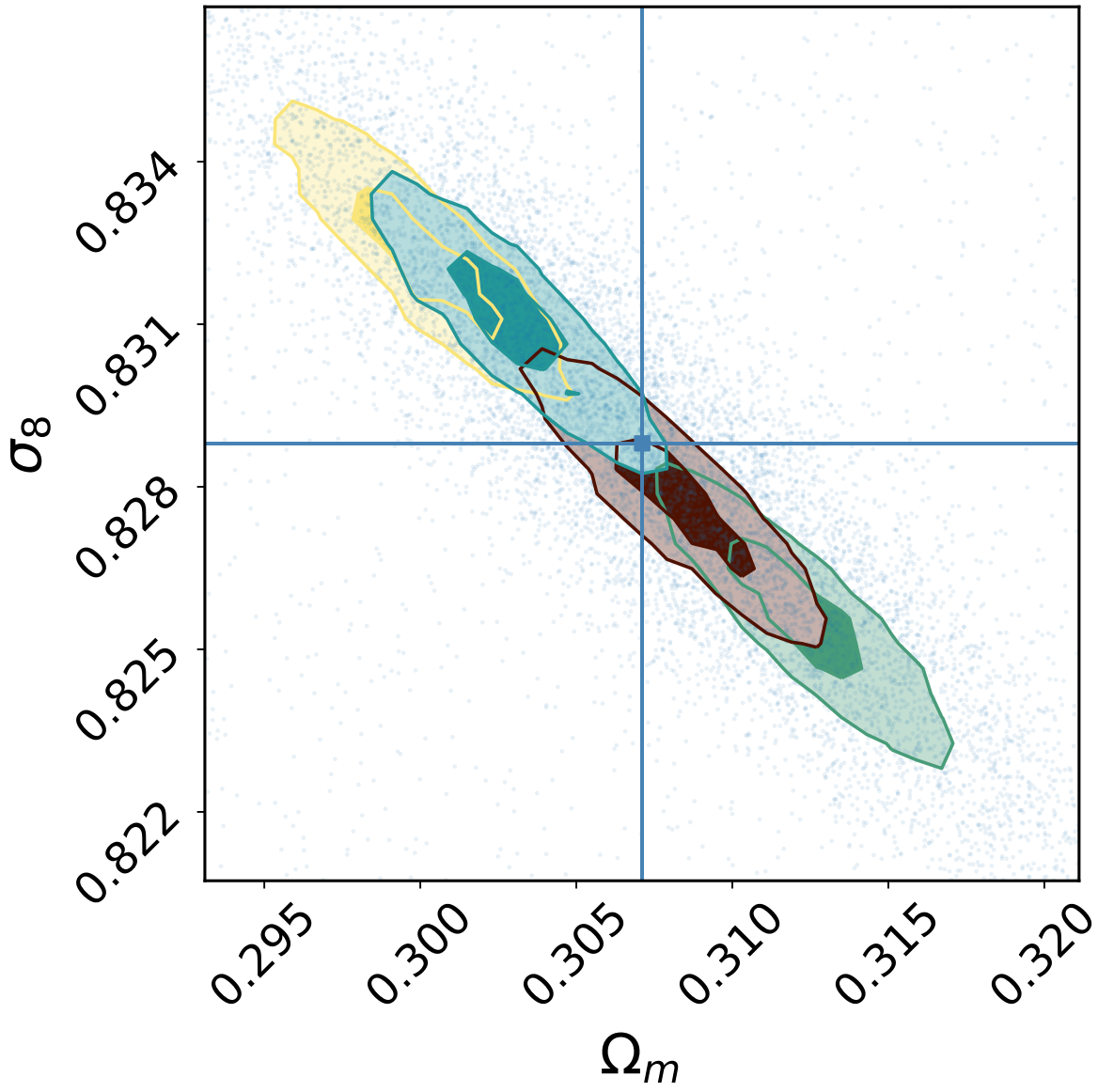

Figure 5.5: Illustration of the importance sampling input and output. Background blue points: sample of 10^4 points following the proposal distribution q. The contour plots correspond to the 1 and $2-\sigma$ levels of the posterior distributions of Ω_m and σ_8, for four different PINOCCHIO simulations. Vertical and horizontal lines correspond to the input simulation parameters.

We used the various functionalities of Slurm, the CC-IN2P3 scheduling software, to submit thousands of jobs in parallel to first pre-compute the count modelings and second to compute the IS weights and estimate the individual covariance matrices for each PINOCCHIO mocks, according to the various inference setups we defined in Table 5.2. We also used multi-processing to optimize our codes[7].

We use the `PySSC` code, necessary to compute the amplitude of matter fluctuations, that we combined with our code (based on the Core Cosmology Library that predicts the halo bias and halo mass function) to derive the full data covariance. To optimize even more the run time of the cluster abundance prediction at each location of the $q-$sample, we compute the double integral in Eq. (5.4) by interpolating the Despali et al. (2015) halo mass and redshift distribution (using the Core Cosmology Library Chisari et al. (2019)), that is previously tabulated on a two-dimensional grid between the minimum and maximum of each mass and redshift axes. The several choices we made enabled us to have a satisfactory repeatability of $\sim 10-20$ hours to re-do the analysis for all the binning set-ups.

5.1.6 Results

In this section, we present the results of testing the accuracy of likelihoods using importance sampling and the set of the 1,000 PINOCCHIO simulations. First, we compare the individual posterior means to the ensemble posterior mean, in order to check for a bias in the likelihoods. Second, we compare the individual posterior variances and individual posterior correlation coefficients to the ensemble posterior variance and ensemble correlation coefficients, to check the robustness of the estimated posteriors. Finally, we respectively reduce the sky area and increase the lower mass-cut value of the analysis to see if a regime exists in which the Poisson likelihood or the GPC likelihood obtain tighter, yet still robust, cosmological constraints than the Gaussian likelihood.

5.1.6.1 Bias on the posterior mean

Figure 5.6 shows, for the `coarse` 4×4 binning, the histograms of the 1,000 means of the posteriors for Ω_m (top row) and σ_8 (bottom row), computed with Eq. (5.45). From left to right, the columns correspond to the three likelihoods; Poisson, Gaussian and GPC. The input values of the parameters are given by the vertical solid lines. The distributions of the estimated cosmological parameters are very similar for each likelihood and they are roughly normally distributed about the input cosmology.

In Fig. 5.7 we summarize the results for the bias on the mean estimated cosmological parameters for each of the three binnings and the three likelihoods. The black dots show the mean of the 1,000 estimated cosmological parameters. The black solid lines show the standard deviation of the mean. The red intervals show the error on the mean of the 1,000 estimated cosmological parameters (the standard deviation divided by $\sqrt{1,000}$). The black dotted line indicates the input cosmological parameters to the simulations. We see that there is a small consistent bias between the input value of the cosmological parameters and the mean of the estimated parameters (the dotted black line does not fall within the red error bars for any of the binnings or the likelihoods). There are many potential explanations for this bias. It suggests that the likelihood used and therefore the underlying halo model is on some level inaccurate. As the bias is consistent between the Poisson likelihood, which does not use the halo bias, and the other two likelihoods which do, the halo bias modeling is not the source of this effect. Another possible source of bias is the small parametrization difference between the Despali halo mass

[7]The codes to compute the covariance predictions, the likelihood, the forecasts and run the importance sampling are available at `https://github.com/payerne/LikelihoodsClusterAbundance`.

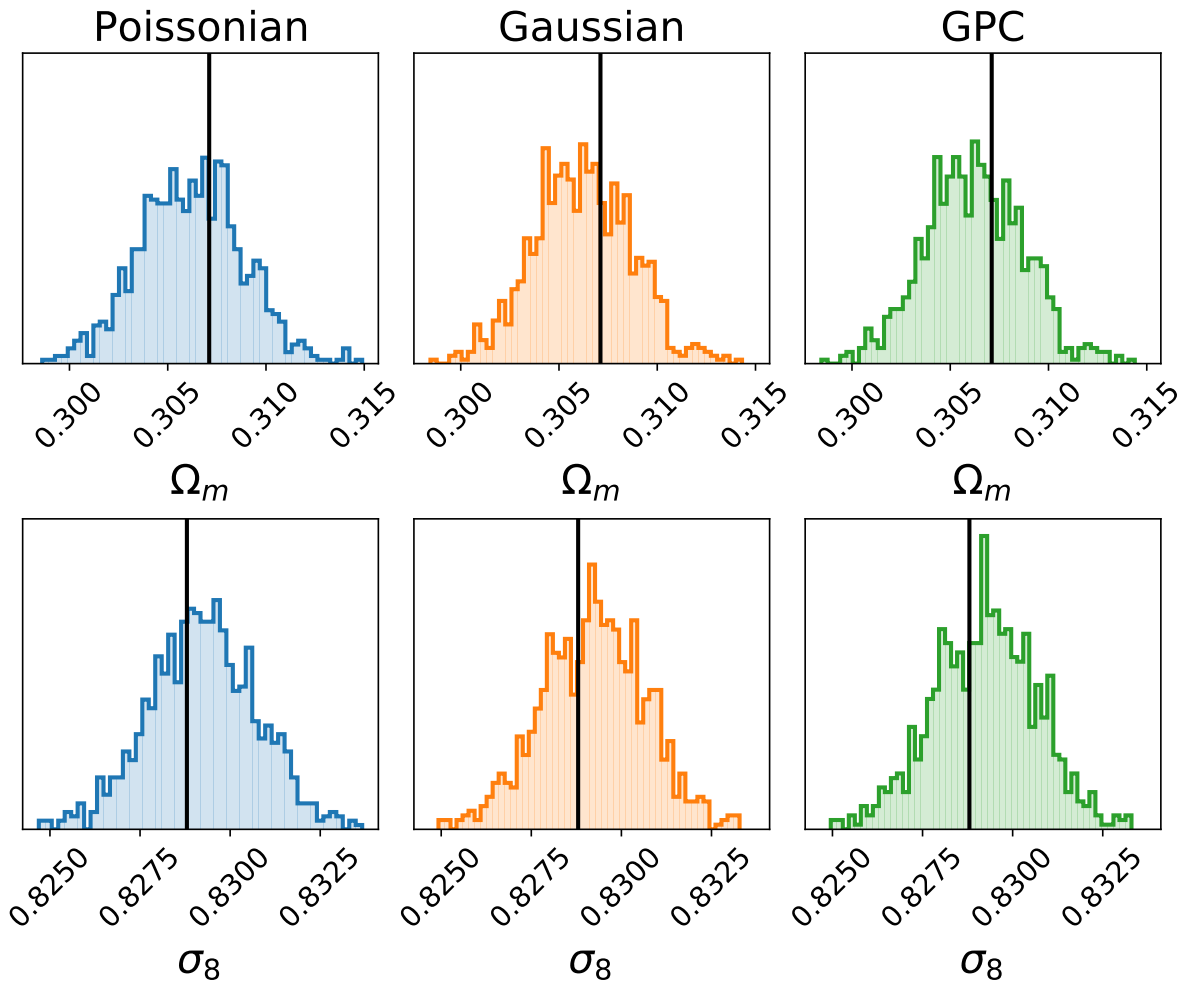

Figure 5.6: Histogram of the 1,000 posterior means of Ω_m (top) and σ_8 (bottom) for the 3 likelihoods (columns) for the first binning setup in Table 5.2. The vertical black lines correspond to the input cosmology of the PINOCCHIO simulations.

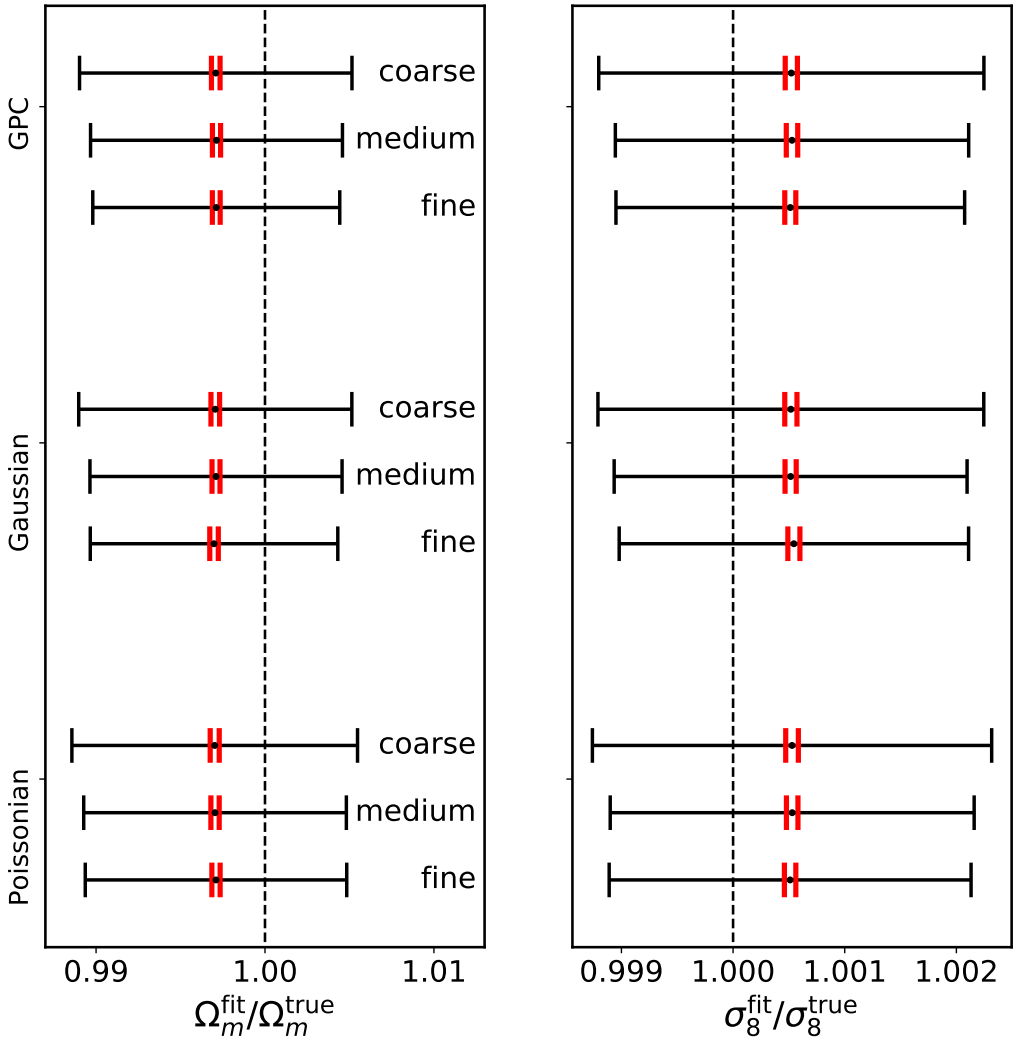

Figure 5.7: Average of individual posterior means for Ω_m and σ_8 (left and right respectively) with standard deviations of posterior means (black error bars) and errors on the mean (red error bars) for the three likelihoods and the three binning setups in Table 5.2.

function implementation in CCL and the one that was used to calibrate PINOCCHIO masses. The bias is nonetheless small, at the sub-percent level, and much smaller than the variance of an individual posterior, which we present in the following subsection.

5.1.6.2 Robustness of the posterior covariance

Now we consider the errors and correlations of the parameters obtained from individual simulations by comparing them to the dispersion of the mean values of the posteriors over the 1,000 simulations. As discussed in Section 5.1.2, this provides a relevant test to check the robustness of the error bars.

Figure 5.8 summarizes the error estimations for the three binning schemes (columns) and the two parameters (σ_8 and Ω_m on the top and bottom row respectively). In each panel, the likelihood used for the analysis is given on the x-axis. The blue dots with error bars correspond to the mean and dispersion of the individual errors that were computed using the importance sampling covariance matrix in Eq. (5.46). The red crosses are the values of the standard deviation of the ensemble of mean values, calculated from the ensemble covariance matrix presented in Eq. (5.49). The red crosses are the same errors as the black error bars in Fig. 5.7.

The errors slightly decrease when using more bins, regardless of the likelihood, with an improvement of 10%, between the `coarse` (left) and `medium` (middle) binning scheme. There might therefore be a small benefit to consider a finer binning when dealing with the upcoming survey data of Rubin LSST or *Euclid*. Pushing things further, there seems however to be no gain (or only a very marginal one) in using the `fine` binning compared to the `medium` one.

The Poisson likelihood underestimates the individual errors by up to 30%. This is evident as the blue dots are below the red crosses (the individual errors are smaller than the ensemble errors). This remains true even when a significant fraction of the mass-redshift space has low number counts for `medium` and `fine` binnings. This is because the Poisson likelihood cannot account for the SSC contribution (the errors are underestimated).

We see that the agreement is much better for the Gaussian and GPC descriptions, that account for super-sample variance. The individual errors are only slightly under or over-estimated ($\sim 2-5\%$ for the `coarse` and `medium` cases) compared to the dispersion of the 1,000 means. It seems likely that this slight underestimation arises from our use of only the block-diagonal terms in the Gaussian and GPC covariance matrices[8].

Fumagalli et al. (2021) also find that the parameter constraints for the average individual posteriors for the Poisson likelihood are tighter than that of the Gaussian likelihood. In this work, we were able to show explicitly that this is not because the Poisson uses information beyond the Gaussian likelihood but rather that the Poisson neglects the super-sample variance contribution.

We can also better understand the results using two different forecasts, motivated by the discussion presented in Section 5.1.2. The first is the widely used Fisher forecast, introduced in Eq. (5.25). This forecast should always agree with the estimated individual errors since the likelihood used in the analysis (which we call the *analysis* likelihood) is assumed to be equivalent to the latent likelihood driving the data. These forecasts are the black circles in Fig. 5.8. There is indeed an agreement between the blue dots and the black circles therefore these forecasts provide a consistency check on the analysis. We did not compute the Fisher forecast of the GPC individual parameter covariance, since it requires the computation of as many multidimensional integrals as the total number of mass-redshift bins in Eq. (5.24), making it very complex from a computational point of view. However, regarding the difference between Gaussian and GPC

[8]We ignore correlations between redshift bins as detailed in Section 5.1.1, to make this analysis computationally feasible for the GPC likelihood.

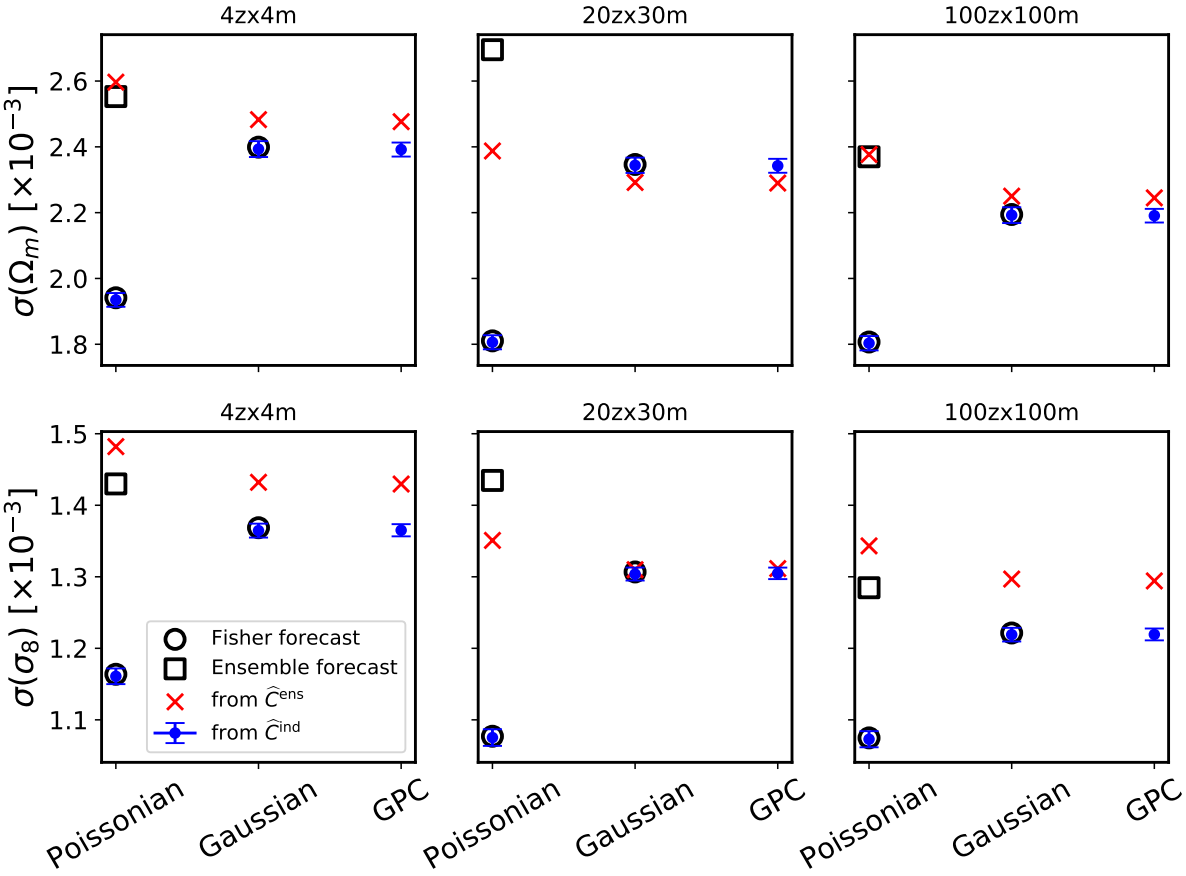

Figure 5.8: Errors on Ω_m (top) and σ_8 (bottom) for the different binning setups in Table 5.2 (columns). For each plot, the chosen likelihood is given on the x-axis while errors are on the y-axis. Blue dots correspond to the average of individual errors, and the error bars represent the standard deviation of these individual errors. Red crosses correspond to the standard deviation of the 1,000 best fits. Black circles correspond to Fisher forecasts in Eq. (5.25) (for the Poisson and Gaussian likelihoods) and black boxes to Ensemble forecast in Eq. (5.26) (only for the Poisson likelihood, see text for details).

individual errors for all binning schemes, we can assume that the GPC Fisher constraints are roughly the same as for the Gaussian case.

Secondly, we look at the ensemble forecast introduced in Eq. (5.26). This gives a prediction of the errors on the cosmological parameters where the analysis likelihood differs from the latent likelihood. As we do not know the latent likelihood we use the Gaussian likelihood as a substitute for the latent likelihood in this forecast, therefore stating explicitly that $\mathcal{C}^{\rm ens} = \mathcal{C}^{\rm ind}$ for the Gaussian analysis likelihood. Therefore we do not show the ensemble forecast for the Gaussian analysis likelihood, since it would overlap with the forecasted individual error (black circles). This forecast is shown for the Poisson analysis likelihood with the black boxes in Fig. 5.8. We find that for the Poisson analysis likelihood, these forecasts predict that the ensemble variance is higher than individual errors. Still, for the Poisson likelihood, both ensemble and Fisher forecast are respectively larger and smaller than the Gaussian individual errors, the latter representing the *correct* error if the Gaussian analysis likelihood was the latent one. This forecast gives reasonable agreement with the calculated ensemble errors, demonstrating that the underestimation of the Poisson errors indeed appears to be due to the lack of the super-sample variance in the Poisson analysis likelihood.

In addition to the errors on Ω_m and σ_8, we also compare the correlation coefficients using the three likelihoods and three binnings and present the results in Fig. 5.9. For each binning setup, we find a negative correlation coefficient between the two cosmological parameters, which is the expected degeneracy between Ω_m and σ_8 with cluster abundance cosmology. Additionally, we compute the correlation coefficients from the Fisher and the ensemble forecasts. We find that the correlation is accurate for the Poisson likelihood, the blue dots are close to the red crosses. The correlation is slightly underestimated for the GPC and Gaussian likelihoods (at most on the order of 5% in the case of `medium` binning). Once again, this is likely due to using only the block-diagonal terms in the covariance matrices for these likelihoods. This correlation would be important when calculating $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$, which is often used to quantify the tension between CMB and large-scale structure cosmological probes (see Valentino et al. (2021) for a recent review). However, we note a 5% bias on the correlation coefficient, as we see for the GPC and Gaussian cases, would lead to a bias of 1% on the error of S_8 and is therefore unimportant.

Finally, in the three binning regimes where the tests above were undertaken, the Gaussian and GPC likelihoods always yield the same parameter constraints. The GPC provides, in principle, a more accurate description of the statistical behavior of the data, but the Gaussian approximation is sufficient in the three use cases.

In conclusion, we find that the Gaussian likelihood remains an accurate description of the data, up to a very large number of bins, when the cluster analysis is performed with clusters above 10^{14} $M_\odot$ in a Rubin or *Euclid*-like survey. This is despite the bins being poorly populated. We would have expected that at a sufficiently fine binning the Gaussian likelihood would no longer be valid, which we do not see in our results. But even with the `fine` binning, there are still a significant number of bins containing ~ 100 halos (see right panel in Fig. 5.3). Choosing an even finer binning may allow us to reach a regime in which the Gaussian approximation is no longer valid and GPC a better description. However as we see little improvement between the cosmological constraints from 600 to 10,000 bins an even finer binning than that used appears to be of little cosmological interest.

5.1.6.3 Reduced volume sample and high mass only sample

As shown in Section 5.1.1, the Gaussian and Poisson likelihoods are two limit cases of the GPC description. In the previous section, where the number of clusters per bin is large, the GPC and Gaussian descriptions behave similarly and give robust constraints, conversely to the Poisson

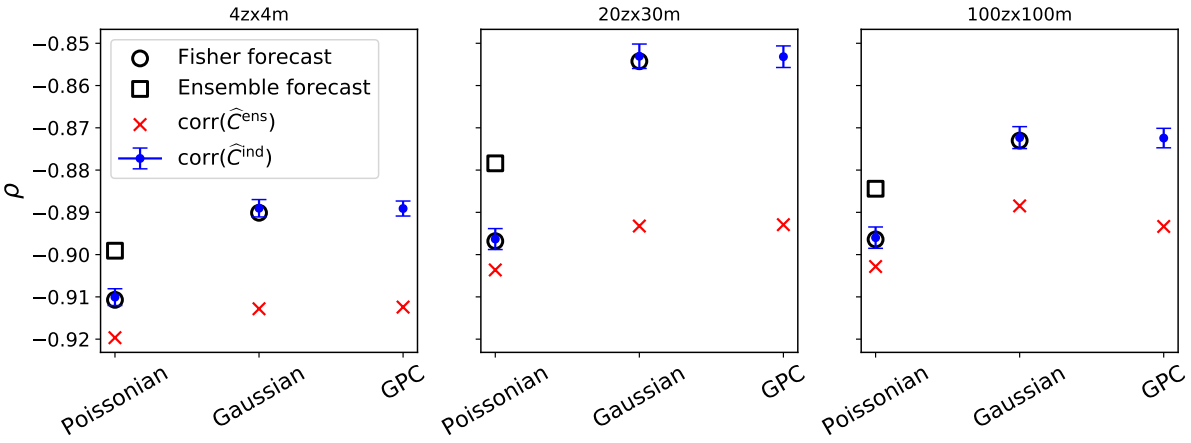

Figure 5.9: Correlation coefficients between Ω_m and σ_8 for the different binning setups in Table 5.2 (left to right). For each plot, the chosen likelihood is indicated on the x-axis while correlation is on the y-axis. Blue dots correspond to the average of individual correlations, and the error bars represent the standard deviation of these individual correlations. Red crosses correspond to the correlation of the 1,000 two-dimensional best fits. Black circles correspond to Fisher forecasts in Eq. (5.25) (for the Poisson and Gaussian likelihoods) and black boxes to the ensemble forecast in Eq. (5.26) (only for the Poisson likelihood, see text for details).

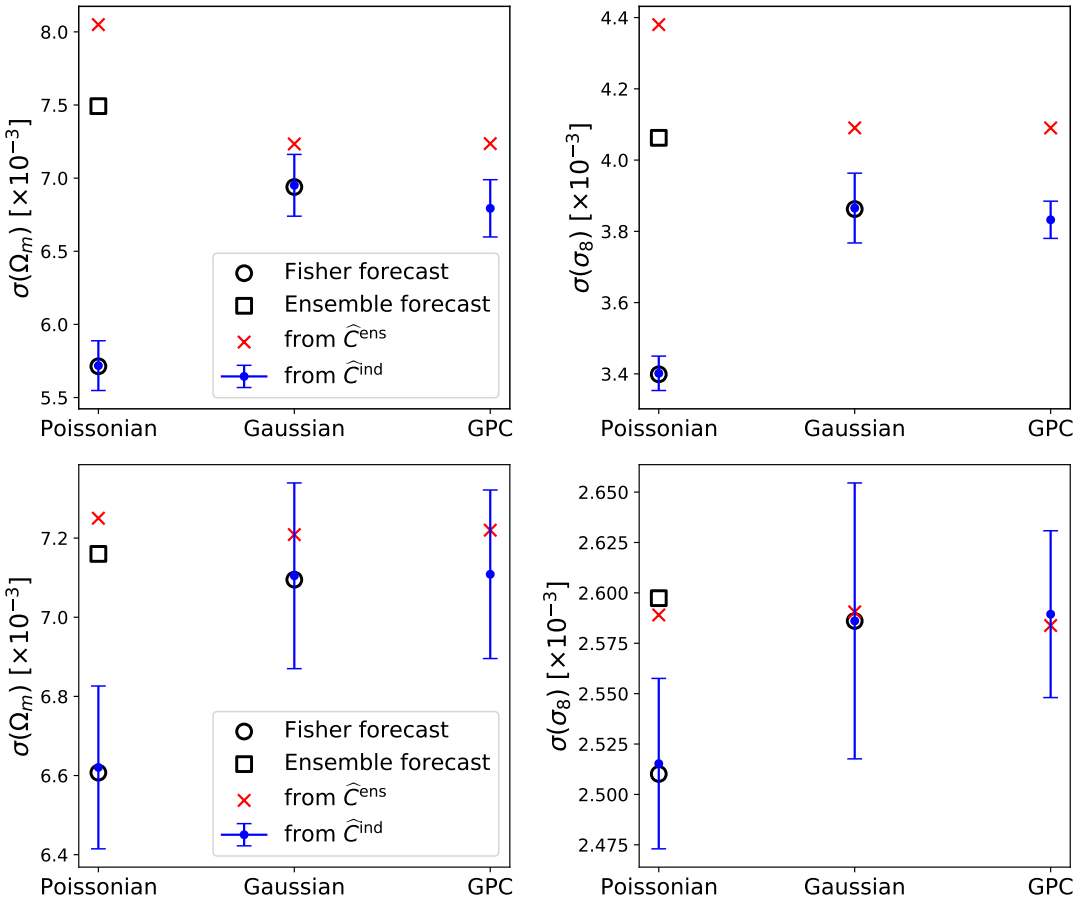

Figure 5.10: Same as Fig. 5.8 but for the reduced volume sample (top) and the high mass only cluster sample (bottom).

case. In this section, we perform the following exercise: we reduce the number of clusters per bin to see if we can reach a regime where GPC and Poisson would give similar results and if those results could be more robust than the Gaussian case. To do so, we first perform another analysis in which we reduce the sky fraction by a factor of 10, therefore reducing the total number of halo to $\sim 10^4$ per simulation.

We show the results of the three likelihoods by using this reduced volume sample and the `fine` mass-redshift binning in Fig. 5.10. By reducing the sky fraction (top panel), errors on Ω_m and σ_8 increased by a factor of $\sim 1/f_{\rm sky}$ compared to the constraints obtained using the full PINOCCHIO sky area. The Poisson likelihood still underestimates errors by $\sim 25-30\%$, and the Gaussian and GPC individual errors are closer to the ensemble errors, but negatively biased by a factor $\sim 5\%$.

In the same spirit, we now reduce the number of objects by selecting high mass clusters with $M > 5.10^{14} M_\odot$ (bottom panel), but keeping the full PINOCCHIO sky area. We repeat the analysis with 100 mass bins from $5.10^{14} M_\odot$ to $10^{15.6}$ $M_\odot$ (we use the same upper mass limit as defined in Section 5.1.5), and 100 redshift bins defined in Table 5.2. The total number of halos used in the analysis now drops to $\sim 5 \times 10^3$. We see that the Poisson likelihood still underestimates error on cosmological parameters from $5-20\%$ (depending on the cosmological parameter). With this result, we see that off-diagonal terms of the super-sample covariance and the super-sample variance are still important. The Gaussian and GPC individual errors are fully consistent with the ensemble errors. In this regime and under our model assumptions, Gaussian and GPC still correctly describe the latent likelihood.

5.1.7 Conclusions and perspectives

Cluster abundance cosmology relies on the likelihood to extract cosmological information from the data. It has been assumed in the past to follow a Poisson, Gaussian or Gauss-Poisson Compound distribution. In the context of the forthcoming cluster surveys, such as the Rubin LSST or the *Euclid* survey, it is important to consider if these likelihoods are sufficiently accurate and robust.

To do so, we make use of the PINOCCHIO simulation set, that provides us with 1,000 realizations covering 10,313 deg^2 (i.e., 1/4th of the full sky), for a given cosmology. For each likelihood and each simulation, we perform a cosmological analysis using importance sampling to recover posteriors for σ_8 and Ω_m. The comparison of the average and covariance of the 1,000 posterior means to that of the individual posteriors, provides us with the metric to assess the reliability of the likelihood under scrutiny. This is repeated for three binning schemes in the mass-redshift plane, ranging from 16 to 10,000 bins. Therefore in this work, we have sampled a total of 9,000 posteriors.

For each binning and each likelihood, we find a consistent bias in the recovered parameters despite this idealized setup. This bias is small with respect to the individual posterior variance and will therefore be negligible for future cosmological analyses.

For the three binning schemes we have explored, over the full survey area, the Poisson likelihood always underestimates the errors on the parameters by $\sim 30\%$; this is unsurprising as the super-sample variance cannot be accounted for in that description. Reducing the sky fraction to $\sim 1,030$ deg^2 or limiting the analysis to the most massive clusters does not change that conclusion despite yielding bins populated by less than five clusters. The Gaussian and Gauss-Poisson Compound likelihoods perform similarly in all cases, underestimating the errors on the recovered parameters by at most 5%.

Increasing the number of bins from the `coarse` (16 bins) to `fine` (10,000 bins) cases only yields a marginal $\sim 10\%$ improvement on σ_8 and Ω_m error bars. The advantage of a finer binning may therefore disappear in a real analysis when using a Gaussian likelihood. This may

become clearer in the future as we are also performing a study about the impact of binned versus unbinned cluster count likelihoods on cosmological constraints (complementary to this work), focusing on the calibration of mass-proxy relations. Unbinned cluster count likelihoods may also be useful when studying the correlations between different observable properties of galaxy clusters; such as the cluster richness, weak lensing, X-ray and SZ mass (for example see Rozo et al. (2009); Mahdavi et al. (2013); Umetsu (2020); Murray et al. (2022)).

This analysis is also an opportunity to revisit the question of forecasts. We compute the Fisher forecasts for the Poisson and Gaussian likelihoods, finding them (as expected) in agreement with the individual posterior covariances. We also develop a forecast approach that, conversely to the Fisher forecast, allows us to account for a latent likelihood that is not necessarily the same as the *analysis* likelihood. In our work this allowed us to show that the difference between the Poisson individual posterior errors and the ensemble errors can be explained by the non-inclusion of the super-sample covariance in the Poisson likelihood.

In more general terms, this new type of forecast, which in this work we call the ensemble forecast, would allow us to assess the reduction of the constraining power of an approximate likelihood in the case where the latent likelihood is known. This would be useful where the latent likelihood is computationally expensive to sample, for example, with many off-diagonal terms in the covariance matrix.

In this work, we were limited by the numerical feasibility of forecasting the GPC likelihood. We explored in Appendix E how to forecast the parameter errors when using a composite likelihood, i.e., the effective likelihood is a weighted average of likelihoods (e.g., GPC is a weighted average of Poisson likelihood with different means). The calculation has yet not resulted in a confident solution, we aim to pursue this work more in detail to compare directly the different forecasts of the three cluster likelihoods we tested in this work.

To conclude, from this work and for the general purpose of cluster abundance cosmology with future large-scale surveys, the Gaussian likelihood remains a sufficient description of the data and we find no clear benefits to make the analysis more complex and slower by using the GPC description or a large number of bins.

5.2 Including sample variance to unbinned cluster abundance likelihood

Everything described above dealt with the binned approach, i.e., we measure the number of clusters in pre-defined bins over the mass-redshift window (or proxy-redshift window). An alternative approach consists in considering an unbinned likelihood (Penna-Lima et al., 2014; Penna-Lima, 2010) that uses the cluster properties in the Bayesian inference at the individual level, such as cluster redshift, individual mass proxy and individual mass (when they have been previously inferred, for instance using weak lensing). To do so, the unbinned regime is obtained by considering the binning of the mass-redshift plane small enough to have at most one cluster per two-dimensional bin (so the name "unbinned" does not perfectly describe this methodology, since bins are still there, but very small). So the infinitesimal $k-$th mass-redshift bin has a volume $\Delta z \Delta m$ (for this paragraph we put together the two indexes ij in a single index $k \in [1, c]$), and the observed count $\widehat{N}_k$ in that bin only has two possible values $\{0, 1\}$, with corresponding probabilities

$$P(\widehat{N}_k|N_k) = \begin{cases} p_0 & \text{if } \widehat{N}_k = 0, \\ 1 - p_0 & \text{if } \widehat{N}_k = 1, \end{cases} \tag{5.50}$$

where N_k is the predicted small abundance in the $\Delta z \Delta m$ volume surrounding m_k and z_k and is given by

$$N_k = \int_{\Delta z} \int_{\Delta m} \Omega_S \frac{dn(m,z)}{dM} \frac{d^2V(z)}{dz d\Omega} dz dm \approx n_h(m_k, z_k) \Delta z \Delta m, \tag{5.51}$$

where $\Delta m, \Delta z \ll 1$. Here, $n_h(m,z)$ is the predicted halo number density per mass and redshift range (given by the product of the halo mass function and the partial volume).

5.2.1 The standard unbinned approach

The *standard* derivation of the unbinned likelihood is based on the Poisson statistics, i.e., $p_0 = e^{-N_k}$ and $p_1 = N_k e^{-N_k}$, such that $p_0 + p_1 = e^{-N_k}(1 + N_k) \approx 1$ (when $N_k \ll 1$), and correlations between counts are null. Then, given a survey of $\widehat{N}_{\rm tot}$ clusters with masses m_k and redshifts z_k, the standard unbinned likelihood is given as a product of individual likelihoods for each mass-redshift bin with index k, whether they are filled or not by at most one halo. So we can derive the standard *Unbinned likelihood for Clusters* (ULC[9])

$$\mathcal{L}_{\rm ULC} = \prod_{k=1}^{\text{all bins}} \mathcal{P}(\widehat{N}_k | N_k) = \prod_{k'=1}^{\text{empty bins}} \mathcal{P}(0|N_{k'}) \times \prod_{k=1}^{\text{full bins}} \mathcal{P}(1|N_k) = e^{-N_{\rm th}} \prod_{k=1}^{\widehat{N}_{\rm tot}} N_k, \tag{5.52}$$

where $N_{\rm th} = \langle \widehat{N}_{\rm tot} \rangle$ is the predicted total cluster abundance at given cosmology.

The unbinned analysis accounts for each individual position on the mass-redshift plan and it is sensitive to the local cluster mass-redshift distribution contrary to the binned approach. Since unbinned approaches are more sensitive to local variations of the halo distribution, they have the advantage to provide tighter constraints on cosmology compared to the binned cases. We show in Fig. 5.11 the joint posteriors of Ω_m, σ_8 and the matter power spectrum index n_s. As for Ω_m and σ_8, the n_s parameter impacts the growth of structures in the Universe by modifying the shape of the primordial power spectrum, and in this work, we test how it can be constrained from abundance, and how it correlates with the *standard* growth parameters (that were previously considered). We use two different binned likelihoods that we mentioned in Section 5.1 and the standard unbinned likelihood (we used one of the PINOCCHIO mock presented in Section 5.1.3, and the details of the cluster sample "full" are shown in Table 5.3). For this example, the binned likelihoods (Gaussian: blue, Poisson: red) are computed for 3 mass bins and 3 redshift bins over the full mass and redshift ranges. We recover the results of the previous section, where the binned Poisson likelihood gives lower/underestimated errors compared to the binned Gaussian, which we showed to be the most accurate description of the data in the binned case. Furthermore, the constraints from the standard unbinned likelihood (dashed black lines) are tighter by a factor of 20 to 40% compared to the binned Gaussian approach and reduced by a factor 10 to 30% compared to the binned Poisson results.

Unbinned likelihoods have been used in several cosmological analyses, namely SPT (Bocquet et al. (2019), that used 343 clusters), XMM-XXL (Pacaud et al. (2018), 178 clusters), WtG (Mantz et al. (2015), 50 clusters). Artis et al. (2022); Chiu et al. (2023); Bocquet et al. (2019, 2023) showed that the unbinned framework is perfectly adapted to constrain the uncertainty and correlation between the different cluster observables (e.g., between lensing masses and richness estimates, but also between optical and SZ/X-ray observables). Moreover, the standard unbinned method does not account for SSC. Some steps in that direction have been explored in Pacaud et al. (2018), where the full likelihood is given by

$$\mathcal{L}_{\rm Pacaud} = e^{-N_{\rm th}} \prod_{k=1}^{\widehat{N}_{\rm tot}} N_k \times \mathcal{L}_{\rm GPC}(\widehat{N}_{\rm tot}), \tag{5.53}$$

[9]This terminology is taken from Penna-Lima (2010).

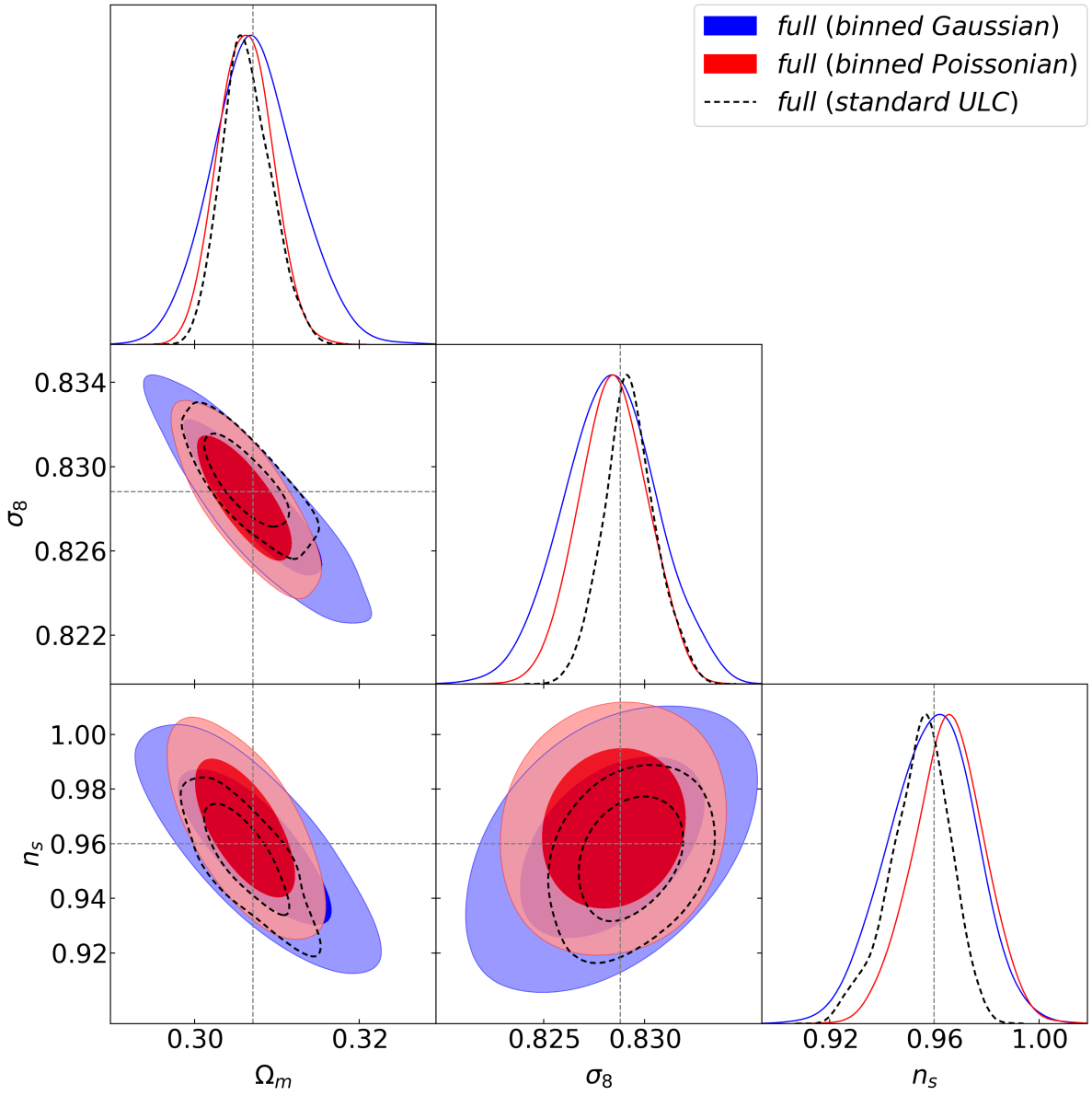

Figure 5.11: Joint posteriors on three cosmological parameters using one PINOCCHIO mock ("full" sample presented in Table 5.3). Blue and red contours correspond respectively to the binned Gaussian and Poisson posteriors (where we used 3 mass and 3 redshift bins). Black dashed contours are obtained with the standard unbinned likelihood. PINOCCHIO cosmology is represented by vertical/horizontal dashed lines.

namely, the product between the standard unbinned approach (left side of the above equation) and a Gauss-Poisson Compound statistics applied to the total cluster count $\widehat{N}_{\rm tot}$ (so one single mass-redshift bin).

5.2.2 Including SSC in the unbinned framework: formalism

The Pacaud et al. (2018) approach only accounts for SSC when applied to the total (binned) number of clusters and not for each one of them. In this section, we aim to account for this effect at the individual cluster level. Moreover, Takada & Spergel (2014) developed a formalism to approximate the GPC likelihood in the low abundance regime that is a first step toward the unbinned regime, and in this section, we propose a more extended rigorous derivation of the unbinned likelihood accounting for the effect of super sample covariance.

Let us consider a series of $c = N_z \times N_m$ mass-redshift bins with respective mass indexes $i \in [1, c_m]$ and redshift indexes $j \in [1, c_z]$. For a fixed overdensity realization $\{\delta_{ij}\}$ in each ij-th mass-redshift bin, the binned Gauss-Poisson Compound likelihood is given by

$$\mathcal{L}(\{\widehat{N}_{ij}\}|\{x_{ij}\}) = \prod_{i,j=1}^{c} \mathcal{P}(\widehat{N}_{ij}|x_{ij}), \tag{5.54}$$

where $x_{ij} = N_{ij}(1 + b_{ij}\delta_j)$ with N_{ij} and b_{ij} the abundance prediction and the average halo bias in the ij-th mass-redshift bin respectively (the quantity δ_j is a *spatial* overdensity, thus only depending on the redshift index). Here, the $\mathcal{P}$ functions are individual Poisson likelihoods. Then, the *full* GPC likelihood is obtained when averaging to above expression over the many realizations of the density perturbations, such as $\mathcal{L}(\{\widehat{N}_{ij}\}) = \langle \mathcal{L}(\{\widehat{N}_{ij}\}|\{x_{ij}\})\rangle$. We first approximate the *local* GPC likelihood for the low abundance regime, denoted by two conditions $b_{ij}\delta_j \ll 1$ and $b_{ij}N_{ij}\delta_j \ll 1$ (Takada & Spergel, 2014). The first condition denotes that the smoothed density perturbation in each redshift bin is small [10] and the second that each bin has a small number of clusters. By removing term $o(\delta_k^3)$ (that are considered to be small in the low abundance regime) and taking the ensemble average of $\mathcal{L}(\{\widehat{N}_{ij}\}|\{x_{ij}\})$ to recover the full GPC likelihood (see the full details of the calculation in Appendix F), we get that

$$\mathcal{L}(\{\widehat{N}_{ij}\}) = \prod_{k=1}^{c} \mathcal{P}(\widehat{N}_{ij}|N_{ij})(1 + f_{\rm SSC}). \tag{5.55}$$

The first term in Eq. (5.55) describes the previous Poisson statistics of each $\widehat{N}_{ij}$ with Poisson mean N_{ij}, i.e., the *standard* Poisson likelihood for binned cluster abundance. Now, the second term $f_{\rm SSC}$ denotes the super-sample covariance contribution (SSC) to the total likelihood. So the full GPC likelihood is obtained by correcting the standard unbinned likelihood (that only accounts for Poisson statistics) by a multiplying factor $(1 + f_{\rm SSC})$. After a lot of algebra (see Appendix F), we get

$$f_{\rm SSC} = \frac{1}{2}([1] - 2 \times [2] + [3] - [4]), \tag{5.56}$$

[10] Takada & Spergel (2014) have considered that the smoothed overdensity $\delta_j \ll 1$ in large redshift bins. In our case, we consider the redshift bins to be small, then $\delta_j \ll 1$ may no longer be valid. However, we have checked analytically that the unbinned formalism we develop hereafter is the same without this condition. We have discussed this aspect in Appendix F.

where

$$[1]: \sum_{m,n=1}^{c_z} S_{mn} \sum_{i,j=1}^{c_m} b_{im} b_{jn} N_{im} N_{jn} = N_{\rm th}^2 \langle Sbb \rangle_{\rm th}, \tag{5.57}$$

$$[2]: \sum_{m,n=1}^{c_z} S_{mn} \sum_{i,j=1}^{c_m} b_{im} b_{jn} N_{im} \widehat{N}_{jn} = N_{\rm th} \widehat{N}_{\rm obs} \langle Sbb \rangle_{\rm th/obs}, \tag{5.58}$$

$$[3]: \sum_{m,n=1}^{c_z} S_{mn} \sum_{i,j=1}^{c_m} b_{im} b_{jn} \widehat{N}_{im} \widehat{N}_{jn} = \widehat{N}_{\rm obs}^2 \langle Sbb \rangle_{\rm obs}, \tag{5.59}$$

$$[4]: \sum_{m=1}^{c_z} S_{mm} \sum_{i=1}^{c_m} b_{im} b_{im} \widehat{N}_{im} = \widehat{N}_{\rm obs} \langle Sb^2 \rangle_{\rm obs}. \tag{5.60}$$

Here, $\langle \cdot \rangle_{\rm th}$ denotes a theoretical binned average over the 4 dimensions (2 mass dimensions, 2 redshift dimensions), accounting for the predicted abundances, halo bias and $S_{mn} = \langle \delta_m \delta_n \rangle$ is the covariance between the over-densities within the redshift bins m and n. The average $\langle \cdot \rangle_{\rm obs}$ in [3] is the 4-dimension binned average over the observed counts $\widehat{N}_{jn}$ (in [4], the average only runs over 1 mass dimension and 1 redshift dimension). Finally, $\langle \cdot \rangle_{\rm th/obs}$ denotes the binned average over 1+1 dimensions with the predicted N_{jn} and 1+1 dimensions with the observed $\widehat{N}_{jn}$. At this point, we find an expression that generalizes the result found by Takada & Spergel (2014), that derived the GPC likelihood by considering halos at a single redshift $z_{jn} = z_0$ (such as all clusters share the same amplitude of matter fluctuation $S_{kl} = S_0$). In that sense, our derivation generalizes this result to a population of clusters with different redshifts. We checked that applying the condition $S_{kl} = S_0 \delta_{kl}^K$ gives back the Takada and Spergel's results (again, see Appendix F for more details about the calculation).

The last step to obtain the unbinned likelihood is *pushing* toward the unbinned regime. It means that the mass-redshift bins are small enough to contain at most one cluster. The predicted abundance N_{ik} in the narrow mass-redshift bin ik is asymptotically equivalent to the formula given in Eq. (5.51), namely $N_{ik} = n_h(m_i, z_k) \Delta z \Delta m$ for $\Delta m, \Delta z \ll 1$. The terms $\langle Sbb \rangle$ and $\langle Sb^2 \rangle$ are computed considering $S_{kl} \to \sigma^2(z_k, z_l)$ for narrow redshift bins. The unbinned regime permits to turn sums over the predicted abundance N_{ik} into integrals,

$$[1] = \int dz_1 dm_1 n_h(m_1, z_1) b(m_1, z_1) \int dz_2 dm_2 \; n_h(m_2, z_2) b(m_2, z_2) \sigma^2(z_1, z_2), \tag{5.61}$$

$$[2] = \sum_{k=1}^{\widehat{N}_{\rm tot}} b(m_k, z_k) \int dz_1 dm_1 n_h(m_1, z_1) b(m_1, z_1) \sigma^2(z_1, z_k) \tag{5.62}$$

$$[3] = \sum_{k,l=1}^{\widehat{N}_{\rm tot}} b(m_k, z_k) b(m_l, z_l) \sigma^2(z_k, z_l) \tag{5.63}$$

$$[4] = \sum_{k=1}^{\widehat{N}_{\rm tot}} b(m_k, z_k)^2 \sigma^2(z_k, z_k). \tag{5.64}$$

Finally, we find that the (log) Unbinned Likelihood for Clusters with Super-Sample Covariance (ULC-SSC) is given by

$$\ln \mathcal{L}_{\rm ULC}^{\rm SSC} = \ln \mathcal{L}_{\rm ULC} + \ln(1 + f_{\rm SSC}), \tag{5.65}$$

where the full expression of $f_{\rm SSC}$ is given by

$$f_{\rm SSC} = \frac{1}{2} [N_{\rm th}^2 \langle Sbb \rangle_{\rm th} - 2 N_{\rm th} \widehat{N}_{\rm obs} \langle Sbb \rangle_{\rm th/obs} + \widehat{N}_{\rm obs}^2 \langle Sbb \rangle_{\rm obs} - \widehat{N}_{\rm obs} \langle Sb^2 \rangle_{\rm obs}]. \tag{5.66}$$

The term $\mathcal{L}_{\rm ULC}$ is the standard Poisson likelihood in Eq. (5.52), seen in the previous Section 5.2.1. Depending on the sensitivity of the correcting factor $(1 + f_{\rm SSC})$ to a change in cosmological parameter, it may behave differently from the standard likelihood and would impact the Bayesian inference of cosmological parameters. This is what we check hereafter.

5.2.3 Impact on cosmological parameters

We have seen in Section 5.1 that accounting for super-sample covariance in a binned approach increases the parameter errors by $\sim 10\%$ on Ω_m and σ_8, compared to only accounting for the standard Poisson shot noise. Formally, the binned SSC contributes to increase the scatter of measured counts (then increases the errors on the recovered cosmological parameters). However, the impact of the unbinned SSC term is not as intuitive.

According to the expression of the $f_{\rm SSC}$ term in Eq. (5.66), we find that it has many dependencies with respect to the characteristics of the cluster sample. Namely $N_{\rm th}^2\langle Sbb\rangle_{\rm th}$, $N_{\rm th}\widehat{N}_{\rm obs}\langle Sbb\rangle_{\rm th/obs}$ and $\widehat{N}_{\rm obs}^2\langle Sbb\rangle_{\rm obs}$ are proportional to the sky area of the survey Ω_S and to the square of the total cluster count $N_{\rm tot}^2$, and the term $\widehat{N}_{\rm obs}\langle Sb^2\rangle_{\rm obs}$ is proportional to $N_{\rm tot}$. We can intuitively expect that the importance of the SSC contribution at the level of cosmological parameter inference would depend on these quantities, namely the sky area of the survey and/or on the total number of clusters used in the analysis. Before the Bayesian inference, we first study the behavior of the $1 + f_{\rm SSC}$ term with respect to a change in cosmological parameters. To propagate the change in the SSC term to the posterior constraints on cosmological parameters, we use that

$$\frac{\mathcal{P}_{\rm ULC}^{\rm SSC}}{\mathcal{P}_{\rm ULC}} = \frac{\widetilde{\mathcal{L}}_{\rm ULC}}{\widetilde{\mathcal{L}}_{\rm ULC}^{\rm SSC}}(1 + f_{\rm SSC}), \tag{5.67}$$

where $\mathcal{P}_{\rm ULC}^{\rm SSC}$ is the posterior distribution of cosmological parameter using the unbinned SSC likelihood, $\mathcal{P}_{\rm ULC}$ corresponds to the *standard* unbinned posterior, and the terms $\widetilde{\mathcal{L}}$ correspond to the *evidence* of a likelihood defined by

$$\widetilde{\mathcal{L}}(\{\widehat{N}_i\}) = \int d^d\theta \mathcal{L}(\{\widehat{N}_i\}|\vec{\theta})\pi(\vec{\theta}). \tag{5.68}$$

We use one PINOCCHIO catalog (the one that was used in the Bayesian inference in Fig. 5.11). The computation of the standard unbinned likelihood at fixed cosmology is based on the interpolation of the Despali et al. (2015) halo mass and redshift distribution, that is previously tabulated on a two-dimensional grid between the minimum and maximum of each mass and redshift axes (we use the same approach as for the Section 5.1). For each dark matter halo in the catalog, the interpolated function is evaluated at each halo mass and redshift position in the sample. Again, for the halo bias, we use the Tinker et al. (2010) parametrization as we did in the last Section 5.1 and we use the `PySSC` code to compute the amplitude of the matter fluctuations[11].

The Fig. 5.12 shows the correcting factor $1 + f_{\rm SSC}$ as a function of Ω_m, for different cluster samples that have the common redshift range $0.2 - 1$, but different mass ranges. These mass ranges have different colors, from blue to purple, the smallest cluster sample (only low mass) is

[11] We use the *Planck* cosmology to compute $\sigma(z1, z_2)^2$ once, and we do not vary this quantity in the Bayesian inference pipeline as we did for the approaches developed in Section 5.1. So the full SSC contribution varies with respect to the cosmology through the halo mass function and partial volume that are re-computed at each location in the parameter space. This implies a higher sensitivity to a change in cosmology for the SSC effect compared to the covariance of the binned Gaussian likelihood (see Section 5.1), where both the "abundance" ingredients (halo mass function, partial volume and halo bias) and matter amplitude fluctuation within $\sigma(z1, z_2)^2$ where computed once at the *Planck* cosmology.

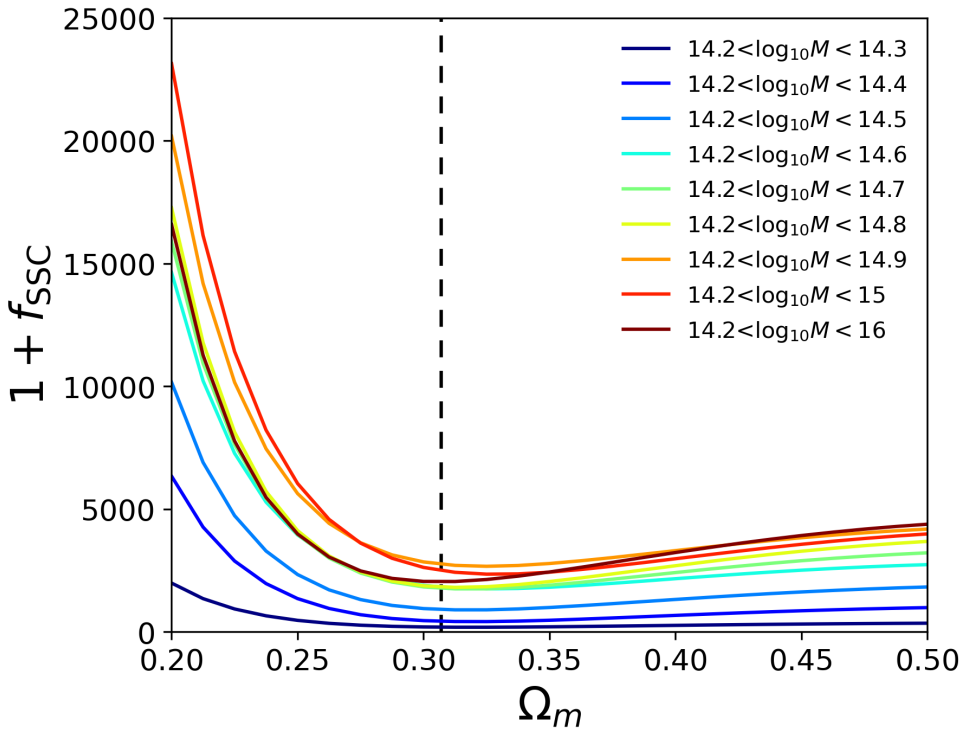

Figure 5.12: Super-sample covariance contribution to the unbinned likelihood for cluster count $1 + f_{\rm SSC}$ as a function of Ω_m. The different colors correspond to the different truncated cluster samples. Blue corresponds to low-mass clusters, we progressively include more massive clusters from blue to purple.

in blue, and the full PINOCCHIO sample is in purple. All other cosmological parameters are fixed to *Planck* values. We see that the SSC contribution on the Ω_m space is not flat (so when re-scaled by the ratio of evidences, the ratio between the SSC unbinned and standard unbinned posteriors is not equal to 1). Moreover, we find that it shows a distorted parabola whose minimum corresponds roughly to the input Ω_m parameter. At the end of Appendix F, we show that the width of the parabola roughly scales with $1/\langle Sbb\rangle N_{\rm th}$, so the spatial variation of the SSC term is even stronger when the total number of clusters increases. However the effect on cosmological constraints is not distinguishable from this plot: each curve needs to be multiplied by the ratio of evidences, and by the standard posterior distribution. Nonetheless, we see that the SSC contribution should impact the recovered error of cosmological parameters, due to its specific convex shape. The Fig. 5.12 helps understand more *qualitatively* rather than quantitatively the impact of the SSC contribution on parameter posteriors (here on the Ω_m Bayesian inference). The SSC contribution seems stronger for large cluster samples since the spread of the SSC parabola decreases with the total count of clusters. However, as $N_{\rm tot}$ increases, the standard unbinned posterior (denoted by $\mathcal{P}_{\rm ULC}$) is more "peaked" around the input cosmology, so its very local variations may mitigate the effect of SSC contribution. More precisely, this last one may be roughly constant on the range of strong variation of $\mathcal{P}_{\rm ULC}$.

To test the impact of SSC at the parameter inference level, we first considered performing a Fisher forecast. As seen in Section 5.1, a Fisher forecast gives *with little effort* the intrinsic level of uncertainty on cosmological parameters for a given likelihood. It is quite easy to numerically

sample	redshift range	($\log_{10}$)mass range	Number of clusters
full	0.2 - 1	14.4 - 15.6	∼83,300
low−redshift	0.2 - 0.4	14.4 - 15.6	∼16,500
low−mass	0.2 - 1	14.2 - 14.4	∼52,700
high−mass	0.2 - 1	14.5 - 15.6	∼17,600

Table 5.3: Details of the mass-redshift samples used to test the impact of the SSC contribution in an unbinned analysis compared to the standard unbinned approach.

derive the Fisher matrix for the standard unbinned likelihood since it only involves a two-dimensional integral of the derivatives of the halo number density (see Appendix G for more details). However, for the unbinned SSC likelihood, the Fisher forecast involves multivariate integrals with as many dimensions as the number of clusters in the full sample (from a few hundred to tens of thousands). It is too computationally expansive and we opted not to consider this option in this work.

Instead, we decided to estimate directly the posterior on cosmological parameters using the PINOCCHIO idealistic data set. We explore varying the number of clusters used for the analysis, that will impact the precision of the recovered parameters, by using different mass and redshift ranges.

For that, we identify different cluster samples that differ from one another by the selected mass and redshift regime covered by the halo number density. The "full" sample of galaxy clusters is comprised between redshift $0.2 - 1$, and between ($\log_{10}$)mass 14.2 - 15.6. We define the "low-redshift" and "low-mass" samples respectively by restricting the "full" sample to the redshift interval 0.2 - 0.5 and the ($\log_{10}$)mass interval 14.2 - 14.4. Finally, we consider only the more massive halos with ($\log_{10}$)masses between 14.5 to 15.6 in the "high−mass" cluster sample (that has approximately the same number of clusters as the "low-redshift" sample). The names, mass and redshift ranges for each sample are summarized in Table 5.3.

For each sample, we estimate the joint posterior distribution of the growth parameters in ΛCDM namely Ω_m, σ_8 and n_s. Contrary to the previous section where we needed to estimate thousands of posteriors, we now use the `emcee` package (Foreman-Mackey et al., 2013) to run MCMCs and estimate the posterior distribution of parameters from one PINOCCHIO simulation.

We represent the posterior distributions for the four different samples in Fig. 5.13. For the different cluster samples and the two different likelihoods, we see that the parameter errors are different from one another, and are all compatible at $\sim 1 - 2\ \sigma$ with the input cosmology, due to the difference in size of the several cluster samples. We find that the mass-redshift region probed by the cluster sample has a strong impact on the correlation of parameters, since the correlation between n_s and σ_8 are positive and negative respectively for the "low-redshift" and "low-mass" sample. The "full" sample gives the smallest error bars. Fig. 5.13 is a busy plot and we extract the main features in Fig. 5.14.

The Fig. 5.14 (left) shows the best fits and errors obtained with the two different likelihoods (unbinned standard: squares, unbinned SSC: circles) for the different cluster samples in Table 5.3. We see that all samples are compatible at $1-2\sigma$ with the input value, with and without accounting for SSC, for the three different parameters. We see that the error bars reflect the size of the sample (a larger cluster sample gives smaller error bars on the recovered parameters). We also see very little difference in the recovered values between the standard ULC and the SSC unbinned likelihood, whatever the sample.

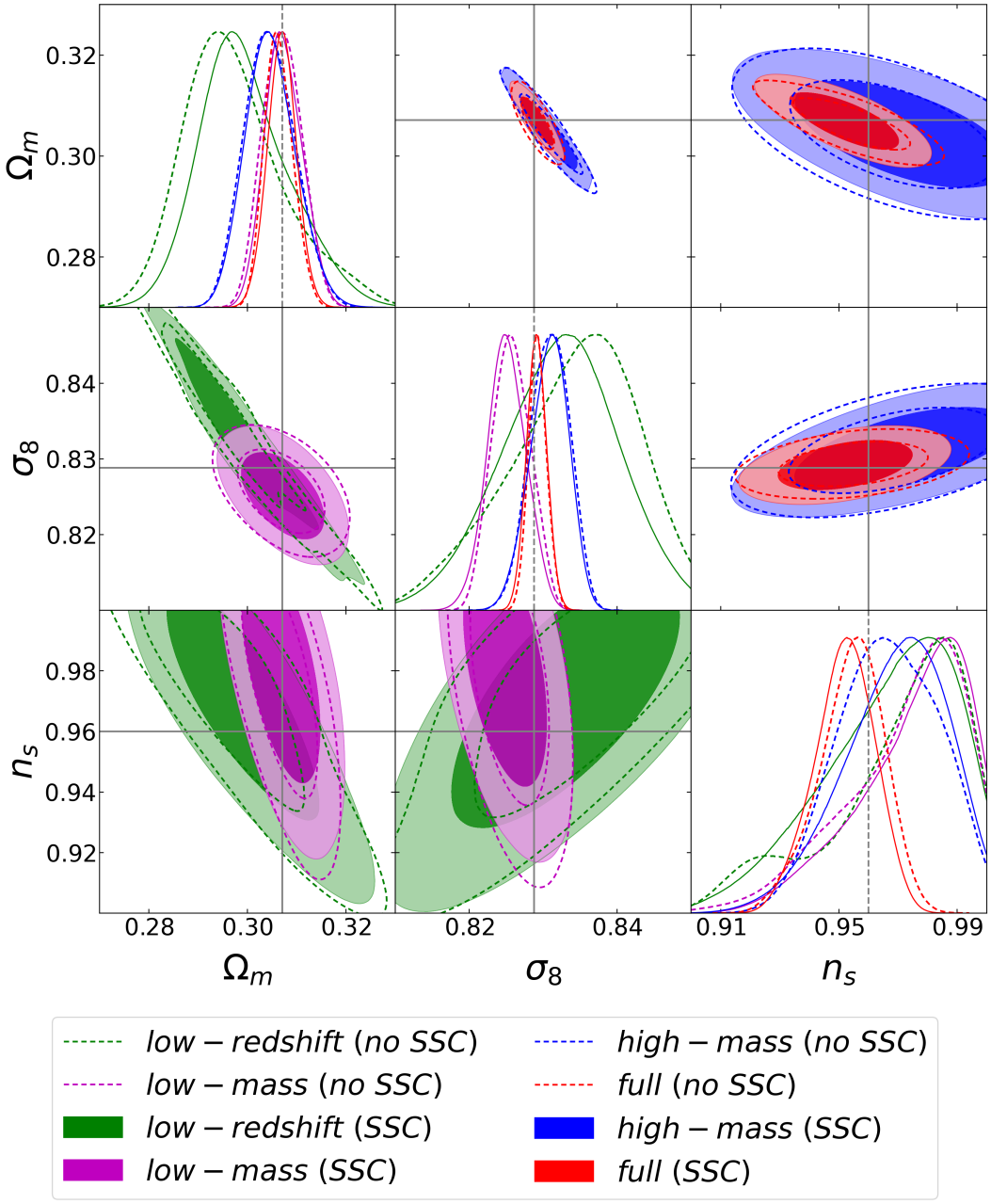

Figure 5.13: Joint posterior distribution of Ω_m, σ_8 and n_s considering the unbinned standard and unbinned SSC likelihoods. Upper right corner plot: Contour plots for the "high−mass (blue) and "full" (red) samples. Lower left corner plot: Contour plots for the "low−redshift" (green) and "low−mass" (magenta) samples. The diagonal shows the marginalized posteriors for all the configurations.

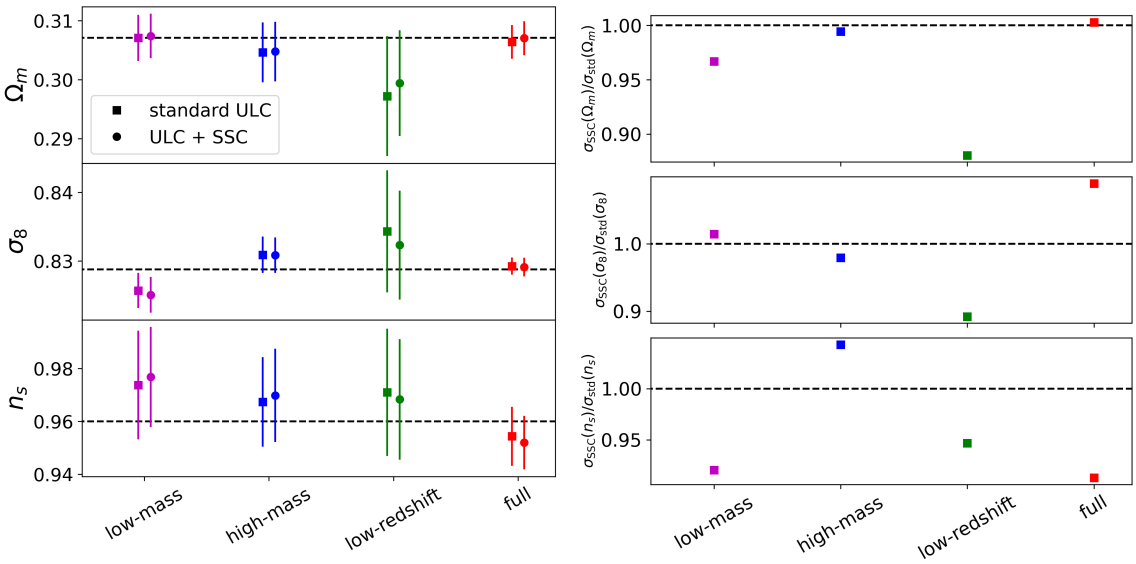

Figure 5.14: Left: Best-fitted Ω_m, σ_8 and n_s (and error bars) for the four different cluster samples (we used the same colors as in Fig. 5.13). Constraints obtained with the unbinned SSC likelihood are represented with the dotted points, and with the standard unbinned likelihood with squared points. Right: Ratio between the errors on cosmological parameters obtained with the unbinned SSC likelihood, and the errors obtained with the standard unbinned likelihood.

The Fig. 5.14 (right) shows the ratio between the unbinned SSC and the unbinned standard likelihood. The SSC contribution affects the cosmological parameter errors at most 10% for all of these cosmological inference setups. However, we do not identify a tendency of the SSC term to impact the correlation between cosmological parameters compared to the standard approach in a particular direction (using more massive or more recent clusters). We see that for some parameters and some cluster samples, the SSC error bars are slightly smaller, which seems counter-intuitive when compared to the effect of SSC on binned statistics, that tends to increase the dispersion of parameter posteriors (for the binned approach, the use of a binned Poisson likelihood leads to an underestimation of parameter errors compared to the Gaussian likelihood, that account for SSC).

We studied other configurations that are detailed in Appendix H by varying separately the mass range and the redshift range of the cluster sample. For both study cases, we find the same conclusion, such that errors are affected at most $10 - 15\%$ compared to the standard unbinned likelihood, without distinguishing a specific trend in the recovered parameter or their errors with respect to the inference setup we defined. The results above have been obtained on a single PINOCCHIO mock catalog. A direct extension to this work would be to check if the (lack of) trends we see here are present for other realizations.

We are currently consolidating and writing up these findings for publication (Payerne et al. in prep). Follow-up studies we are considering include several questions regarding the impact of SSC in the unbinned framework: How do the results we obtained compare with the unbinned/binned approach proposed by Pacaud et al. (2018)? How does accounting for more systematic effects (photometric cluster redshift, proxy-mass relation) mitigate or emphasize the effect of SSC on parameter posteriors in an unbinned analysis? Is the difference bigger between the standard unbinned and unbinned SSC when considering more/less than three cosmological

parameters to fit? We aim to investigate more in detail all of these aspects in future works. Moreover, another perspective is to test if the unbinned SSC likelihood is *robust* in the sense defined in Section 5.1, i.e., comparing the ensemble error to individual errors for both standard unbinned and SSC unbinned likelihoods.

5.3 Conclusions

Likelihoods are central in Bayesian inference and are crucial for accurate and robust constraints on cosmological parameters. In this chapter we addressed fundamental questions about cluster abundance likelihoods, that will be very relevant in the context of large surveys for precision cosmology. We built a framework to test the robustness and accuracy of likelihoods generally used for cluster abundance cosmology. We showed that the binned Poisson likelihood always underestimates parameter variances compared to the Gaussian and Gauss-Poisson Compound likelihoods, that both account for super-sample covariance. We also find that the Gaussian and GPC likelihoods provide robust constraints, for a variety of cosmological inference setups using a *Euclid* or LSST-like surveys.

We also explored how to account for super-sample covariance (SSC) in the formalism of unbinned abundance. The unbinned methods generally provide tighter constraints on cosmological parameters compared to binned approaches. For most of the cosmological inference setups we tested, we have found at most 10% difference between the standard unbinned and unbinned SSC constraints. However, the computational cost to evaluate the unbinned SSC likelihood is much higher than the standard approach (high dimensional integrals). Regarding other complications appearing in abundance cosmology, such as the control of systematic, the calibration uncertainties of the cluster scaling relation, the issues related to the likelihood function in the unbinned framework, although potentially important, are not the most time-critical given the imminent arrival of stage IV survey data and can be addressed at a later time.

Chapter 6

Weak lensing shear multipole analysis in The Three Hundred

Contents

As presented in Chapter 2, weak gravitational lensing is a unique tool to probe the underlying dark matter distribution around clusters. The most widely used approach, like the one we considered in Chapter 4 to work on DC2 data, consists in estimating the cluster mass and concentration from the measurement of the "monopole" of the tangential shear, i.e., the average in radial bins of the background galaxy tangential ellipticities. However, more information could be extracted from the weak lensing data when considering higher-order multipoles of the shear that are highly sensitive to the ellipticity of the projected mass density. I started this project with a student as part of his Master's end-of-year project. He developed preliminary analysis tools and I pursued and complemented this analysis that is presented in this chapter.

In Section 6.1, we present the motivations to use the shear multipole moments to gain insights on the geometrical properties of the cluster mass distribution beyond sphericity, and the corresponding formalism in Section 6.2. In Section 6.3, we introduce the data set from the project The Three Hundred and we detail in Section 6.4 how to measure the lensing multipoles from background galaxies and how to extract information about the halo triaxiality. The results are presented in Section 6.5, before concluding in Section 6.6.

6.1 Motivations

At optical wavelength, the measurement of the cluster weak gravitational lensing shear is a powerful way to constrain cluster masses, as well as calibrate the cluster scaling relation. However, the inferred weak lensing mass is strongly dependent on the modeling of the underlying three-dimensional matter distribution.

The collapsed dark matter halos are not expected to be spherical, due to the non-spherical initial density peaks from which they form, and also due to their complex individual accretion history in the cosmic web (Sheth et al., 2001). The underlying dark matter halos are shown to have complex triaxial structures in simulations (Jing & Suto, 2002; Schneider et al., 2012; Despali et al., 2014), with a preference of being prolate rather than oblate[1]. Evidence for non-sphericity is also given by the distribution of cluster member galaxies (e.g., Binggeli (1982)) or the X-rays surface brightness (e.g., Lau et al. (2012)). The shear field around clusters has also demonstrated that the cluster mass distribution is not spherical; Oguri et al. (2010) have measured the average axis ratio $q = 0.54 \pm 0.04$ for 25 X-rays selected clusters from HSC two-dimensional shear maps ($q = b/a$, where b and a are respectively the minor and major axis of the elliptical projected mass density).

The *standard* weak lensing approach that we used in Chapter 4 considers that clusters are spherical. However, considering spherical halos may lead to significant bias in the lensing mass and concentration estimates. A variety of studies based on simulations e.g., Becker & Kravtsov (2011); Meneghetti et al. (2010); Lee et al. (2018); Corless & King (2007) (see also Giocoli et al. (2014); Herbonnet et al. (2022); Muñoz-Echeverría et al. (2023) who analyzed The Three Hundred dataset) shows that the lensing masses inferred assuming sphericity are significantly impacted by halo elliptical shape. This effect originates from the combination of elongation and orientation of the mass distribution that may lead to a bias of 30% to the recovered lensing mass. In general, for prolate-shaped halos whose major axes are aligned along the line-of-sight (LOS) the average tangential shear is *boosted* and the inferred weak lensing masses are overestimated, and reversely when their major axes are perpendicular to the LOS, i.e., the matter distribution is elliptical on the sky plane. Cluster triaxiality is therefore a source of systematic uncertainty in the interpretation of the weak lensing signal, that will propagate to cosmological constraints[2].

When the projected dark matter distribution is not spherical, the local shear has non-trivial radial and polar dependencies and can be studied through its multipole decomposition. Using this, Gouin et al. (2017) have measured the aperture multipolar moments of the convergence field in the N-body PLUS simulations to quantify the geometry and topology of the cosmic web. Clampitt & Jain (2016)[3] and Robison et al. (2023)[4] have measured the ellipticity of the galaxy-sized dark matter halos around SDSS Luminous Red Galaxies by analyzing the stacked lensing shear multipoles. In a cluster perspective, Gonzalez et al. (2021)[5] have measured the average projected ellipticity[6] of SDSS redMaPPer galaxy clusters from stacked shear multipoles and found $\langle\epsilon\rangle = 0.21 \pm 0.04$, and Shin et al. (2018)[7] analyzed jointly the shear multipolar moments and the distribution of member galaxies to find $\langle\epsilon\rangle = 0.27 \pm 0.03$.

In this chapter, we aim to test if constraints of several key halo parameters (e.g., the mass, the concentration, the ellipticity of the dark matter distribution) can be improved by using the cluster shear multipoles, at the level of individual clusters.

Indeed, we have seen in Chapter 5 that using an unbinned cluster abundance analysis generally provides tighter cosmological constraints. Therefore, to take full advantage of the next generation of cluster cosmology data, being able to detect asphericity and measure robust masses on individual galaxy clusters (see e.g. Murray et al. (2022)) is a requirement.

[1] Most of them look like a rugby ball, with different ellipticities and orientation with respect to the observer.

[2] Stacked cluster lensing analysis as explored in Chapter 4, should in principle mitigate the projection effect because of their random orientations. But selection bias from cluster finders (Wu et al., 2022) may still yield projection-related effects in a stacked signal.

[3] Using SDSS galaxy shape catalog.

[4] Using UNIONS galaxy shape catalog.

[5] Combining several galaxy shape catalogs including CFHTLenS and KiDS.

[6] In this chapter, we adopt the definition $\epsilon = (1-q)/(1+q)$ for the ellipticity.

[7] Using SDSS galaxy shape catalog.

6.2 Weak lensing shear multipoles in cluster field

Let us first recap a couple of results from weak lensing. The observed ellipticity $\epsilon^{\rm obs}$ of a source galaxy is related to the reduced shear g by

$$\epsilon^{\rm obs} = \frac{\epsilon^{\rm int} + g}{1 + g^* \epsilon^{\rm int}}, \tag{6.1}$$

where $g = \gamma/(1-\kappa)$ and γ is the reduced shear, g^* is the complex conjugate of g, and $\epsilon^{\rm int}$ is the intrinsic ellipticity of the galaxy (see Sections 2.2 and 2.4, and Appendix B). The weak lensing regime is given by $\epsilon^{\rm obs} \approx \gamma + \epsilon^{\rm int}$. In most weak lensing analyses, it is relevant to consider the tangential shear γ_+ and cross shear $\gamma_\times$ instead of the two complex components γ_1 and γ_2 namely $\gamma_+ = -{\rm Re}[\gamma {\rm e}^{-2i\varphi}]$ and $\gamma_\times = -{\rm Im}[\gamma {\rm e}^{-2i\varphi}]$ where φ is the polar angle relative to the cluster center. Replacing the shear γ by the galaxy ellipticity ϵ in the two above equations gives the tangential and cross ellipticities of individual galaxies.

Formally (see e.g. Bernstein & Nakajima (2009)), the local two-dimensional tangential and cross shear fields $\gamma_{+/\times}$ can be expressed following their multipole expansion

$$\gamma_{+/\times}(R,\varphi) = \frac{1}{2}\sum_{m=-\infty}^{m=+\infty}(1+\delta_m)\gamma_{+/\times}^{(m)} e^{im\varphi}, \tag{6.2}$$

where $\gamma_{+/\times}^{(m)}$ is the multi-pole coefficient defined as

$$\gamma_{+/\times}^{(m)} = \frac{1}{(1+\delta_m)\pi}\int_0^{2\pi}\gamma_{+/\times}(R,\varphi)e^{-im\varphi}d\varphi. \tag{6.3}$$

For the tangential shear, the real and imaginary part of the multipole moment $\gamma_+^{(m)} = \gamma_{+,\Re}^{(m)} + i\gamma_{+,\Im}^{(m)}$ are given by

$$\gamma_{+,\Re}^{(m)} = \Re[\gamma_{+/\times}^{(m)}] = \frac{1}{(1+\delta_m)\pi}\int_0^{2\pi}\gamma_{+/\times}\cos(-m\varphi)d\varphi, \tag{6.4}$$

$$\gamma_{+,\Im}^{(m)} = \Im[\gamma_{+/\times}^{(m)}] = \frac{1}{(1+\delta_m)\pi}\int_0^{2\pi}\gamma_{+/\times}\sin(-m\varphi)d\varphi. \tag{6.5}$$

As introduced in Section 2.4, the azimuth averaged tangential/cross shears correspond to the monopole ($m=0$) of the local tangential/cross fields. We now denote them as $\widehat{\gamma}_\times^{(0)}$ and $\widehat{\gamma}_+^{(0)}$ or equivalently the excess surface density $\widehat{\Delta\Sigma}_\times^{(0)}$ and $\widehat{\Delta\Sigma}_+^{(0)}$. They can be measured from galaxy shapes, by making the weighted average of the tangential/cross components of source galaxies in circular annuli such as

$$\widehat{\gamma}_{+/\times}^{(0)}(R) = \frac{1}{\sum_{s=1} w_{ls}}\sum_{s=1} w_{ls}\epsilon_{+/\times}^{ls}. \tag{6.6}$$

The multipole $m \neq 0$ coefficients $\widehat{\gamma}_{+/\times}^{(m)} = \widehat{\gamma}_{+/\times,\Re}^{(m)} + i\widehat{\gamma}_{+/\times,\Im}^{(m)}$ in Eq. (6.4) and Eq. (6.5) can be similarly estimated from the galaxy shapes, such as

$$\widehat{\gamma}_{+/\times,\Re}^{(m)} = \frac{1}{\sum_{s=1} w_{ls}(\cos m\varphi_s)^2}\sum_{s=1} w_{ls}\epsilon_{+/\times}^{ls}\cos(-m\varphi_s), \tag{6.7}$$

$$\widehat{\gamma}_{+/\times,\Im}^{(m)} = \frac{1}{\sum_{s=1} w_{ls}(\sin m\varphi_s)^2}\sum_{s=1} w_{ls}\epsilon_{+/\times}^{ls}\sin(-m\varphi_s), \tag{6.8}$$

where φ_s is the polar angle of the galaxy s with respect to the lens center. From this, we see that the shear multipoles can be easily estimated from the background source sample[8] (performing weighted averages in circular annulus).

Moreover, the shear multipoles can reveal a wealth of information about the halo morphology (Schneider & Weiss, 1991; Bernstein & Nakajima, 2009). We can show that each multipole moment $\gamma_{+/\times}^{(m)}$ is sensitive to the convergence multipole $\kappa^{(m)} = \kappa_{\Re}^{(m)} + i\kappa_{\Im}^{(m)}$ of the same order m (we also refer the reader to Appendix I for the some derivation of these results). Given a source galaxy at redshift z_s and a cluster at redshift z_l, the *local* convergence field writes

$$\kappa(R, \varphi|z_s) = \frac{\Sigma(R, \varphi)}{\Sigma_{\rm crit}(z_s, z_l)}, \tag{6.9}$$

where the projected surface mass density of the galaxy cluster is defined by

$$\Sigma(R, \varphi) = \int dy \rho(R, \varphi, y), \tag{6.10}$$

and where ρ is the three-dimensional matter density of the dark matter halo. In the above equation, y is the LOS coordinate, and the quantity $\Sigma_{\rm crit}(z_s, z_l)$ is the critical surface mass density of the lens-source system. When $\rho = \rho(r)$, i.e., that the three-dimensional matter distribution is radial, we have that $\kappa = \kappa(R)$, and by symmetry $\kappa^{(m)} = \gamma^{(m)} = 0$ if $m \neq 0$. When the halo is spherical, the shear multipoles vanish, and only the average tangential shear $\gamma_+^{(0)}$ and the convergence monopole $\kappa^{(0)}$ remain.

However, when the projected matter distribution is not spherical along the LOS, the convergence/shear multipoles are not null. To describe more accurately the mass distribution, the halo shape is usually well-approximated by a triaxial ellipsoid (see e.g. Jing & Suto (2002); Knebe & Wießner (2006)) that is specified by deformation factors (semi-major and semi-minor axis) in addition to three orientation (Euler) angles. In that formalism, the three-dimensional density is given by

$$\rho(\vec{r}) = \rho_{\rm s}(r_\epsilon), \tag{6.11}$$

where $\rho_{\rm s}$ is a given function of the modulus r_ϵ (such as NFW or Einasto) and r_ϵ is given by

$$r_\epsilon^2 = \frac{x^2}{a_x^2} + \frac{y^2}{a_y^2} + \frac{z^2}{a_z^2}. \tag{6.12}$$

Here, we have introduced the deformation factors a_x, a_y and a_z on each axis (we have not specified in the above equation the Euler angles to orientate the ellipsoid in the Cartesian reference frame for simplicity). The iso-density surfaces are defined by $r_\epsilon = {\rm cst}$ instead of $r = \sqrt{x^2 + y^2 + z^2} = {\rm cst}$ as it is the case for the spherical assumption. When projecting the ellipsoid ρ along the LOS, the surface mass density Σ is then described by two-dimensional elliptical iso-density contours (see e.g. Oguri et al. (2003); Oguri & Keeton (2004)). However, all six deformation and rotation parameters (semi-axes and Euler angles) of the three-dimensional density cannot be deduced from the projected two-dimensional density (see Appendix J). The surface density is instead usually described more simply, accounting for only two *effective* ellipse parameters. When assuming that the total mass of the halo within a sphere with radius r is the same after elliptical deformation, we get $a_z = (a_x a_z)^{-1}$. By projecting the dark matter distribution along the LOS, we have

$$\kappa(R, \varphi) = \frac{1}{a_x a_y} \kappa_{\rm sph} \left(R \sqrt{\frac{\cos^2(\varphi - \varphi_0)}{a_x^2} + \frac{\sin^2(\varphi - \varphi_0)}{a_y^2}} \right), \tag{6.13}$$

[8]The measurement of every shear multipoles enables to access the same level information as with a two-dimensional modeling of the local convergence field by solving the local Poisson equation as it is performed in Oguri et al. (2010).

where $\kappa_{\rm sph}$ is a spherical convergence profile and φ_0 is the projected orientation angle of the ellipse. This expression describes two possible features of dark matter halo shape. If $a_x = a_y = 1$, ρ is spherical. If $a_x = a_y \neq 1$ the halo is prolate/oblate shaped and aligned along the LOS, whereas κ is still radial. In this work, we consider that if the projected matter density is spherical, then the three-dimensional density is spherical too, i.e., we fix $a_x a_y = 1$[9] (see e.g. Gonzalez et al. (2021); van Uitert et al. (2017); Adhikari et al. (2015)). We then define the axis ratio $q = a_y/a_x = a_x^{-2} = a_y^2$, so the convergence in Eq. (6.13) simplifies to

$$\kappa(R,\varphi) = \kappa_{\rm sph}\left(R\sqrt{q\cos^2(\varphi-\varphi_0) + \frac{\sin^2(\varphi-\varphi_0)}{q}}\right). \tag{6.14}$$

In this modeling, φ_0 is considered as a free parameter, compared to Gonzalez et al. (2021); van Uitert et al. (2017) that estimated the average multipoles around several lenses by stacking on a preferred orientation (the elongated axis of the projected mass density probed by the orientation of the BCG or the spatial distribution of the member galaxies). This rotation allowed them to consider $\varphi_0 = 0$, then using that $\kappa_{\Im}^{(m)} = 0$, and then $\gamma_{+,\Im}^{(m)} = 0$. In this work, we aim to constraint φ_0 using weak lensing, so we will not use any tracers of φ_0.

We can use the formalism developed in Adhikari et al. (2015) to capture the trend of the convergence multipoles with respect to the ellipse parameters, namely the orientation angle and the ellipticity of the mass distribution given by $\varepsilon = (1-q^2)/(1+q^2)$. When q is close to 1, the convergence is then given by (see Appendix I)

$$\kappa(R,\varphi) = \kappa_{\Re}^{(0)}(R) + \kappa_{\Re}^{(2)}(R)\cos(2\varphi) + o(\epsilon^2), \tag{6.15}$$

where the second multipole of the convergence is given by

$$\kappa_{\Re}^{(2)}(R) = \frac{\epsilon}{2}\frac{d\ln\kappa_{\Re}^{(0)}}{d\ln R}\kappa_{\Re}^{(0)}(R)\cos(2\varphi_0). \tag{6.16}$$

The second-order multipole reveals the halo morphology, since it is proportional to the dark matter halo projected ellipticity ϵ, so it can be estimated to test the validity of the spherical assumption. We also see that the multipole is proportional to $\cos(2\varphi_0)$, so the orientation of the ellipse on the sky plane defines the sign of $\kappa_{\Re}^{(2)}(R)$.

We represent in Fig. 6.1 the zeroth ($m = 0$), second ($m = 2$) and fourth ($m = 4$) multipoles of the convergence and shear fields by using a NFW profile[10] and choosing the source redshift $z_s = 3$. We consider an angle $\varphi_0 = \pi/3$ and two different axis ratios, namely $q = 0.9$ (blue) and $q = 0.4$ (orange).

The first row corresponds to the convergence, the second and third rows correspond respectively to the tangential and cross shears. The full lines correspond to the real part and the dashed lines correspond to the imaginary part of the multipoles[11]. For $m = 0$ (first column), the "elliptical" shear and convergence monopoles (full lines) are close to that of a spherical distribution (dashed lines), so to disentangle between the shear monopoles produced respectively by an

[9] This particular choice denotes that when $a_x = a_y$, the recovered convergence $\kappa = \kappa_{\rm sph}$. Then, the halo cannot be prolate/oblate shaped with respect to the spherical profile $\kappa_{\rm sph}$. The main argument is that the prolate (or oblate) index (given by a_x is $a_x = a_y$) is strongly correlated with the concentration parameter. In other words, to differentiate between a prolate (respectively oblate) halo aligned along the LOS and a spherical halo with a high (respectively low) concentration parameter is not an easy task to do.

[10] with mass $M = 10^{14} M_\odot$ and concentration $c = 4$.

[11] There is no imaginary part for the monopole ($m = 0$) since $\sin(0 \times \varphi) = 0$ (no blue and orange dashed lines) in the first column.

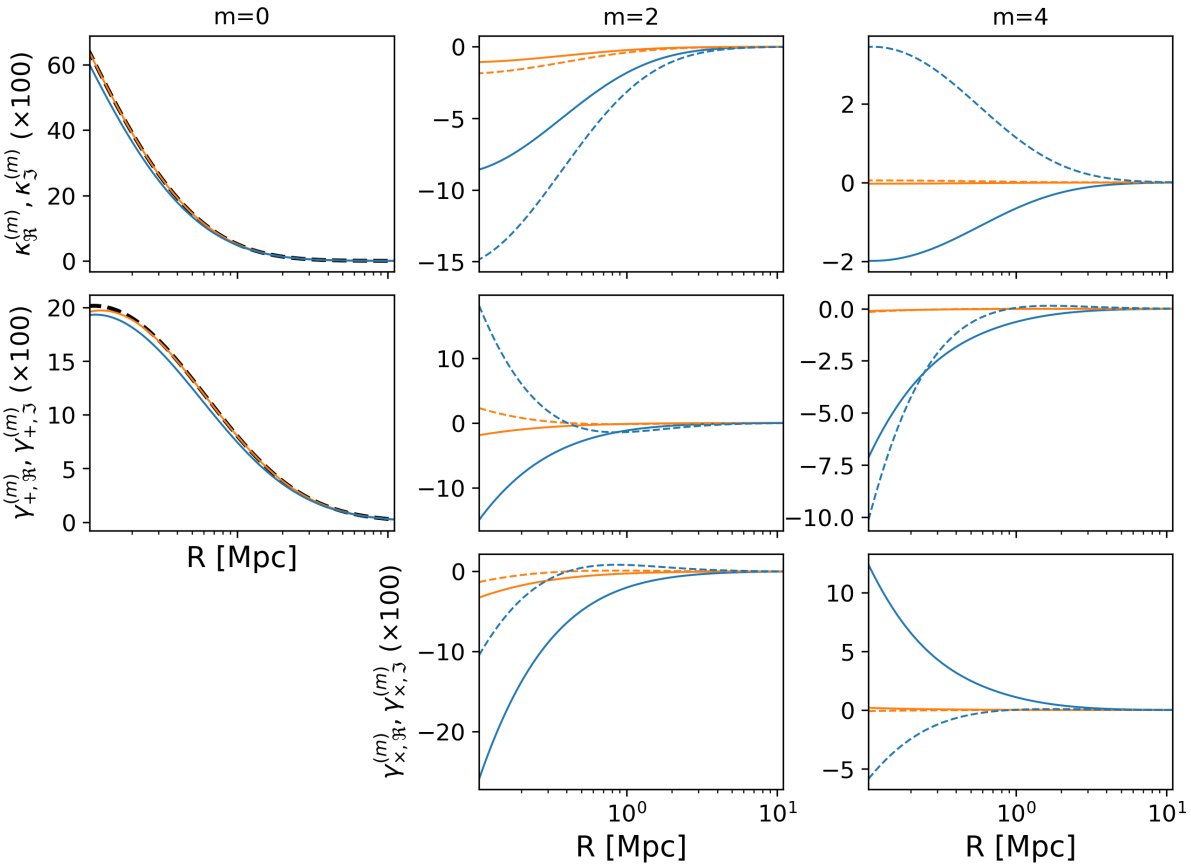

Figure 6.1: Top row: convergence multipoles, Middle row: tangential shear multipoles, Bottom row: cross shear multipole. The dashed and full lines correspond respectively to the real and imaginary parts of the multipoles (for $m > 2$). The blue colored lines correspond to an axis ratio $q = 0.4$ and orange for $q = 0.9$. The black dashed lines in the first column correspond to the radial convergence and shear monopoles for a spherical halo.

elliptical and spherical cluster (i.e., the standard approach) may be difficult. We see that convergence multipoles $m \neq 0$ have different signs, and have a larger amplitude for much deformed halos ($q = 0.4$, in blue). In this example, we can see that the multipoles are highly sensitive to the projected deformation of the mass density. The two other rows in Fig. 6.1 represent the real and imaginary parts of the tangential and cross shear multipole moments, that are derived from the Bernstein & Nakajima (2009) formalism (see the full details Appendix I). There is no monopole for the cross shear field for any type of mass distribution, hence no bottom-left panel (see Appendix K).

6.3 The Three Hundred data

The Three Hundred (The300) project[12] (Cui et al., 2018) provides a large set of simulated galaxy clusters with high mass resolution for cosmological and astrophysical applications, to improve our understanding of the formation history and evolution of galaxy clusters. It consists in the re-simulation of 324 dark matter halos identified by the ROCKSTAR halo finder (Behroozi

[12]The official website of the project is available here: `https://the300-project.org`.

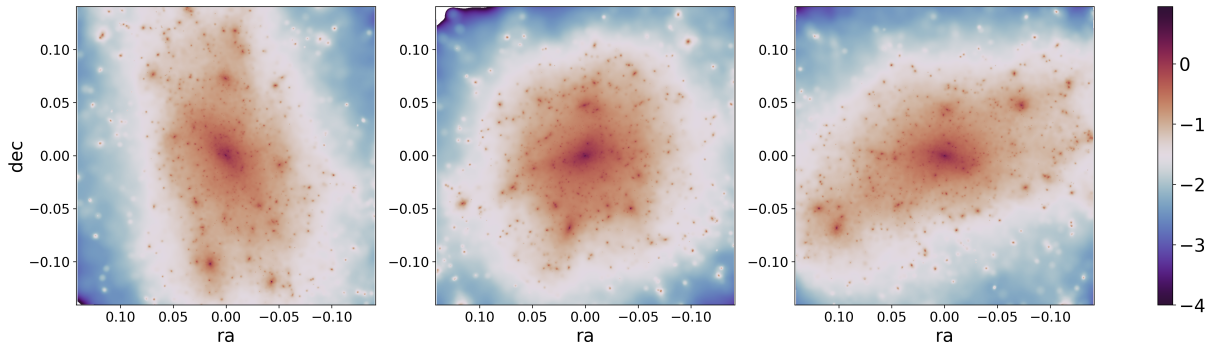

Figure 6.2: Convergence maps for three different orthogonal projections X, Y and Z along the LOS for the same simulated cluster from THE300.

et al., 2013) in the Multi-Dark Planck 2^{13} (MDPL2) N-body simulation (Klypin et al., 2016), the latter having 3840^3 dark matter particles and in a 1 h^{-1} Gpc^3 box. MDPL2 follows the Planck Collaboration et al. (2016) cosmology, i.e., $h = 0.678$, $n_s = 0.96$, $\sigma_8 = 0.823$, $\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, and $\Omega_b = 0.048$.

To produce full-physics hydrodynamics simulations of galaxy clusters, a spherical region of radius $15h^{-1}$Mpc has been identified around each of the 324 most massive objects within the host MDPL2 box. These regions have been re-simulated at higher resolution using the full-physics hydro-dynamical codes GADGET−MUSIC (Sembolini et al., 2014) and GADGET−X (Rasia et al., 2015) to include several baryon processes (radiative processes, star formation, thermal feedback from supernovae, supermassive black holes with AGN feedback, etc., see full details in Cui et al. (2018)).

THE300 galaxy cluster masses range from $6.4 \times 10^{14} M_\odot < M_{200\mathrm{m}} < 26.5 \times 10^{14} M_\odot$ and 128 simulation snapshots were saved between $0 < z < 17$. From the total cluster mass distribution, including gas, stars, and dark matter, the lensing properties of each cluster have been derived following the procedure detailed in Herbonnet et al. (2022), based on Meneghetti et al. (2010, 2020)[14]. The resulting weak lensing maps γ_1, γ_2 and κ have a size of 5 × 5 Mpc, spatially resolved with 2048 × 2048 pixels. Finally, the weak lensing maps for each cluster are available for different projections along the LOS. Fig. 6.2 shows the convergence maps for cluster #19 of the THE300 sample, for three different orthogonal projections. We see that the cluster has numerous substructures (the redder points). Each *average* projected shape represented in Fig. 6.2 seems to be well approximated by an ellipsoid, since the first and third projections seem roughly elliptical and the second is almost circular.

[13]The simulations are publicly available at the `https://www.cosmosim.orgdatabase`.

[14]The method used to derive the lensing maps is different from ray-tracing, that was used to compute the convergence and shear at galaxy position in cosmoDC2. Here, the deflection angle of a light ray coming from a point source at $z_s = 3$ and at location $\vec{\theta}$ is given by (Schneider, 2006)

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi} \int d^2\vec{\theta} \kappa(\vec{\theta}') \frac{\vec{\theta} - \vec{\theta}'}{|\vec{\theta} - \vec{\theta}'|^2}, \tag{6.17}$$

where κ is the cluster convergence. The shear maps γ_1 and γ_2 are obtained via

$$\gamma_1 = \frac{1}{2}\left(\frac{\partial \alpha_1}{\partial \theta_1} - \frac{\partial \alpha_2}{\partial \theta_2}\right) \quad \text{and} \quad \gamma_2 = \frac{\partial \alpha_2}{\partial \theta_1}. \tag{6.18}$$

6.4 Methodology

In this work, we aim to go beyond the spherical modeling to measure the weak lensing mass of clusters from the project THE THREE HUNDRED. In particular, we will test several inference setups (considering spherical or elliptical modeling, monopole only or combined multipole analysis) by comparing the lensing mass estimates obtained for each cluster projection. The compatibility between these per-projection mass estimates provides a direct test of the reliability of the halo mass modeling.

6.4.1 Rubin LSST-like background galaxy sample

We want to perform realistic weak lensing analysis to infer the properties of clusters from THE300. To do so, we first need to generate a mock galaxy catalog, representative of the LSST data. We start by generating a mock sample of un-lensed galaxies. Their distribution in redshift follows the Chang et al. (2013) distribution and we choose a LSST-like galaxy number density of 30 gal.arcmin^{-2}. We then generate intrinsic ellipticities using a shape noise $\sigma = 0.25$ on each ellipticity component.

From this unlensed catalog, we use Eq. (6.1) to derive lensed ellipticities from THE300 lensing shear $\gamma_{\rm The300}(z_0)$ and convergence $\kappa_{\rm The300}(z_0)$ maps computed at source redshift $z_0 = 3$. To do so, we need to compute the lensing maps for different source redshifts. We follow the methodology in Herbonnet et al. (2022) (see their Section 2.2) such that the lensing maps $\gamma(z_i)$ and $\kappa(z_i)$ at source redshift z_i are given by

$$(\gamma(z_i), \kappa(z_i)) = \frac{\Sigma_{\rm crit}(z_l, z_0)}{\Sigma_{\rm crit}(z_l, z_i)} \times (\gamma_{\rm The300}(z_0), \kappa_{\rm The300}(z_0)), \tag{6.19}$$

where z_l is the cluster redshift. The re-scaling of the THE300 lensing maps is valid only if we consider that the distortion of background galaxy shapes is caused by the presence of a single lens, i.e., the cluster[15]. In reality, this would not be the case as all the structures lying along the LOS contribute to the lensing signal. So, we created a realistic redshift catalog, but it still needs to account for observational effects, as we explored in Chapter 4 (errors in shape measurement, photometric redshift, etc.).

6.4.2 Cluster lensing multipole analysis

As in Chapter 4, we use the excess surface density $\Delta\Sigma$ which follows the same multipole decomposition as the shear by replacing $\gamma(R,\varphi) \to \Delta\Sigma(R,\varphi)$, $\gamma_+(R,\varphi) \to \Delta\Sigma_+(R,\varphi)$, $\gamma_\times(R,\varphi) \to \Delta\Sigma_\times(R,\varphi)$ and $\kappa(R,\varphi) \to \Sigma(R,\varphi)$. From Eqs. (6.7) and (6.8), the excess surface mass density multipole estimators are then given by

$$\widehat{\Delta\Sigma}^{(m)}_{+/\times,\Re} = \frac{1}{\sum_{s=1} w_{ls}(\cos m\varphi_s)^2} \sum_{s=1} w_{ls}\Sigma^{s,l}_{\rm crit}\epsilon^{ls}_{+/\times}\cos(-m\varphi_s), \tag{6.20}$$

$$\widehat{\Delta\Sigma}^{(m)}_{+/\times,\Im} = \frac{1}{\sum_{s=1} w_{ls}(\sin m\varphi_s)^2} \sum_{s=1} w_{ls}\Sigma^{s,l}_{\rm crit}\epsilon^{ls}_{+/\times}\sin(-m\varphi_s). \tag{6.21}$$

The sums run over all lens-source (l,s) pairs, the weights w_{ls} are defined in Chapter 2 and account for the lens and source redshifts.

Since we have access to the true underlying shear maps, we have performed a sanity check by comparing the estimated multipole moments (from background sources) to the prediction

[15]Refereed as the "single-lens" approximation (Schneider et al., 1992).

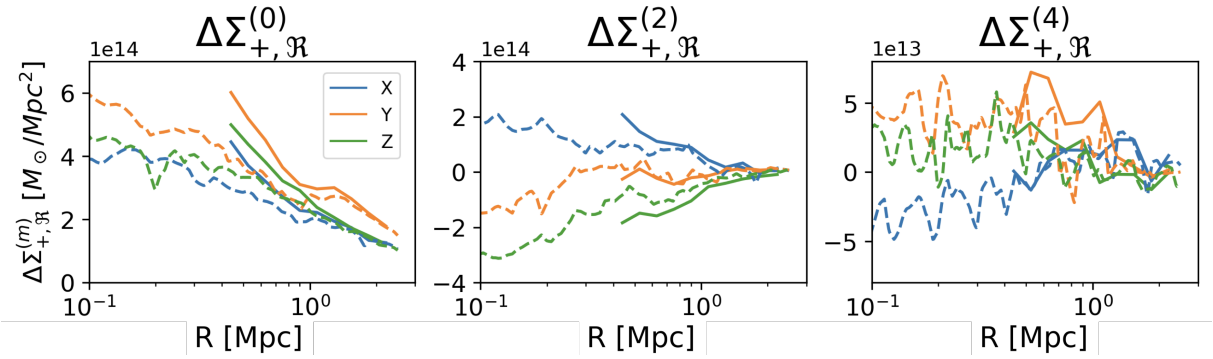

Figure 6.3: From left to right: Tangential excess surface density multipoles (real part, $m = 0$: left panel, $m = 2$, middle panel, $m = 4$: right panel). The different colors correspond to the three orthogonal projections along the LOS. The full lines correspond to the estimated profiles from background source samples and the dashed lines to the multipole prediction derived from the shear and convergence maps of THE300.

using directly the THE300 convergence and shear maps, as presented in Appendix I. To do so, we have considered a very idealistic catalog of background galaxies compared to the criteria we defined in Section 6.4.1. We use $n_{\rm gal} = 100$ gal.arcmin^{-2} to decrease significantly the shot noise variance on the recovered signal (for this test, we also consider that background sources have no intrinsic ellipticity, i.e. $\sigma = 0$). We consider 10 radial bins within the range $[0.5, 3.5]$ Mpc[16].

Fig. 6.3 shows the $m = 0, 2, 4$ multipole moments of the tangential excess surface density[17], estimated from the background source sample (full lines, from left to right) using Eq. (6.20) and for the three orthogonal projections X (blue), Y (orange) and Z (green) along the LOS of a THE300 cluster. We see that the per-projection signals differ from one another in sign (for $m = 2, 4$) and amplitude (for $m = 0, 2, 4$). The dashed lines correspond to the direct derivation of the excess surface density multipoles from the shear and convergence maps. For scales $R \geq 1$ Mpc, where the weak lensing approximation is valid, the estimated and predicted signals are in reasonable agreement which validates our overall procedure. They start diverging at lower radii. This is due to the non-weak lensing contribution[18] that affects the measurement of the cluster lensing signal (see e.g. Mandelbaum et al. (2006); Johnston et al. (2007)). This effect is discussed in Appendix I. In that regime, each m-th shear multipole depends on every convergence multipole. Here, we will consider $R > 0.7$ Mpc for the Bayesian inference of halo parameters to stay in the weak lensing regime, and we plan to account for them as a perspective for this work. We also mention that we consider that the true cluster center is given by the center of the convergence map. However, any deviation from the true cluster center (called the miscentering) would generate lensing multipoles, even for a perfect spherical cluster. In this work, we do not consider miscentering in the modeling of the lensing signal. Moreover, we do not extract information from odd multipoles, since we choose an elliptical modeling of the cluster density, that predicts null odd multipoles, even if odd multipoles would provide a wealth of information about complex morphologies of the cluster mass distribution.

[16]The maximum radial value is given by the size of the aperture of the THE300 maps.

[17]We do not show the error bars in this example.

[18]The weak lensing approximation is $\kappa \ll 1$ such that $g = \gamma/(1-\kappa) \approx \gamma$, which is valid at large radii. In the innermost regions, the convergence is larger and the weak lensing approximation may be no longer valid, especially for very massive clusters. For a massive cluster with mass $M = 10^{15} M_\odot$ with concentration $c = 4$, the convergence $\kappa(R = 0.1) \sim 0.3$ (the reduced shear is given by $g = \gamma/(1-\kappa) = 1.5\gamma$), $\kappa(R = 0.7) \sim 0.05$ (such as $g = 1.1\gamma$) and $\kappa(R = 1) \sim 0.03$ (then, $g = 1.05\gamma$).

From the estimated excess surface density multipoles (cross and tangential components) given in a more synthetic way

$$\widehat{\Delta\Sigma} = \{\widehat{\Delta\Sigma}^{(m)}_{+/\times,\Re} + i\widehat{\Delta\Sigma}^{(m)}_{+/\times,\Re}\}_{m\geq 0}, \tag{6.22}$$

we estimate the posterior distribution of halo parameters $\vec{\theta} = (\log_{10} M, q, \phi_0)$ (we use a $c(M)$ relation from Bhattacharya et al. (2011)) such as

$$\mathcal{P}(\vec{\theta}|\{\widehat{\Delta\Sigma}\}) \propto \mathcal{L}(\{\widehat{\Delta\Sigma}\}|\vec{\theta})\pi(\vec{\theta}), \tag{6.23}$$

where π is the prior, and $\mathcal{L}(\widehat{\Delta\Sigma}|\vec{\theta})$ is a Gaussian likelihood. We estimate the covariance matrix of the Gaussian likelihood using jackknife re-sampling. We define $N_{\rm JK}$ jackknife regions, defined as equal-sized sky areas with constant angular aperture from the cluster center. The covariance of the cluster lensing multipoles is then obtained by applying the same methodology as for the stacked lensing profile presented in Section 2.4. Moreover, we use a flat prior for the cluster (log)mass within $[14, 16]$. We restrict the axis ratio to $q < 1$. The symmetry property $\phi_0 \to \phi_0+\pi$ that gives $\kappa^{(m)} \to \kappa^{(m)}$ implies that $\phi_0 \in [0, \pi]$. Since $\kappa^{(2)}_{\Re} \propto \cos(2\varphi_0)$ according to Eq. (6.16), $\gamma^{(2)}_{+,\Re}$ is invariant under the transformation $\varphi_0' = \pi - \varphi_0$. We then restrict the rotation angle between 0 and $\pi/2$.

6.5 Preliminary results

I developed a software[19] that provides modules dedicated to the prediction of the lensing shear multipoles, from an elliptical convergence map[20]. This software allows us to estimate the shear multipole from background galaxies and provides several analysis tools to compute the lensing of background galaxies from the THE300 lensing maps. We now turn to the preliminary results we obtained using this code, by considering first the monopole of the tangential excess surface density, and then its monopole and multipole $m = 2$ (real part) this time using LSST-like statistics, i.e., $n_{\rm gal} = 30$ gal.arcmin^{-2}.

Monopole analysis

For a subset of 40 clusters with redshift $z_{\rm cl} = 0.33$, we first consider the monopole moment of the tangential excess surface density $\widehat{\Delta\Sigma}^{(0)}_{+,\Re}$ and we use both a spherical and an elliptical model for the cluster mass density.

Fig. 6.4 shows the lensing mass estimates with respect to their virial mass $M_{\rm 200m}$ calculated as the sum of dark matter particles in the simulations. By using the methodology presented in Section 6.4, the first row shows the masses obtained assuming halos to be elliptical and the second row for a spherical modeling[21]. The inferred masses correlate with the true dark matter halo masses. We also note that the higher the true mass is, the lower the error on the recovered lensing mass, since more massive clusters have high SNR. For more massive clusters, the mass estimates are slightly larger (we find $M_{\rm WL}/M_{\rm The300} \approx 1.25 \pm 0.4$); this possibly originates from the modeling of the spherical density profile we choose for the analysis, but also from the not-so-weak lensing regime that the modeling we use does not account for and that still occurs for very

[19] available here: `https://github.com/payerne/ThreeHundredWL`.

[20] For this work I used CLMM (Aguena et al., 2021b) to predict the convergence.

[21] The modeling of the excess surface density profile for a spherical halo is given by $\langle\Sigma(< R)\rangle - \Sigma(R)$ where Σ is the radial projected mass density profile. For an elliptical halo, it is given by $\langle\Sigma^{(0)}(< R)\rangle - \Sigma^{(0)}(R)$ where $\Sigma^{(0)}$ is the azimuth averaged projected mass density profile.

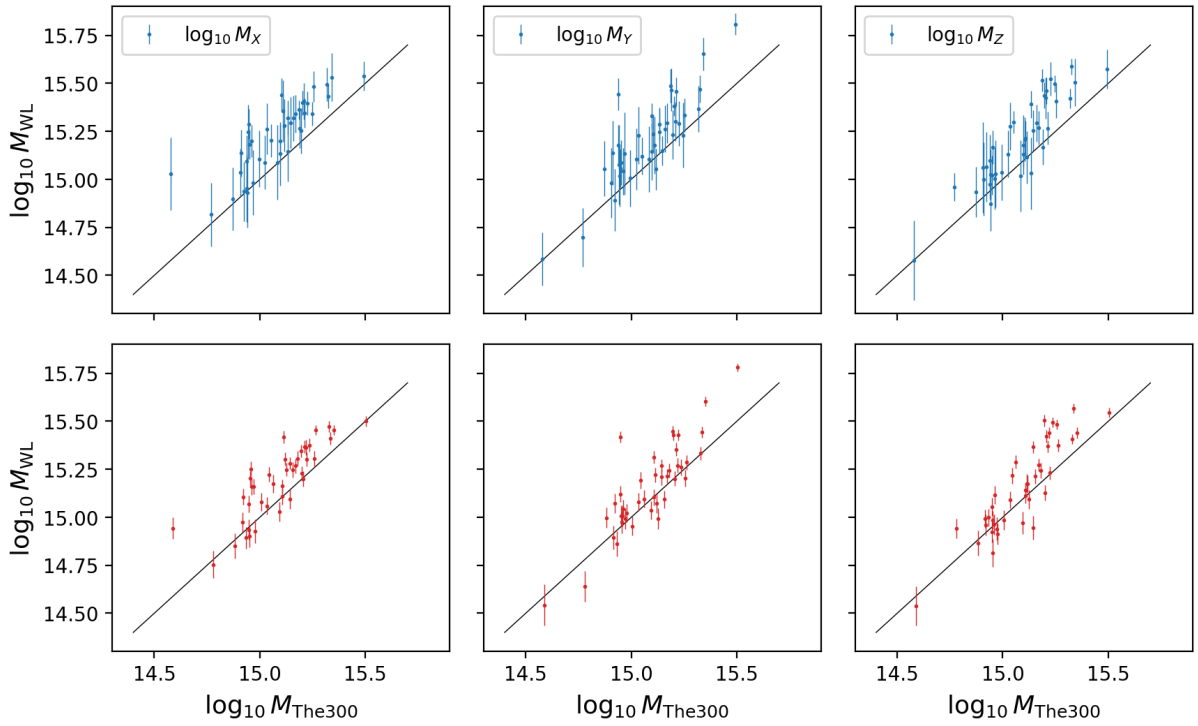

Figure 6.4: Weak lensing mass estimates for a shear monopole analysis for the three different projections (columns). First row: elliptical modeling. Second row: spherical modeling.

massive halos. For masses obtained with the elliptical modeling, the error bars are significantly larger than the spherical modeling, this is due to the strong correlation between the cluster mass and the axis ratio q.

The Fig. 6.6 (top) shows the joint posterior of the cluster mass, the axis ratio and the orientation angle for three different projections of the same cluster. We see (for example for the X projection in blue) that the axis ratio reaches all possible values permitted by the prior range up to very large deformations of the mass distribution (up to $q \approx 0$). This is easy to understand since the monopole of an elliptical distribution is invariant under a change in its orientation angle φ_0, so φ_0 is not constrained. The axis ratio q relies on the identification of the major axis (that is defined by the rotation angle φ_0) so it is also not well constrained. We find that for the Y projection, the inferred mass is larger than that obtained for the two other projections.

Multipole analysis

After considering only the monopole moment of the tangential excess surface density, we test the impact of considering higher multipole moments on the inference of cluster properties. For this analysis, we restrict ourselves to the multipole $m = 2$ of the tangential excess surface density (real part only[22]). The corresponding data vectors for the three orthogonal projections are represented in blue (X projection), orange (Y) and green (Z) in Fig. 6.5 for the cluster presented in Fig. 6.2.

The first result of the multipole analysis is that with LSST-like statistics (level of shape

[22]In this analysis we only test the impact of considering the real part of the tangential shear (as it is done in Gonzalez et al. (2021)) but in principle, we could also use the imaginary part, that has another dependency with respect to the orientation angle φ_0.

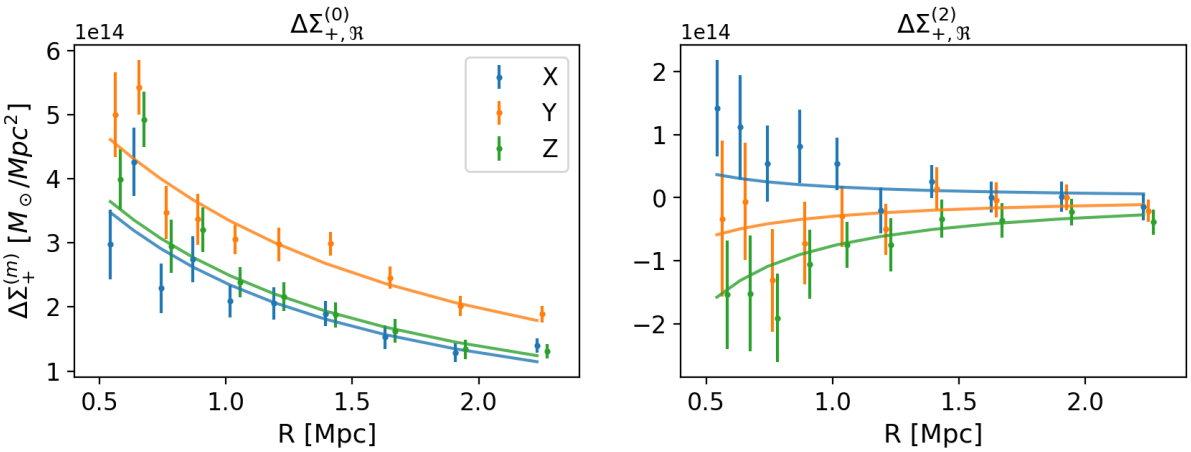

Figure 6.5: Estimated (data points and error bars) and best fit (full lines) $\Delta\Sigma^{(0)}_{+,\Re}$ (left panel) and $\Delta\Sigma^{(2)}_{+,\Re}$ (right panel) for the three different projections along the line-of-sight (respectively blue, orange and green). See text for details.

noise, number density and redshift distribution of galaxies), it is possible to measure the shear multipoles $m = 2$ above 2σ from 0 for individual massive clusters while being a noisy measurement. We see that the Y projection gives a $\Delta\Sigma^{(0)}_{+,\Re}$ higher in amplitude compared to the two other projections. Moreover, the Y-$\Delta\Sigma^{(2)}_{+,\Re}$ is smaller in amplitude compared to the two others. The multipole moments for the two other orientations show significant positive and negative values, revealing the different orientations. Since the three orientations describe the same matter distribution but seen from different angles, we can state that in the context of an ellipsoidal matter distribution, the Y orientation corresponds to the major axis aligned along the LOS (the standard lensing signal is boosted, with no traces of projected ellipticity). The two other projections "look" more elliptical on the sky plane.

We show in Fig. 6.6 (bottom) the posterior distribution of $\log_{10} M, q$ and φ_0 for this cluster, by fitting the estimated lensing signals $\widehat{\Delta\Sigma}^{(0)}_{+,\Re} + \widehat{\Delta\Sigma}^{(2)}_{+,\Re}$ for each of the three projections. We find that the per-projection orientation angle and axis ratio are better constrained compared to the monopole analysis (only fitting $\widehat{\Delta\Sigma}^{(0)}_{+,\Re}$ see Fig. 6.6, top). However, the best-fit masses and corresponding errors are not very different from the ones obtained with the monopole. Regarding the Y projection, the axis ratio is compatible with $q = 1$. So the projected mass distribution is roughly spherical, inducing small multipole $m = 2$ (as we can deduce from the Fig. 6.2). This illustrates the inherent problem behind the interpretation of cluster lensing for an "almost" spherical projected mass distribution: it is difficult to disentangle between a massive spherical cluster and a less massive ellipsoidal cluster but with a major axis oriented along the LOS. In that specific orientation of the cluster mass distribution, the cluster seems prolate-shaped with a major axis aligned along the Y axis. The best-fit predictions (using the mean parameters over the joint posterior) are over-plotted on the data vectors in Fig. 6.5.

We repeated this analysis to our sample of 40 clusters and compared the recovered masses to the ones obtained fitting only the monopole + spherical modeling.

Fig. 6.7 shows the M_Y masses versus the M_Z masses, and M_Y versus M_X. The red points correspond to the masses inferred from the monopole and quadrupole lensing profiles (elliptical modeling), and the blue points to that obtained with the monopole + spherical modeling. If the mass reconstruction did not suffer from projection effects, every point should be on the $x = y$ line

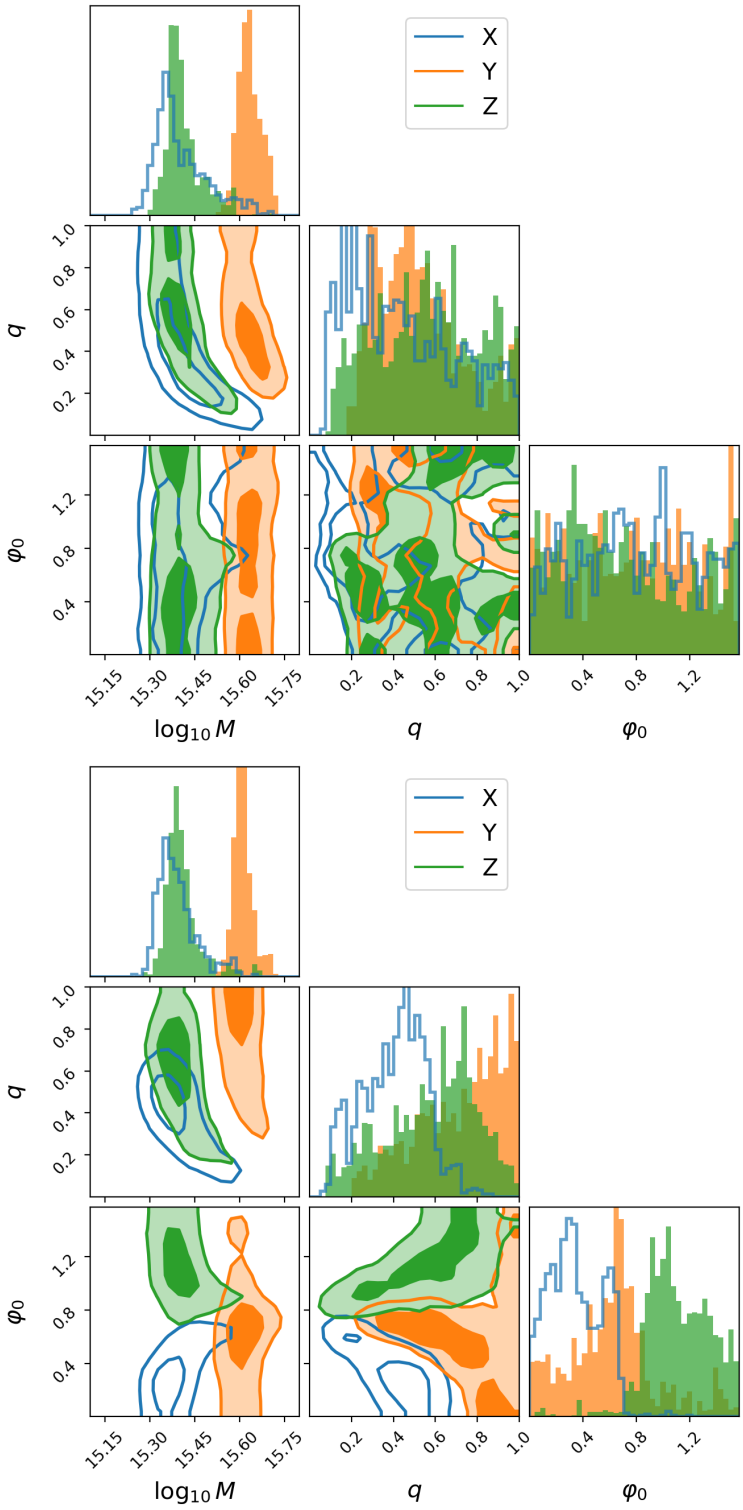

Figure 6.6: Top: Posterior distributions for the parameters $\log_{10} M, q, \varphi_0$, obtained by analyzing the tangential monopole (real part) $\widehat{\Delta\Sigma}^{(0)}_{+,\Re}$. The different colors (blue, orange and green) correspond to the three orthogonal projections along the LOS for a given cluster. Bottom: Posteriors obtained with the multipole analysis, i.e., with $\widehat{\Delta\Sigma}^{(0)}_{+,\Re}$ and $\widehat{\Delta\Sigma}^{(2)}_{+,\Re}$.

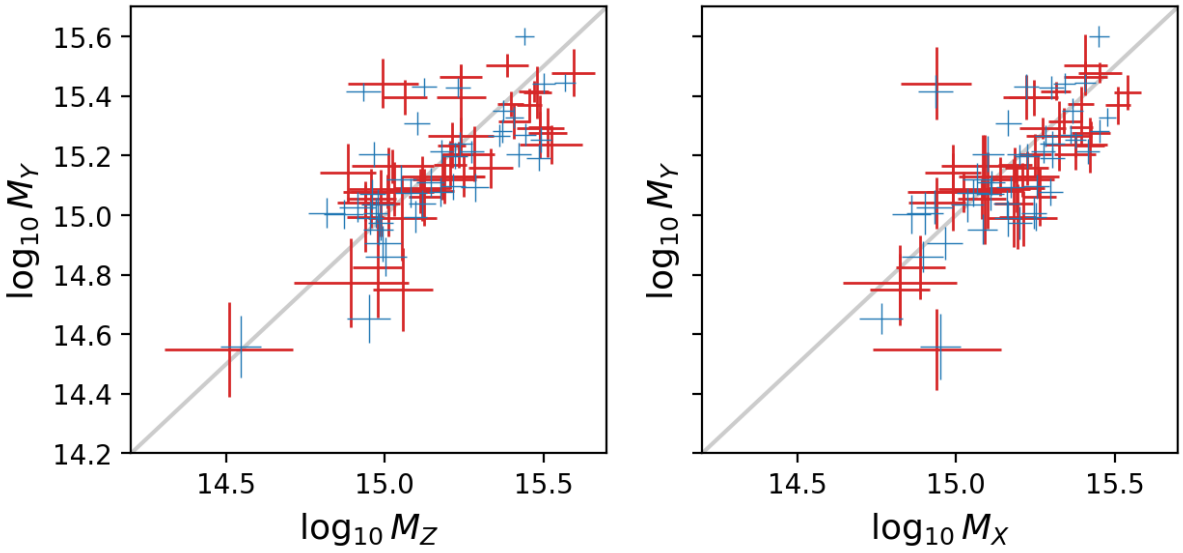

Figure 6.7: Lensing mass estimates for the projection Y versus the projection Z, and projection Y versus projection X. Blue: monopole + spherical modeling, red: monopole and quadrupole + elliptical modeling.

(in blue), i.e., the mass would be the same independent of the projection, i.e., $M_X = M_Y = M_Z$. However, due to the complex shape of the dark matter halo, each projected lensing mass is different from one another. We note that using shear multipole + elliptical modeling (red), the compatibility between per-projection lensing masses is improved, since mass posteriors are more overlapping compared to the monopole + spherical modeling case (blue); this is however due to larger error bars resulting from the added degrees of freedom rather than an improvement on the mass.

We show in Fig. 6.8 the difference between the lensing mass estimates obtained for the projection i and j given by $\Delta(\log_{10} M)_{ij} = \log_{10} M_i - \log_{10} M_j$, from the monopole + spherical modeling, versus the ones obtained from the quadrupole analysis. Again, we see that the impact of using the shear multipole $m = 2$ on mass reconstruction has no significant effect, since the mass differences correlate positively between the monopole and multipole cases.

So, from the above, we find that the mass is not better constrained when using the multipole analysis. We now turn to see the effect on the ellipticity of the cluster. The average projected ellipticity per projection is defined by $\epsilon = (1-q)/(1+q)$. The ellipticity estimates per projection are shown in Fig. 6.9. We find that from the multipole analysis $\langle\epsilon\rangle = 0.25 \pm 0.09$ compared to the monopole + elliptical modeling that provided $\langle\epsilon\rangle = 0.32 \pm 0.07$. The value inferred with multipoles is more compatible with other estimates of the mean cluster ellipticity such as that Gonzalez et al. (2021) who found $\langle\epsilon\rangle = 0.21 \pm 0.04$ from stacked shear multipoles.

6.6 Conclusions and perspectives

It is essential to interpret correctly the complex features of the shear to infer the properties of the cluster mass distribution from weak gravitational lensing. In this chapter we present how the lensing shear can be described with a series of multipole coefficients, that can be used to reconstruct the overall shear map in the surroundings of the cluster. These multipoles are efficient to test whether the cluster mass distribution deviates from sphericity, and if so they

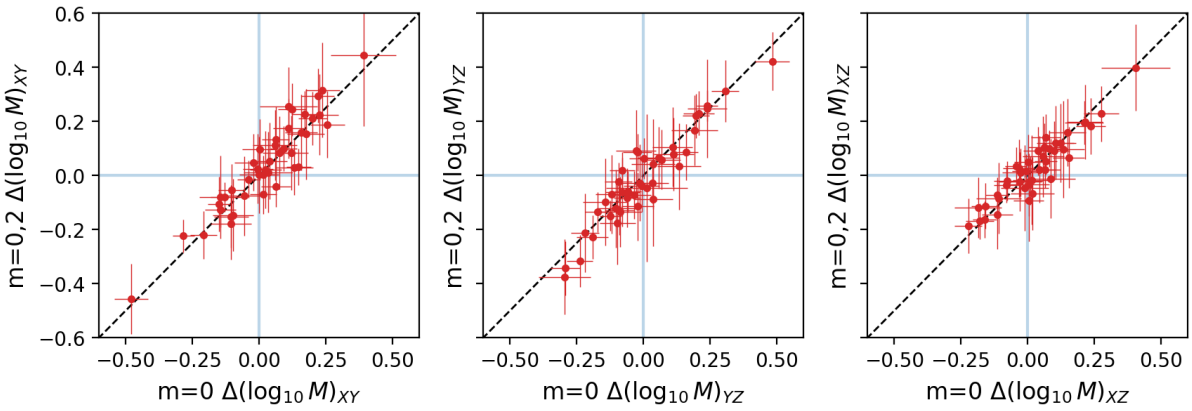

Figure 6.8: Difference between the lensing mass estimates obtained for the projection i and j given by $\Delta(\log_{10} M)_{ij} = \log_{10} M_i - \log_{10} M_j$, from the monopole + spherical modeling (x-axis), versus the ones obtained from the multipole analysis + elliptical modeling (y-axis).

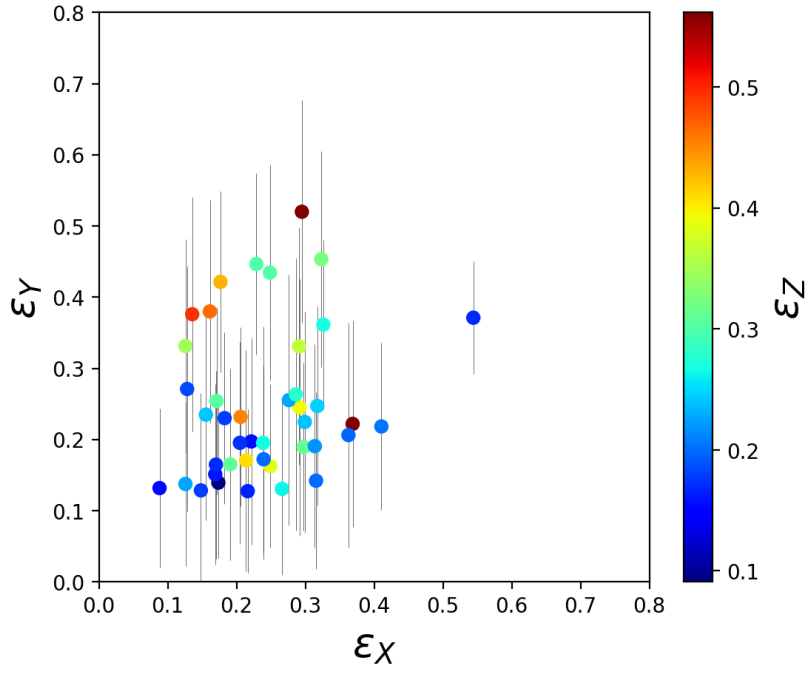

Figure 6.9: Ellipticity estimates for the Y projection against that obtained for the X projection. The color bar corresponds to the ellipticity on the Z projection. We only show the error bars of ϵ_Y.

provide information on the projected ellipticity when releasing the spherical assumption. We developed the software to perform lensing multipole analysis that we applied on the simulated galaxy clusters from the project The Three Hundred, which enabled us to reconstruct the halo properties for different projections along the LOS.

To do so, we first generated LSST-like galaxy catalogs, namely with a typical density $n_{\rm gal} = 30$ gal.arcmin^{-2}. With these statistics, we show that it is possible to measure/detect the order 2 shear multipoles. However, using this information, we did not find a strong impact on the weak lensing mass; In fact, this work says that three observers seeing the same galaxy cluster from three different views will not measure the same mass using multipoles or not[23]. Nonetheless, this information is found to be a promising approach to infer the halo orientation and ellipticity at the individual level, for massive clusters showing high lensing SNR, and independently to baryonic tracers such as the BCG orientation.

These results represent only a fraction of the total The300 data, and we aim to use the whole dataset as a next step. The analysis pipeline is still under construction and I still need to validate the prediction module and the data analysis tools on ideal elliptical mock clusters. The validation methodology is straightforward since ideal mocks can be generated from elliptical $\kappa_{\rm mock}$ map and using the Kaiser & Squires (1993) inversion formula[24]. From that, we can test the analysis pipeline by recovering the input mass and ellipticity. I am currently pursuing this validation, by using the implementation of the Kaiser and Squires formulae in the `lenspack`[25] code.

Since the baryon information is available for each The300 cluster, we aim as a perspective to compare the baryonic tracer of the halo ellipticity to the weak lensing estimate.

We also aim to study the impact of accounting for more multipoles to refine the constraints on the cluster ellipticity, a first step will be necessary to evaluate the feasibility (e.g., forecasting the signal-to-noise ratio) of analyzing these signals with LSST-like statistics. We also want to account for prolate/oblate deformation in cluster mass inference, to check whether it improves the statistical compatibility between projected masses. We expect the prolate/oblate index to be highly correlated to the concentration parameter. Moreover, in this work, we only considered the weak lensing regime, we aim to account for the not-so-weak lensing correction that is not negligible for the most massive clusters. We started some preliminary calculations in Appendix I, and we aim to re-analyze the lensing data accounting for this effect.

[23] This is due to the hypothesis $a_x a_y = 1$ in the modeling of the cluster density, that does not account for flatness/elongation of the mass distribution along the LOS.

[24] Following Kaiser & Squires (1993), the shear is given in terms of the convergence maps by

$$\gamma_{\rm mock}(\vec{\theta}) = \int d^2\vec{\theta}' \mathcal{D}(\vec{\theta}' - \vec{\theta}) \kappa_{\rm mock}(\vec{\theta}') \quad \text{where} \quad \mathcal{D}(\vec{\theta}' - \vec{\theta}) = \frac{(\theta_2 - \theta_2')^2 - (\theta_1 - \theta_1')^2 - 2i\,(\theta_1 - \theta_1')\,(\theta_2 - \theta_2')}{|(\vec{\theta} - \vec{\theta}')|^4}. \tag{6.24}$$

[25] Available here `https://github.com/CosmoStat/lenspack`.

Chapter 7

Conclusions and Perspectives

Contents

Galaxy clusters are primary cosmological probes to study the growth of structures and the expansion history of the late Universe. The next-generation surveys will provide the community with large cluster catalogs that will enable to refine cosmological models at unprecedented level of precision. The current limiting factor is the uncertainties in the cluster scaling relation, which can be calibrated from weak gravitational lensing. These last three years have allowed me to work on a variety of topics related to cluster cosmology, both in the context of the large DESC collaboration, and of joint work with a few colleagues (as part of the likelihood and The300 project). I first summarize the several projects that I presented in this manuscript, and the last section is dedicated to present my view of the various directions that can be pursued as a continuation of my Ph.D. activities.

7.1 Conclusions

In this thesis, I have explored many aspects of cosmology with optically selected galaxy clusters, in the era of the next-generation surveys such as the Rubin LSST, ranging from the cluster catalog, the use of lensing for deriving cluster masses and the calibration of the scaling relation, to the Bayesian inference of cosmological parameters.

I first contributed to the Dark Energy Science Collaboration pipeline in preparation for the analysis of the upcoming LSST data. I developed several modules to perform the cosmological analysis of the abundance and lensing of clusters in the CLCosmoSim code, and I contributed to the CLMM code, dedicated to the lensing measurement and lensing cluster mass inference. Using the DC2 simulated dataset produced by the collaboration, I focused on how cluster lensing can be exploited efficiently to infer cluster masses and how it is impacted by systematic effects. Using the pipeline codes that I developed I obtained preliminary cosmological constraints from a joint analysis of redMaPPer cluster abundance and cluster lensing.

I also developed a methodology based on cluster lensing multipoles to extract information on cluster morphology, by using high-resolution simulations from The Three Hundred project. The goal of this work was to check whether this could improve the lensing mass calibration, by comparing the lensing mass estimates obtained from different projections along the line-of-sight. From this work, I found that including multipole moments of the local shear permits to measure the ellipticity and orientation of the projected mass distribution. However, it does not strongly impact the lensing mass reconstruction compared to the spherical approach, but increases the error bars. I identified a strong degeneracy between the lensing mass estimate and elongation/flattening parameter for aligned clusters along the LOS.

I also focused on testing the accuracy of likelihoods for cluster abundance that are prerequisites to draw the posterior distribution of cosmological parameters. Using the wrong likelihood may lead to biased cosmological constraints, and/or to under/overestimated errors. With a few collaborators, I have developed a methodology to test the robustness of cluster count likelihoods widely used in the literature, namely the Poisson, the Gaussian and the Gauss-Poisson Compound likelihood (GPC). For this work, we used the PINOCCHIO simulated dark matter halo catalogs, and we compared the individual posterior variance of relevant cosmological parameters obtained on each simulation, to the dispersion of best fits over the full set of simulations. The main finding is that for a variety of cosmological inference setups, the Poisson likelihood underestimates errors by 20-30%, whereas the Gaussian and GPC give similar and robust errors.

In addition, I worked on the unbinned cluster count formalism, that improves constraints on cosmology directly from the local dark matter halo distribution in mass and redshift. I have developed a new unbinned cluster count formalism allowing us to account for the effect of Super-Sample Covariance (SSC), which is an inherent source of scatter for large-scale structure observables, such as galaxy clusters. I obtained preliminary cosmological constraints and I compared the standard and the SSC methods. From the preliminary analysis of one PINOCCHIO mock, I did not find a strong impact of SSC on parameters posteriors (the constraints are roughly similar compared to the standard unbinned approach, with a much higher computational cost), but we aim to push the analysis further and to test several cosmological inference setups.

7.2 Perspectives

These are several directions into which the projects I undertook could evolve in. I give a non-exhaustive list of options below, before turning back to the overall experimental landscape of the next decade and to the synergies opportunities it will present.

7.2.1 Cluster cosmology with LSST data

As part of my work within DESC, I contributed to develop several tools dedicated to the cosmological analysis of the LSST cluster catalogs, that are still in preparation and being tested. I aim to pursue getting involved in the pipeline discussion, especially contributing to the official DESC likelihood code FireCrown[1]. These new tools will need to be tested on the DC2 data such as the DESC code CLCosmoSim that I have developed, that served as a place for discussion about FireCrown. Beyond my work on simulations, the LSST data will arrive very soon, and further improvements of the cluster pipeline still need to be done before making the most of the very large LSST cluster sample and the high-quality weak lensing data.

[1] https://github.com/LSSTDESC/firecrown

7.2.2 Cluster abundance likelihood

The correct description of cluster abundance statistics is central in the context of the future wide surveys of galaxy clusters such as Rubin or *Euclid*. As explored in Chapter 4, the cluster scaling relation is one of the main limiting factors when inferring cosmological parameters and one of the biggest sources of uncertainty.

Likelihood robustness test with scaling relation: The Chapter 5 was dedicated to test the robustness of widely used cluster likelihoods in the abundance regime that is aimed by the future surveys. We developed a methodology only accounting for the simplest inference setup, namely binned statistics in intervals of mass and redshift. We aim to study how scaling relation challenges the likelihood robustness, defined by the equivalence between the individual posterior variance to the variance of posterior means. This would imply a higher computational cost that should be carefully investigated.

Testing the accuracy of cluster likelihoods with Likelihood-Free Inference: I am also very interested in testing cluster abundance likelihoods using *Likelihood-Free Inference* (LFI) techniques (see e.g. Lemos et al. (2023)). The advantage of LFI is that we can reconstruct the parameter posterior even if the data likelihood is intractable. In a nutshell (i) we draw samples of cosmological parameters $\vec{\theta}_k \sim \pi$ (π is the parameter prior) (ii) for each cosmology $\vec{\theta}_k$, we perform $N_{\rm sim}$ cosmological simulations and we extract $N_{\rm sim}$ binned galaxy cluster catalogs in mass and redshift intervals (as it was done 1,000 times with the PINOCCHIO algorithm, but in this case, it was made for a single *Planck* cosmology) (iii) from the simulated counts for each fixed cosmology, we can reconstruct the joint distribution of the counts and cosmological parameters $f(\{\widehat{N}_{ij}^k\}, \vec{\theta})$. For an observed set of counts $\{\widehat{N}_{ij}^{\rm obs}\}$ (that we may observe in the Universe from a cluster survey), the posterior distribution of cosmological parameters is given by

$$\mathcal{P}(\vec{\theta}|\{\widehat{N}_{ij}^{\rm obs}\}) = \frac{f(\{\widehat{N}_{ij}^{\rm obs}\}, \vec{\theta})}{\mathcal{L}(\{\widehat{N}_{ij}^{\rm obs}\})}, \tag{7.1}$$

where $\mathcal{L}(\{\widehat{N}_{ij}^{\rm obs}\})$ is the evidence. We see that there is no need to compute the likelihood $\mathcal{L}(\{\widehat{N}_{ij}^{\rm obs}\}|\vec{\theta})$, whose information is fully encoded in $f(\{\widehat{N}_{ij}^{\rm obs}\}, \vec{\theta})$.

I show a simple example in Fig. 7.1, where I used 3,000 Ω_m parameters drawn from a flat prior. For each Ω_m, I sampled 500 cluster abundances in a single mass-redshift bin.[2] Here I use a likelihood to sample the cluster counts at fixed Ω_m, but LFI replaces this step with cosmological simulations (as likelihood-"free" refers to). For three different observed counts (here respectively 300, 500 and 700) I can reconstruct the parameter posterior (respectively the blue, orange and red points).

LFI techniques enable to draw the parameter posterior without any assumption of the "latent" likelihood, only relying on the forward modeling of cluster abundance through simulations. Thus, LFI accounts for complex physical processes beyond the Poisson, Gaussian or GPC formalism. It was already explored to derive cosmological constraints from the abundance of clusters (Ishida et al., 2015; Tam et al., 2022; Reza et al., 2022). This procedure is particularly costly since it requires a large quantity of cosmological simulations at different cosmologies. Then, we can propose a test of the cluster abundance likelihood approximations (Poisson, Gaussian, GPC), by drawing the most accurate posterior distribution from LFI and comparing to the "analysis" posteriors (obtained with the Poisson, the Gaussian or the GPC likelihood). This

[2] I used a simple model $\langle\widehat{N}|\Omega_m\rangle = N_0(\Omega_m/0.3)^{2.9}$ from Bolliet et al. (2020) (black dashed line in Fig. 7.1) and each cluster count $\widehat{N}$ at fixed Ω_m is sampled using a Gaussian likelihood that accounts for Poisson noise only.

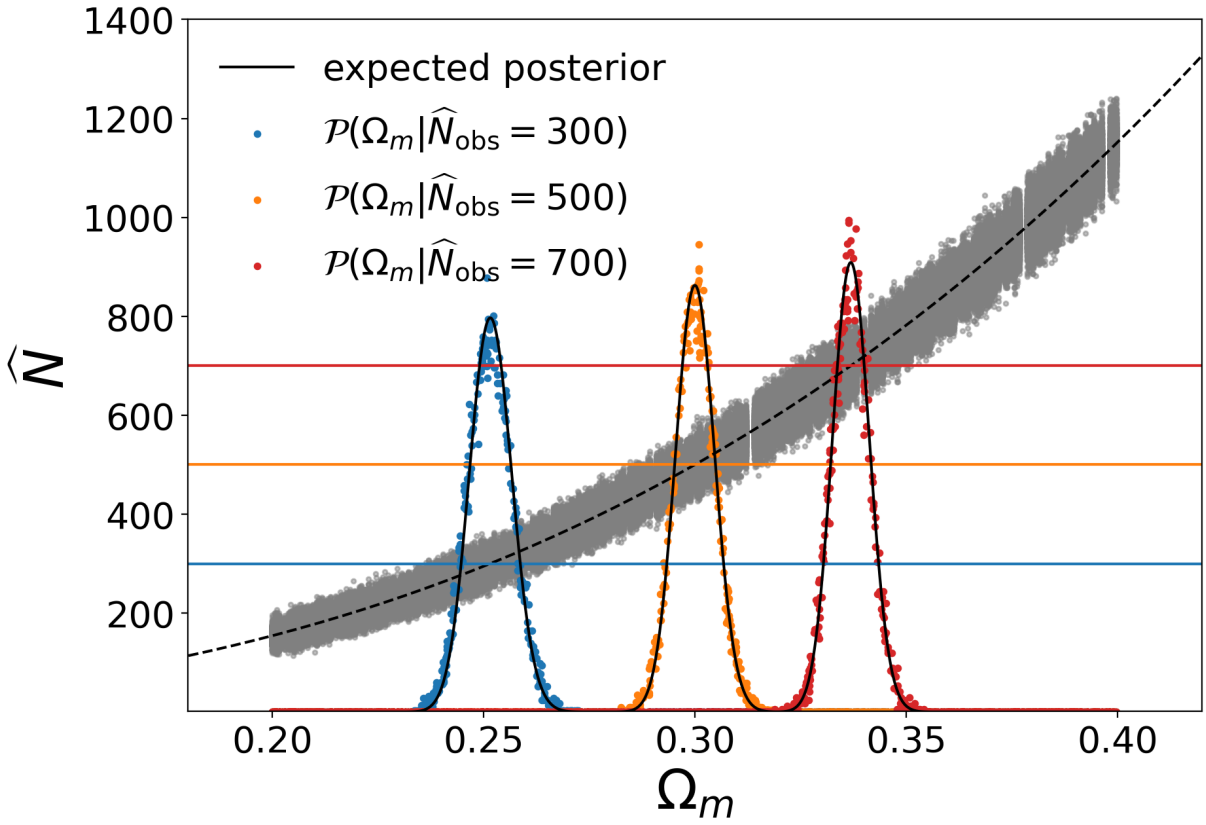

Figure 7.1: Joint distribution of counts $\widehat{N}$ and Ω_m (grey points), $\langle \widehat{N}|\Omega_m \rangle$ model in black dashed line. Reconstructed posterior distribution for the two observed counts 300 (blue points), 500 (orange points) and 700 (red points). The expected posterior distributions are represented in black full lines.

methodology is significantly different from what we did in Section 5.1 since it provides the most accurate posterior distribution without assuming a "latent" likelihood (in Section 5.1 we measured the latent likelihood at fiducial cosmology, not the true posterior). The methodology is however limited by the number of available simulations and the difficulty in simulating realistic observations.

Unbinned cluster count likelihood with SSC: In Section 5.2 we have explored a new formalism to account for the Super-Sample Covariance in the cluster count unbinned regime. We aim to study the impact on cosmological constraints for various other inference setups, and we aim to publish this work soon. This work in the unbinned regime could also be the place to discuss the inference of the individual covariance between the lensing properties of clusters and optical richnesses, that cannot be investigated easily in the binned statistics, and has a great impact on cosmological constraints. As we developed a robustness test for cluster abundance likelihood that is even applicable directly to the unbinned formalism, we aim to test the validity of such a new unbinned framework for the description of abundance statistics.

Inference from joint binned and unbinned cluster abundance: As seen in Chapter 5, the unbinned formalism for cluster abundance allows to probe locally the underlying mass and redshift distribution of clusters, and provides generally tighter cosmological constraints than the binned approach. However, it may be particularly difficult to compute in the context of

the very large upcoming *Euclid* and LSST cluster samples. In that context, the binned and unbinned framework can also benefit from each other, by performing joint analysis with two separate cluster samples, intuitively a large catalog of low mass clusters (binned) and a smaller one with massive clusters, reaching the tail of the halo mass function (unbinned). Thanks to the large statistical power of the binned statistics and the sensitivity of the unbinned formalism to individual cluster features, this will permit to combine different degeneracies and improve cosmological constraints.

7.2.3 Weak lensing mass calibration

Joint shear and magnification multipole analysis: The other lensing effect that I did not explore in this thesis is the magnification (see e.g. Tudorica et al. (2017)). In continuation to my work on the weak lensing mass reconstruction, I am interested in improving weak lensing mass inference from the joint analysis of lensing shear and magnification (Umetsu et al., 2016), that are both affected by very different systematic effects that can be more controlled when combined. For elliptical clusters, the multipole decomposition of the magnification effect is directly proportional to the multipole moment of the convergence field (since the derivation of the shear multipoles from convergence multipole is much more complicated), thus allowing us to access an even more direct probe of the projected halo shape.

Simulation-based inference of cluster mass: For our modeling choice of the elliptical projected density profile in the The300 project, the elongation/flatness of the cluster mass distribution is not taken into account, thus biasing the mass reconstruction when the major axis of a prolate/oblate cluster is aligned along the line-of-sight. We made preliminary studies accounting for this degree of freedom, and we found that the lensing mass is strongly correlated with the prolate/oblate index. This limitation of the elliptical halo modeling can be improved using machine learning techniques based on the reconstructed lensing multipoles for the several available orientations of The300 clusters. The cluster mass reconstruction from Simulation-Based Inference (SBI) and machine learning techniques was already explored for dynamical cluster mass by Ramanah et al. (2020).

The methodology could be the following:

- First, split the The300 cluster sample in half (one training and one test dataset). For each cluster of the training sample, measure the monopoles of the cluster lensing shear from a realistic background galaxy catalog, thus accounting for the intrinsic source of scatter from galaxies (shape noise and shot noise).

- Train a neural network to map between the observed lensing multipoles and the posterior distribution of the simulated mass, that turns out to be a Dirac distribution (and even other simulated quantities such as the "true" halo ellipticity). For a single The300 cluster (with unique mass), the several projections will lead to different multipoles as we have seen in Chapter 6.

- The second cluster sample can be used to test the performance of the neural network regression. When using the neural network to recover the posterior distribution of an almost spherical object in the sky (i.e., no multipoles), we expect the recovered error on the lensing mass to be larger than a projected elliptical mass distribution, since it accounts for all possible elongation/flatness degrees that are present in the training sample.

7.2.4 Synergies between experiments and other cosmological probes

The upcoming stage IV surveys will provide a large amount of data. Synergies between them will allow us to make the most of these unprecedented statistics, thanks to the large overlap between respective sky coverage.

The Rubin LSST and the *Euclid* space mission: For instance, the *Euclid* space mission[3] has optimal wide-field optical photometry, and near-infrared photometry and spectroscopy to observe large sample of galaxy clusters up to redshift $z \approx 2$, i.e. to larger redshifts than what Rubin LSST will provide ($z \approx 1.2$). In a synergistic point of view (Rhodes et al., 2017), robust photo-z and shear calibration will also be achieved by combining the *Euclid* data with LSST photometry. Approximately 50% of LSST galaxies will be blended that may lead to several biases at a per-cent level for weak-lensing shear measurements. To combine the LSST data with *Euclid* high-resolution space-based images will help improve the deblending of objects. In a cluster perspective, the synergies between *Euclid* and LSST will improve cluster detection, pushing cluster analyses to higher redshifts, and enable accurate cluster mass reconstruction from weak lensing (better separation between lensed and foreground galaxies).

Combining probes and surveys: It is known that better constraints can be obtained by combining different cosmological probes. Combining (meaning joint analysis and/or cross-correlating datasets) has the great advantage of beating down systematic effects (if each individual dataset suffers from independent and/or uncorrelated noises), also to have a more constraining power through new estimators (especially when cross-correlating probes).

For instance, the DES analysis in Abbott et al. (2022) combined and cross-correlated the galaxy weak lensing and clustering, to better calibrate galaxy biases and get more accurate cosmological constraints compared to each individual result.

In a cluster perspective, To et al. (2021) combined the DES cluster sample with the DES galaxy catalog in a full "6×2pt+N" analysis, and that has improved both constraints of cosmological and cluster scaling relation parameters, compared to cluster-alone (abundance and lensing) analyses. Moreover, the combination of different surveys will permit to combine multi-wavelength observations of galaxy clusters. Over the past decades, several cluster analyses have benefited from these multi-wavelength follow-ups to improve cosmological constraints. For instance, the SPT cluster analysis (Bocquet et al., 2019) combined X-rays and lensing data to improve the calibration of the cluster scaling relation. In the next decades, the joint analysis of multi-wavelength observations of galaxy clusters in optics (LSST, *Euclid*), X-rays (*eROSITA*) and at millimeter wavelengths (Simons Observatory, CMB-S4) will permit to address each independent systematic effects (hydro-static bias, modeling of the mass distribution, cluster scaling relation).

In the next decade, large CMB experiments will measure with high precision the LSS distribution in the Universe. CMB observables are also highly correlated to LSS probed by galaxy surveys, for instance through the gravitational lensing of the CMB anisotropies and polarization. Science from CMB cross-correlations with LSS tracers permits to map in a unique way the spatial distribution and dynamics of the gravitational potentials in the Universe. Many studies[4] have been carried out to gain cosmological insights from the cross-correlation between CMB data and LSS surveys. With the upcoming powerful datasets such as the Rubin LSST, the *Euclid*, the Simons Observatory and CMB-S4 (and their large overlap), CMB-LSS cross-correlations will enable us to extract more information and reach smaller statistical error bars on cosmological

[3] *Euclid* was successfully launched on 1st July 2023 (when finishing this writing).

[4] For CMB lensing-galaxy clustering, see e.g. Saraf et al. (2022), for CMB lensing-cosmic shear, see e.g. Namikawa et al. (2019), for CMB lensing-thermal SZ, see e.g. Hill & Spergel (2014).

parameters.

7.2.5 Others

From this experience in weak gravitational lensing physics by galaxy clusters, I realized that the derivations of some of the widely known equations of the field could not be found in the existing literature. I am currently finishing a review document explicitly showing how to obtain those equations from the theory of weak lensing to lensing in galaxy cluster fields as I believe this could be useful for new scientists entering the field. I presented a preliminary version of the document to the DESC cluster working group and gave a 7-hour-long course on the topic to members of my lab[5]. Once finalized, the document will be accessible on Arxiv.

[5]The lecture slides are available at `https://www.researchgate.net/publication/354238822_Weak_gravitational_lensing_by_galaxy_clusters?channel=doi&linkId=612e0dfc0360302a006cc1f9&showFulltext=true`.

Bibliography

Abazajian K. N., et al., 2016, arXiv e-prints, p. arXiv:1610.02743

Abbott T., et al., 2020, Phys. Rev. D, 102, 023509

Abbott T., et al., 2022, Phys. Rev. D, 105, 023520

Abdalla E., et al., 2022, J. High Energy Phys., 34, 49

Abdullah M. H., Klypin A., Wilson G., 2020, ApJ, 901, 90

Abdullah M. H., Wilson G., Klypin A., Ishiyama T., 2023, ApJ, 955, 26

Abolfathi B., et al., 2021, ApJS, 253, 31

Ade P., et al., 2014, A&A, 571, A20

Ade P., et al., 2016, A&A, 594, A24

Ade P., et al., 2019, J. Cosmology Astropart. Phys., 2019, 056

Adhikari S., Chue C. Y. R., Dalal N., 2015, J. Cosmology Astropart. Phys., 2015, 009

Adler R. J., Casey B., Jacob O. C., 1995, Am. J. Phys., 63, 620

Aghanim N., Majumdar S., Silk J., 2008, Rep. Prog. Phys., 71, 066902

Aguena M., Lima M., 2018, Phys. Rev. D, 98, 123529

Aguena M., et al., 2021a, MNRAS, 502, 4435

Aguena M., et al., 2021b, MNRAS, 508, 6092

Aguena M., et al., 2023, arXiv e-prints, p. arXiv:2309.06593

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

Aitchison J., Ho C. H., 1989, Biometrika, 76, 643–653

Akeson R., et al., 2019, arXiv e-prints, p. arXiv:1902.05569

Alam S., et al., 2021, Phys. Rev. D, 103

Albrecht A., Steinhardt P. J., 1982, Phys. Rev. Lett., 48, 1220

Alcock C., Paczynski B., 1979, Nature, 281, 358

Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, MNRAS, 383, 879

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Arnaud M., Pointecouteau E., Pratt G. W., 2005, A&A, 441, 893

Artis E., Melin J.-B., Bartlett J. G., Murray C., 2021, A&A, 649, A47

Artis E., Melin J. B., Bartlett J. G., Murray C., Euclid Consortium 2022, in mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera. (arXiv:2111.05432)

Astropy Collaboration et al., 2022, ApJ, 935, 167

Bagla J. S., 2005, Curr. Sci., 88, 1088

Baker T., 2017, Weak Gravitational Lensing (lecture)

Bardeen J. M., 1980, Phys. Rev. D, 22, 1882

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291–472

Bartlett J. G., 1997, in Valls-Gabaud D., Hendry M. A., Molaro P., Chamcham K., eds, Astronomical Society of the Pacific Conference Series Vol. 126, From Quantum Fluctuations to Cosmological Structures. p. 365 (arXiv:astro-ph/9703090), doi:10.48550/arXiv.astro-ph/9703090

Bassett B., Hlozek R., 2010, in Ruiz-Lapuente P., ed., , Dark Energy: Observational and Theoretical Approaches. p. 246, doi:10.48550/arXiv.0910.5224

Baumann D., 2009, arXiv e-prints, p. arXiv:0907.5424

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

Benítez N., 2011, ASCL, p. ascl:1108.011

Bennett C. L., et al., 2013, ApJS, 208, 20

Benson A. J., 2012, New A, 17, 175

Bernstein G. M., Nakajima R., 2009, ApJ, 693, 1508

Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122

Bhattacharya S., Habib S., Heitmann K., Vikhlinin A., 2013, ApJ, 766, 32

Binggeli B., 1982, A&A, 107, 338

Bocquet S., Saro A., Dolag K., Mohr J. J., 2015, MNRAS, 456, 2361

Bocquet S., et al., 2019, ApJ, 878, 55

Bocquet S., et al., 2023, arXiv e-prints, p. arXiv:2310.12213

Böhringer H., Chon G., 2016, Mod. Phys. Lett. A, 31, 1640008

Bolliet B., Brinckmann T., Chluba J., Lesgourgues J., 2020, MNRAS, 497, 1332

Boltzmann 1896, Vorlesungen über Gastheorie. Leipzig, J. A. Barthy

Bond J. R., Myers S. T., 1996, ApJS, 103, 1

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, Astrophys. J., 379, 440

Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, A&A, 296, 575

Buchert T., 1992, MNRAS, 254, 729

Buote D. A., Lewis A. D., 2004, ApJ, 604, 116

Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARA&A, 40, 643

Cataneo M., et al., 2015, Phys. Rev. D, 92, 044009

Catelan P., Kamionkowski M., Blandford R. D., 2001, MNRAS, 320, L7

Cavaliere A. G., Gursky H., Tucker W. H., 1971, Nature, 231, 437

Challinor A., 2012, Proceedings of the International Astronomical Union, 8, 42

Chang C., et al., 2013, MNRAS, 434, 2121–2135

Chaubal P. S., et al., 2022, ApJ, 931, 139

Chevallier M., Polarski D., 2001, Int. J. Mod. Phys. D, 10, 213

Chisari N. E., et al., 2019, ApJS, 242, 2

Chiu I. N., Klein M., Mohr J., Bocquet S., 2023, MNRAS, 522, 1601

Clampitt J., Jain B., 2016, MNRAS, 457, 4135

Clowe D., Lucia G. D., King L., 2004, MNRAS, 350, 1038

Cohen A. G., Rú jula A. D., Glashow S. L., 1998, ApJ, 495, 539

Cooray A., Sheth R., 2002, Phys. Rep., 372, 1

Corless V., King L., 2007, MNRAS, 380, 149

Correa C. M., 2023, BAAA, 64, 159

Costanzi M., et al., 2019, MNRAS, 488, 4779

Costanzi M., et al., 2021, Phys. Rev. D, 103, 043522

Cramer R., 1952

Cromer D., Battaglia N., Miyatake H., Simet M., 2022, J. Cosmology Astropart. Phys., 2022, 034

Cui W., et al., 2018, MNRAS, 480, 2898

DeRose J., et al., 2022, Phys. Rev. D, 105, 123520

Denzel P., Coles J. P., Saha P., Williams L. L. R., 2021, MNRAS, 501, 784

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

Despali G., Giocoli C., Tormen G., 2014, MNRAS, 443, 3208

Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2015, MNRAS, 456, 2486

Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1

Dodelson S., Schneider M. D., 2013, Phys. Rev. D, 88, 063537

Douspis M., Salvati L., Aghanim N., 2019, arXiv e-prints, p. arXiv:1901.05289

Doux C., Ganga K., Aubourg E., 2017, PhD thesis, Paris Diderot University

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, L64

Efron B., 2015, J. R. Stat. Soc., 77, 617

Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

Einstein A., 1916, Annalen Phys., 49, 769

Eisenstein D. J., Loeb A., 1995, ApJ, 439, 520

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Elvira V., Martino L., 2021, arXiv e-prints, p. arXiv:2102.05407

Escoffier S., et al., 2016, arXiv e-prints, p. arXiv:1606.00233

Ettori S., Morandi A., Tozzi P., Balestra I., Borgani S., Rosati P., Lovisari L., Terenziani F., 2009, A&A, 501, 61

Ettori S., Donnarumma A., Pointecouteau E., Reiprich T. H., Giodini S., Lovisari L., Schmidt R. W., 2013, Space Sci. Rev., 177, 119

Evrard A. E., et al., 2002, ApJ, 573, 7

Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562

Fedeli C., Carbone C., Moscardini L., Cimatti A., 2011, MNRAS, 414, 1545

Feng J. L., 2010, ARA&A, 48, 495

Feroz F., Hobson M. P., 2011, MNRAS, 420, 596

Ferraro S., Schmidt F., Hu W., 2011, Phys. Rev. D, 83

Fisher R. A., 1935, J. R. Stat. Soc., 98, 39

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

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

Friedmann A., 1924, Zeitschrift fur Physik, 21, 326

Fumagalli A., et al., 2021, A&A, 652, A21

Fumagalli A., Costanzi M., Saro A., Castro T., Borgani S., 2023, arXiv e-prints, p. arXiv:2310.09146

Georgi H., Glashow S. L., 1974, Phys. Rev. Lett., 32, 438

Giocoli C., Meneghetti M., Metcalf R. B., Ettori S., Moscardini L., 2014, MNRAS, 440, 1899

Glashow S. L., 1959, Nucl. Phys., 10, 107

Gonzalez E. J., Makler M., García Lambas D., Chalela M., Pereira M. E. S., Van Waerbeke L., Shan H., Erben T., 2021, MNRAS, 501, 5239

Gouin C., Gavazzi R., Codis S., Pichon C., Peirani S., Dubois Y., 2017, A&A, 605, A27

Gouyou Beauchamps S., Lacasa F., Tutusaus I., Aubert M., Baratta P., Gorce A., Sakr Z., 2022, A&A, 659, A128

Gruen D., Seitz S., Becker M. R., Friedrich O., Mana A., 2015, MNRAS, 449, 4264

Guth A. H., 1981, Phys. Rev. D, 23, 347

Hamilton A. J. S., 1998, in Hamilton D., ed., Astrophysics and Space Science Library Vol. 231, The Evolving Universe. p. 185 (arXiv:astro-ph/9708102), doi:10.1007/978-94-011-4960-0˙17

Hasselfield M., et al., 2013, J. Cosmology Astropart. Phys., 2013, 008

Hearin A., Korytov D., Kovacs E., Benson A., Aung H., Bradshaw C., Campbell D., LSST Dark Energy Science Collaboration 2020, MNRAS, 495, 5040

Heavens A. F., Seikel M., Nord B. D., Aich M., Bouffanais Y., Bassett B. A., Hobson M. P., 2014, MNRAS, 445, 1687

Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, ApJ, 780, 111

Heitmann K., et al., 2019, ApJS, 245, 16

Herbonnet R., et al., 2022, MNRAS, 513, 2178

Hernquist L., 1990, ApJ, 356, 359

Heymans C., et al., 2012, MNRAS, 427, 146

Heymans C., et al., 2021, A&A, 646, A140

Hill J. C., Spergel D. N., 2014, J. Cosmology Astropart. Phys., 2014, 030

Hirata C., Seljak U., 2003, MNRAS, 343, 459

Hoekstra H., 2003, MNRAS, 339, 1155

Hoffman Y., Shaham J., 1985, ApJ, 297, 16

Hogg D. W., 1999, arXiv e-prints, pp astro–ph/9905116

Howell D. A., 2011, Nature Communications, 2, 350

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Huterer D., Takada M., 2005, Astroparticle Physics, 23, 369

Huterer D., et al., 2015, Astropart. Phys., 63, 23

Ishida E. E. O., et al., 2015, Astron. Comput., 13, 1

Iso K., Kodama H., Sato K., 1986, Phys. Rev. B, 169, 337

Izbicki R., Lee A. B., 2017, arXiv e-prints, p. arXiv:1704.08095

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Jeong D., Komatsu E., Jain B., 2009, Phys. Rev. D, 80

Jing Y. P., Suto Y., 2002, ApJ, 574, 538

Johnston D. E., et al., 2007, arXiv e-prints, p. arXiv:0709.1159

Johnstone I. M., 2015, Inst. Math. Stat., 6

Kaiser N., Squires G., 1993, ApJ, 404, 441

Kamionkowski M., Kovetz E. D., 2016, ARA&A, 54, 227

Kéruzoré F., et al., 2020, A&A, 644, A93

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Knebe A., Wießner V., 2006, PASA, 23, 125

Komatsu E., et al., 2011, ApJS, 192, 18

Korytov D., et al., 2019, ApJS, 245, 26

Kovacs E., et al., 2022, OJAp, 5, 1

Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353

Ku H. H., 1996, NIST, 70, 263

Kurki-Suonio H., 2021, Phys. Rep.

Kurki-Suonio H., 2022, Phys. Rep.

LSST Science Collaboration et al., 2009, arXiv e-prints, p. arXiv:0912.0201

Lacasa F., Lima M., Aguena M., 2018, A&A, 611, A83

Lau E. T., Nagai D., Kravtsov A. V., Vikhlinin A., Zentner A. R., 2012, ApJ, 755, 116

Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193

Lee J., 2012, ApJ, 752, 40

Lee B. E., Le Brun A. M. C., Haq M. E., Deering N. J., King L. J., Applegate D., McCarthy I. G., 2018, MNRAS, 479, 890

Lemaître G., 1931, MNRAS, 91, 483

Lemos P., Cranmer M., Abidi M., Hahn C., Eickenberg M., Massara E., Yallup D., Ho S., 2023, Mach. Learn.: Sci. Technol., 4, 01LT01

Lesci G. F., et al., 2022, A&A, 659, A88

Lesgourgues J., 2011, arXiv e-prints, p. arXiv:1104.2932

Levi M., et al., 2019, in Bulletin of the American Astronomical Society. p. 57 (arXiv:1907.10688), doi:10.48550/arXiv.1907.10688

Lewis A., Challinor A., 2006, Phys. Rep., 429, 1

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Lifshitz E., 1946, J. Phys. (USSR), 10, 116

Lima M., Hu W., 2004, Phys. Rev. D, 70, 043504

Linde A. D., 1982, Phys. Rev. B, 108, 389

Linder E. V., 2003, Phys. Rev. Lett., 90, 091301

Linder E. V., Jenkins A., 2003, MNRAS, 346, 573

LiteBIRD Collaboration et al., 2023, PTEP, 2023, 042F01

Ma C.-P., Bertschinger E., 1995, ApJ, 455, 7

Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116

Mana A., Giannantonio T., Weller J., Hoyle B., Hütsi G., Sartoris B., 2013, MNRAS, 434, 684

Mandelbaum R., et al., 2005, MNRAS, 361, 1287

Mandelbaum R., Seljak U., Cool R. J., Blanton M., Hirata C. M., Brinkmann J., 2006, MNRAS, 372, 758

Mandelbaum R., et al., 2014, ApJS, 212, 5

Mandelbaum R., et al., 2018, MNRAS, 481, 3170

Mantz A. B., et al., 2015, MNRAS, 446, 2205

Mao Y.-Y., et al., 2018, ApJS, 234, 36

Mao Y.-Y., Geha M., Wechsler R. H., Weiner B., Tollerud E. J., Nadler E. O., Kallivayalil N., 2021, ApJ, 907, 85

Marulli F., Veropalumbo A., García-Farieta J. E., Moresco M., Moscardini L., Cimatti A., 2021, ApJ, 920, 13

Maughan B. J., 2007, ApJ, 668, 772

McClintock T., et al., 2019, MNRAS, 482, 1352

McDonald P., Trac H., Contaldi C., 2006, MNRAS, 366, 547

Melchior P., et al., 2017, MNRAS, 469, 4899–4920

Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93

Meneghetti M., et al., 2020, Science, 369, 1347

Milgrom M., 1983, ApJ, 270, 365

Monaco P., 1997, MNRAS, 287, 753

Monaco P., 2016, Galaxies, 4, 53

Monaco P., Theuns T., Taffoni G., 2002, MNRAS, 331, 587

Moutarde F., Alimi J. M., Bouchet F. R., Pellat R., Ramani A., 1991, ApJ, 382, 377

Muñoz-Echeverría M., et al., 2023, arXiv e-prints, p. arXiv:2305.14862

Mukhanov V. F., Feldman H. A., Brandenberger R. H., 1992, Phys. Rep., 215, 203

Munari E., Biviano A., Borgani S., Murante G., Fabjan D., 2013, MNRAS, 430, 2638

Munari E., Monaco P., Sefusatti E., Castorina E., Mohammad F. G., Anselmi S., Borgani S., 2017, MNRAS, 465, 4658

Muñoz-Echeverría M., et al., 2023, A&A, 671, A28

Murata R., et al., 2019, PASJ, 71, 107

Murray C., Bartlett J. G., Artis E., Melin J.-B., 2022, MNRAS, 512, 4785

Namikawa T., et al., 2019, ApJ, 882, 62

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Newman J. A., et al., 2013, ApJS, 208, 5

Oguri M., Keeton C. R., 2004, ApJ, 610, 663

Oguri M., Takada M., 2011, Phys. Rev. D, 83

Oguri M., Lee J., Suto Y., 2003, ApJ, 599, 7

Oguri M., Takada M., Okabe N., Smith G. P., 2010, MNRAS

Pacaud F., et al., 2018, A&A, 620, A10

Park Y., Sunayama T., Takada M., Kobayashi Y., Miyatake H., More S., Nishimichi T., Sugiyama S., 2023, MNRAS, 518, 5171

Parroni C., et al., 2017, ApJ, 848, 114

Payerne C., Murray C., Combet C., Doux C., Fumagalli A., Penna-Lima M., 2023, MNRAS, 520, 6223

Peacock J. A., 1999, Cosmological Physics. Cambridge University Press

Peebles P. J. E., 1980, The large-scale structure of the universe

Penna-Lima M., 2010, Centro Brasileiro de Pesquisas Físicas

Penna-Lima M., Makler M., Wuensche C., 2014, J. Cosmology Astropart. Phys., 2014, 039–039

Percival W. J., et al., 2010, MNRAS, 401, 2148

Percival W. J., Friedrich O., Sellentin E., Heavens A., 2022, MNRAS, 510, 3207

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

Peter P., Uzan J.-P., 2013, Primordial Cosmology. Oxford Graduate Texts, Oxford University Press

Phriksee A., Jullo E., Limousin M., Shan H., Finoguenov A., Komonjinda S., Wannawichian S., Sawangwit U., 2020, MNRAS, 491, 1643

Piattella O., 2018

Planck Collaboration et al., 2014, A&A, 571, A16

Planck Collaboration et al., 2016, A&A, 594, A13

Planck Collaboration et al., 2020a, A&A, 641, A6

Planck Collaboration et al., 2020b, A&A, 641, A7

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, Space Sci. Rev., 215, 25

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

Ramanah D. K., Wojtak R., Arendse N., 2020, MNRAS, 501, 4080

Rasia E., et al., 2015, ApJ, 813, L17

Raveri M., Hu W., 2019, Phys. Rev. D, 99

Reza M., Zhang Y., Nord B., Poh J., Ciprijanovic A., Strigari L., 2022, in Machine Learning for Astrophysics. p. 20 (`arXiv:2208.00134`), doi:10.48550/arXiv.2208.00134

Rhodes J., et al., 2017, ApJS, 233, 21

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

Riess A. G., Casertano S., Yuan W., Bowers J. B., Macri L., Zinn J. C., Scolnic D., 2021, ApJ Lett., 908, L6

Robertson H. P., 1935, ApJ, 82, 284

Robinson J., Gawiser E., Silk J., 2000, ApJ, 532, 1

Robison B., et al., 2023, MNRAS, 523, 1614

Rovelli C., 2023, Anaximander and the Birth of Science. Riverhead Books

Rozo E., et al., 2009, ApJ, 699, 768

Rozo E., et al., 2010, ApJ, 708, 645

Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ, 238, 471

Ruhl J., et al., 2004, in Bradford C. M., et al., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 5498, Z-Spec: a broadband millimeter-wave grating spectrometer: design, construction, and first cryogenic measurements. pp 11–29 (arXiv:astro-ph/0411122), doi:10.1117/12.552473

Rykoff E. S., et al., 2014, ApJ, 785, 104

Salam A., 1968, Conf. Proc. C, 680519, 367

Salvati L., et al., 2022, ApJ, 934, 129

Sanderson A. J. R., Ponman T. J., 2009, MNRAS, 402, 65

Saraf C. S., Bielewicz P., Chodorowski M., 2022, MNRAS, 515, 1993

Saro A., et al., 2015, MNRAS, 454, 2305

Schneider P., 2006, in , Saas-Fee Advanced Courses. Springer Berlin Heidelberg, pp 269–451

Schneider P., Weiss A., 1991, A&A, 247, 269

Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses, doi:10.1007/978-3-662-03758-4.

Schneider M. D., Frenk C. S., Cole S., 2012, J. Cosmology Astropart. Phys., 2012, 030

Schutz B. F., 1985, A First Course In General Relativity

Sehgal N., et al., 2011, ApJ, 732, 44

Sellentin E., Heavens A. F., 2017, MNRAS, 464, 4658

Sellentin E., Heavens A. F., 2018, MNRAS, 473, 2355

Sellentin E., Quartin M., Amendola L., 2014, MNRAS, 441, 1831

Sembolini F., Petris M. D., Yepes G., Foschi E., Lamagna L., Gottlöber S., 2014, MNRAS, 440, 3520

Sheldon E. S., Huff E. M., 2017, ApJ, 841, 24

Sheldon E. S., et al., 2004, AJ, 127, 2544

Sheldon E. S., Becker M. R., Jarvis M., Armstrong R., LSST Dark Energy Science Collaboration 2023, OJAp, 6, 17

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Shin T.-h., Clampitt J., Jain B., Bernstein G., Neil A., Rozo E., Rykoff E., 2018, MNRAS, 475, 2421

Shiralilou B., Martinelli M., Papadomanolakis G., Peirone S., Renzi F., Silvestri A., 2020, J. Cosmology Astropart. Phys., 2020, 057

Shirasaki M., Takada M., 2018, MNRAS, 478, 4277

Shirasaki M., Takada M., Miyatake H., Takahashi R., Hamana T., Nishimichi T., Murata R., 2017, MNRAS, 470, 3476

Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2017, MNRAS, 466, 3103

Smith R. E., et al., 2003, MNRAS, 341, 1311

Springel V., et al., 2018, MNRAS, 475, 676

Stapelberg S., Tchernin C., Hug D., Lau E. T., Bartelmann M., 2022, A&A, 663, A17

Sunayama T., et al., 2023, arXiv e-prints, p. arXiv:2309.13025

Suzuki N., et al., 2012, ApJ, 746, 85

Takada M., Spergel D. N., 2014, MNRAS, 441, 2456–2475

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152

Tam S.-I., Umetsu K., Amara A., 2022, ApJ, 925, 145

Taylor A., Joachimi B., 2014, MNRAS, 442, 2728

Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22

The Dark Energy Survey Collaboration 2005, arXiv e-prints, pp astro–ph/0510346

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

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

To C., et al., 2021, Phys. Rev. Lett., 126, 141301

Trindade A. M. M., Avelino P. P., Viana P. T. P., 2012, MNRAS, 424, 1442

Tudorica A., et al., 2017, A&A, 608, A141

Turner M. S., 2022, ARNPS, 72, 1

Umetsu K., 2020, A&A Rev., 28

Umetsu K., Zitrin A., Gruen D., Merten J., Donahue M., Postman M., 2016, ApJ, 821, 116

Valentino E. D., et al., 2021, Classical and Quantum Gravity, 38, 153001

Varga T. N., et al., 2019, MNRAS, 489, 2511–2524

Verde L., 2010, in Wolschin G., ed., , Vol. 800, Lecture Notes in Physics, Berlin Springer Verlag. pp 147–177, doi:10.1007/978-3-642-10598-2˙4

Vogelsberger M., et al., 2014, Nature, 509, 177

Walker A. G., 1937, Proceedings of the London Mathematical Society, 42, 90

Wang L., Steinhardt P. J., 1998, ApJ, 508, 483

Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, Nature, 585, 39

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230

Weinberg S., 1967, Phys. Rev. Lett., 19, 1264

Weinberg S., 1989, Rev. Mod. Phys., 61, 1

Weinberg D. H., Davé R., Katz N., Kollmeier J. A., 2003, in Holt S. H., Reynolds C. S., eds, American Institute of Physics Conference Series Vol. 666, The Emergence of Cosmic Structure. pp 157–169 (`arXiv:astro-ph/0301186`), doi:10.1063/1.1581786

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

Weltman A., et al., 2020, PASA, 37

White S. D. M., 1984, ApJ, 286, 38

Wolz L., Kilbinger M., Weller J., Giannantonio T., 2012, J. Cosmology Astropart. Phys., 2012, 009

Wu H.-Y., Weinberg D. H., Salcedo A. N., Wibking B. D., Zu Y., 2019, MNRAS, 490, 2606

Wu H.-Y., et al., 2022, MNRAS, 515, 4471

Zeldovich Sunyaev 1969, Ap&SS, 4, 301

Zhang Z., et al., 2023, MNRAS, 523, 1994

Zwicky F., 1937, ApJ, 86, 217

de Haan T., et al., 2016, ApJ, 832, 95

van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649

van Uitert E., et al., 2017, MNRAS, 467, 4131

van Weeren R. J., de Gasperin F., Akamatsu H., Brüggen M., Feretti L., Kang H., Stroe A., Zandanel F., 2019, Space Sci. Rev., 215, 16

Appendix A

Perturbation theory in general relativity

This appendix provides more calculation details about the cosmological perturbation theory in general relativity, introduced in Chapter 1.

The *classical* theory of cosmological perturbations introduced in Section 1.3.2 is valid when the typical spatial extent of density variations is much smaller than the Hubble length $1/H$, as well as for weak gravitational fields. For such scales, the relativistic effects of the metric fluctuations can be ignored and most of the essential physics can be extracted from a Newtonian approach. Moreover, the *Newtonian* description of the gravitational field Φ acting as an external "force" is still inherent to the classical theory of perturbations, whereas general relativity tells us that gravitation is the fact of metric perturbation. For more accurate results, the theory of perturbations needs to be treated from Einstein's general relativity directly (Mukhanov et al., 1992). According to EFE in Eq. (1.4), any density perturbation now perturbs the geometry of space-time itself. We can now write the perturbed energy-momentum tensor $T_{\mu\nu}$ and metric tensor $g_{\mu\nu}$ as

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \quad \text{and} \quad T_{\mu\nu} = \bar{T}_{\mu\nu} + \delta T_{\mu\nu}, \tag{A.1}$$

where $\bar{g}_{\mu\nu}$ is the background space-time metric in the perturbed Universe. Both $T_{\mu\nu}$ and $G_{\mu\nu}$ in the EFE in Eq. (1.4) can be expanded the same way order-by-order. The zero and first-order equations

$$\bar{G}_{\mu\nu} = \frac{8\pi G}{c^4}\bar{T}_{\mu\nu} \quad \text{and} \quad \delta G_{\mu\nu} = \frac{8\pi G}{c^4}\delta T_{\mu\nu}. \tag{A.2}$$

represent respectively the motion of the background and perturbed Universe. The first leads to the Friedmann equations. The second now describes how both metric and energy perturbations are coupled and evolve. From both sides of the second equation, we need to define the number of additional degrees of freedom. However, a problem occurs. The simplest and most intuitive solution for the background metric in our perturbed universe would be to choose the FLRW metric.

However, this *mapping* between two different backgrounds, namely the homogeneous and perturbed backgrounds, is arbitrary. Fixing a *gauge* consists in choosing an appropriate map between the two different backgrounds and choosing a specific framework to define perturbations. This problem is known as the *gauge problem* in cosmology. It can be solved by defining *gauge*-invariant quantities that do not vary under a change of coordinates (Bardeen, 1980), otherwise by fixing a gauge to remove some degrees of freedom and track how the fluctuations grow in this gauge. When fixing a gauge, the metric perturbation can be built upon the SVT (Scalar-Vector-Tensor) decomposition. The commonly used is the conformal *Newtonian gauge* (Ma & Bertschinger, 1995), that is a convenient gauge to deal with scalar perturbations (associated to

the gravitational potential), but others can be useful for other applications (e.g. *synchronous gauge* proposed by Lifshitz (1946)). The "perturbed" interval is given by (Peter & Uzan, 2013)

$$ds^2 = a(\eta)^2[-(1+2\Phi)d\eta^2 + (1+2\Psi)d\vec{x}^2], \tag{A.3}$$

where η is the conformal time defined with respect to time t by $a(\eta)d\eta = dt$. We consider the simplest case of a flat Universe, i.e. $k = 0$. We have introduced two scalar degrees of freedom Φ and Ψ respectively for curvature perturbation of constant space slices $d\vec{x}_i = 0$ and time slices $dt = 0$. In this gauge, tensor and vector modes are zero. We can now compute the tensor perturbation $\delta G_{\mu\nu}$ associated to metric perturbation. Second, we need to define energy-momentum tensor perturbation. The perturbed energy-momentum tensor is given by

$$\delta T_{00} = \rho a^2(\delta + 2\Phi), \tag{A.4}$$

$$\delta T_{0i} = -\rho a^2(\rho + p)v_i, \tag{A.5}$$

$$\delta T_{ij} = a^2 p\left[\left(\frac{\delta p}{p} - 2\Psi\right)\delta_{ij} + \pi_{ij}\right], \tag{A.6}$$

where π_{ij} is the an-isotropic stress tensor. This one is usually neglected, ensuring that the two scalar perturbations verify $\Phi = \Psi$. For the following derivations, we will keep this approximation. Plugged into Eq. (A.2), and using the local conservation equation of the energy-momentum tensor, one can derive the equations of motion for scalar modes coupled to density perturbations, leading to the *relativistic* analogous of classical Euler equations (we use the syntax from the lecture of Kurki-Suonio (2022))

$$\nabla^2\Phi = \frac{3}{2}\mathcal{H}^2[\delta + 3\mathcal{H}(1+w)v], \tag{A.7}$$

$$\Phi' + \mathcal{H}\Phi = \frac{3}{2}\mathcal{H}(1+w)v, \tag{A.8}$$

$$\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi = \frac{3}{2}\mathcal{H}\frac{\delta p}{\bar{\rho}}, \tag{A.9}$$

where we used the modified Hubble parameter $\mathcal{H} = aH$, and the pressure-density relation $w = \bar{p}/\bar{\rho}$. We also consider that the fluid velocity writes $v_i = -\nabla_i v$ with v as a scalar field, and we change the time derivative of a function f given by $\dot{f}$ to the conformal time derivative f'. Eq. (A.7) is the relativistic analogous to the classic Poisson equation in Eq. (1.36), where Φ is now associated to the gravitational potential. We notice that not only it is sourced by δ, but also by *relativistic* corrections that depend on the velocity field. With the manipulation of the energy-momentum conservation, we obtain the continuity equations

$$\delta' = (1+w)(\nabla^2 v + 3\Phi') + 3\mathcal{H}\left(w\delta - \frac{\delta p}{\bar{\rho}}\right), \tag{A.10}$$

$$v' = -\mathcal{H}(1-3w)v - \frac{w'}{1+w}v + \frac{\delta p}{\bar{\rho}(1+w)} + \Phi. \tag{A.11}$$

For adiabatic perturbation, we can solve the pressure perturbation δp in Eq. (A.9), and we can relate to the differential equation describing the evolution of the scalar mode Φ with respect to time and space such as

$$\Phi'' + 3(1+c_s^2)\mathcal{H}\Phi' + 3(c_s^2 - w)\Phi - c_s^2\nabla^2\Phi = 0, \tag{A.12}$$

where we used $\bar{p}' = c_s^2\bar{\rho}'$. We can now track the evolution of the scalar perturbation Φ for the simplest case, matter-dominated era. We can make the matter-domination approximation, i.e.

$\bar{p} = w = c_s = 0$ and apply the same methodology as before, i.e. expanding δ in Fourier space. We first solve the background expansion $\mathcal{H}$ for the specific case of matter-dominated era giving $3\mathcal{H}/2 = 6\eta^2$, then we can solve the motion equation for the potential. The η-dependent solution of Φ decays very rapidly, we find that the potential is constant in time for perturbations living in the flat matter-dominated universe and we note $\Phi(\vec{x}, \eta) = \Phi(\vec{x})$. The velocity is obtained from Eq. (A.8), using $\Phi' = 0$. Plugged into Eq. (A.7) and extracting the perturbation δ, we have finally that

$$\delta = -2\Phi + \frac{2}{3\mathcal{H}}\nabla^2\Phi \quad \text{giving} \quad \delta_k = -2\Phi_k - \frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \Phi_k. \tag{A.13}$$

This is the main result of this section. The "super-Hubble" scales correspond to $k \ll \mathcal{H}$. Converting to physical scales $k_{\rm phys} = k/a$, we have $k_{\rm phys} \ll H$. The perturbation follows $\delta_k = -2\Phi_k$, so the perturbations are fixed or *frozen*, there is no growth. However, when $k \gg \mathcal{H}$, then

$$\delta_k \approx -\frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \Phi_k \propto \eta^2 \propto t^{2/3}. \tag{A.14}$$

In the linearized GR formalism, we see that density perturbations begin to grow when they enter the horizon, which was not considered in the classical theory of perturbation. At lower scales, we find the same results obtained from classical perturbation theory in an expanding universe. Thus, for sub-horizon scales, the magnitude of perturbation at time t should be

$$\delta_k(t) = \begin{cases} (a/a_k)\delta_{k,\rm prim} & \text{if } a > a_k, \\ \delta_{k,\rm prim} & \text{if } a < a_k. \end{cases}, \tag{A.15}$$

where a_k is the scale factor at the time when the scale k enters the horizon, and $\delta_{k,\rm prim} = -2\Phi_k$ is the "primordial" density perturbation, that will be discussed in Section 1.3.4 (for this example we do not consider the Λ-dominated era occurring after the matter-dominated era, that will tend to slow down the growth of perturbations at all scales). Now, the cosmological fluid is filled with multiple fluids. The total energy-momentum tensor is the sum of individual contribution namely

$$T_{\mu\nu} = \sum_i T^{(i)}_{\mu\nu}. \tag{A.16}$$

In the presence of multi-component fluid, the perturbed components will evolve on a common background via the Friedmann equation. We can re-write the relativistic Poisson equation in Eq. (A.7) that now accounts for all possible density and velocity perturbations with label $i = 1, ..., N$. The local conservation of the individual energy-momentum tensor is *apriori* not verified since any species may transfer energy and moment to another, at both background and perturbed levels. In the case of non-interacting fluid components, we can write $\nabla^\mu T^{(i)}_{\mu\nu} = 0$ for all fluid constituents, and we get a set of N differential equations describing the evolution of the perturbations δ_i as in Eq. (A.10) with proper pressure p_i, velocity v_i and equation of state w_i. Let's note that all of the perturbed GR equations are affected, leading to complex coupling between each species and Φ.

We give now a pedagogical example of how to treat perturbations of the matter density field in a multi-component fluid, considering the simplest case of a matter-radiation universe, where local density in the Universe is given by $\rho = \rho_m + \rho_r$. It can be useful to set the time variable $y = \bar{\rho}_m/\bar{\rho}_r$ rather than the time coordinate giving $y = a/a_{eq}$ where a_{eq} is the value of the scale factor when matter and radiation densities are equal. Then $y \gg 1$ corresponds to when the fluid is composed of matter only, and $y \ll 1$ when the fluid is composed of radiation.

The cosmological fluid has now *effective* properties, such as an equation of state w and speed of sound c_s given by

$$w = \frac{1}{3}\frac{1}{1+y} \quad \text{and} \quad c_s^2 = \frac{4}{3(4+3y)}, \tag{A.17}$$

We have the two appropriate limit cases $w = 0 = c_s = 0$ for $y \gg 1$ (fluid of matter only) and $w = c_s^2 = 1/3$ for $y \ll 1$ (radiation only). First, we found that in the matter-dominated era, $\delta_k \sim a/a_{eq} \sim t^{2/3}$ which is also the result obtained from the classical Newtonian approach. However, during the radiation-dominated era, the gravitational potential is fully determined by the radiation perturbation from Eq. (A.12), and the growing modes of matter perturbations are given by $\delta_k \sim \ln a$, and starts growing proportional to the scale factor when entering the matter-dominated era. It becomes much more complicated in our ΛCDM universe since the cosmological fluid has many more components (as studied in Peter & Uzan (2013); Kurki-Suonio (2022)).

Appendix B

Weak gravitational lensing

This appendix presents some calculation details on weak gravitational lensing, introduced in Chapter 1. This appendix is based on the review of Schneider (2006); Baker (2017); Bartelmann & Schneider (2001) but also on the online "Notes on Gravitational Lensing" of S. J. Hilbert. From Eq. (1.69), we define the deformation matrix $\mathcal{A} = \partial\vec{\beta}/\partial\vec{\theta}$ that describes the small change in observed angular separation with respect to the *unperturbed* path and can be decomposed in two terms, an an-isotropic an isotropic deformation represented respectively by the *shear* $\gamma = \gamma_1 + i\gamma_2$ and the *convergence* κ such as

$$\mathcal{A}(\vec{\theta}, z) = 1(1-\kappa) + \begin{pmatrix} -\gamma_1 & -\gamma_2 \\ -\gamma_2 & \gamma_1 \end{pmatrix}, \tag{B.1}$$

where 1 is the 2-dimension identity matrix. The shear and convergence both relate to the lensing potential such as

$$\kappa = \frac{\Delta\psi}{2} = \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi \quad \text{and} \quad \gamma = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi + i\partial_1\partial_2\psi, \tag{B.2}$$

where $\partial_i\partial_j$ corresponds to derivatives with respect to the projected angular positions. Taking the Fourier transform of the two above equations, we get that $\widehat{\kappa}(\vec{l}) = \mathrm{e}^{2i\beta}\widehat{\gamma}(\vec{l})$ where β is the polar angle of the wave vector $\vec{l}$. So shear and convergence are closely related. Using the Poisson equation

$$\Delta\Phi = a^{-1}4\pi G\overline{\rho}^0_{\mathrm{m}}(1+\delta_{\mathrm{m}}), \tag{B.3}$$

where δ_m is the local matter overdensity, the convergence takes the form of the projected matter density along the line of sight

$$\kappa(\vec{\theta}, z) = \overline{\rho}^0_{\mathrm{m}}\int_0^r dr'\Sigma^{-1}_{\mathrm{crit}}(z', z)\delta_{\mathrm{m}}(f_K(r')\vec{\theta}, r'), \tag{B.4}$$

weighted by the (inverse) comoving critical surface mass density

$$\Sigma_{\mathrm{crit}}(z' > z, z) = \frac{c^2}{4\pi G}\frac{\chi(r')}{\chi(r)\chi(r'-r)}\frac{1}{1+z}. \tag{B.5}$$

Now we have related the convergence and shear to the underlying projected matter distribution along the line-of-sight through κ, it is possible to derive the observational effects of weak gravitational lensing of foreground structures on different background emitters.

- *Magnification of background sources*: The weak gravitational lensing preserves the surface brightness of the sources, but modifies the apparent solid angle spanned by the lensed

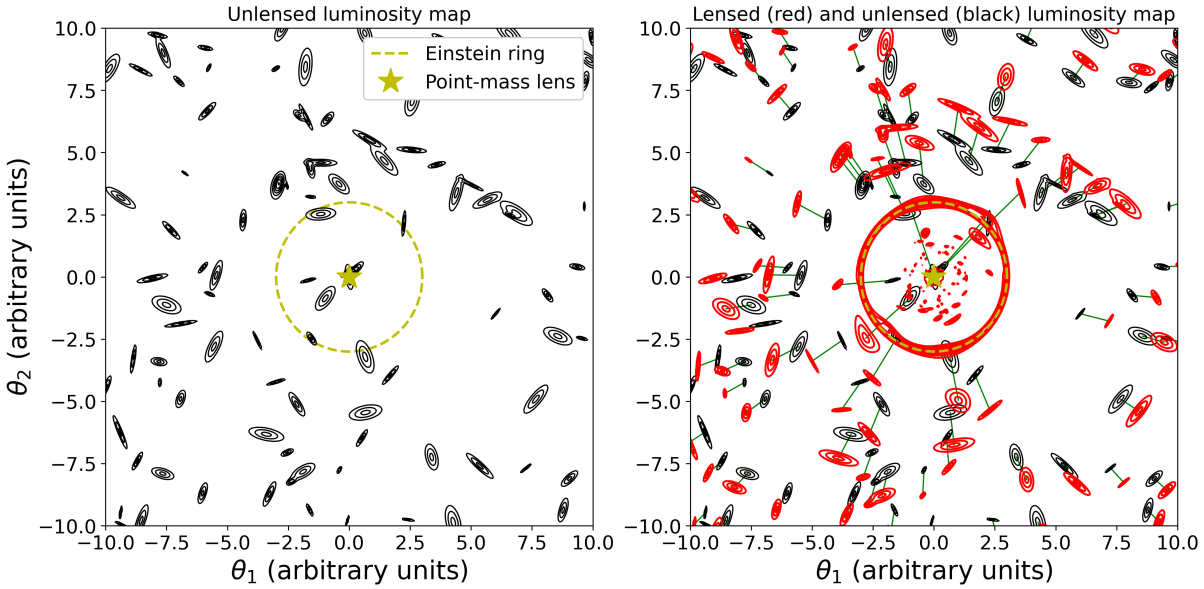

Figure B.1: Left: unlensed luminosity map, right: lensed (red) and unlensed (black) luminosity maps due to the presence of a point-mass lens in the foreground (represented by the yellow star). Proportions, angular scales, lens mass and luminosity map characteristics are arbitrary.

source image. The magnification $\mu(z,\vec{\theta})$ of a source at redshift z and location $\vec{\theta}$ is given by

$$\mu(z,\vec{\theta}) = \frac{1}{\det\mathcal{A}} \approx 1 + 2\kappa(z,\vec{\theta}), \tag{B.6}$$

where the final approximation is obtained by a first-order Taylor expansion in $\kappa, |\gamma| \ll 1$ in the weak lensing regime. The source flux f_0 of a galaxy is magnified and its observed flux is $f_{\rm obs} = \mu f_0$. At the same time, the true solid angle $d^2\vec{\theta}_0$ spanned by the source is distorted by a factor μ, then becomes $d^2\vec{\theta}_{\rm obs} = \mu d^2\vec{\theta}_0$. Taking account of these two effects, the lensed cumulative number density of objects at redshift z with limited flux f and given by

$$n_{\rm obs}(>f,z) = \frac{1}{\mu(\vec{\theta},z)} n_0\left(> \frac{f}{\mu(\vec{\theta},z)}, z\right), \tag{B.7}$$

where n_0 is the intrinsic cumulative number density of galaxies. The lensing of source flux allows us to measure fainter galaxies relative to the true limited flux of the sample. Assuming that the intrinsic cumulative number density expresses as a power law on the flux and is given by $n_0(f,w) \propto f^{-s}$, we get

$$n_{\rm obs}(<m,z) = \mu(\vec{\theta},w)^{2.5s(m)-1} n_0(<m,z), \tag{B.8}$$

where we changed the flux to the apparent magnitude $m(f) = -2.5\log_{10} f$. Then, the galaxy overdensity can be decomposed in two contribution $\delta_g(\vec{\theta}) = \delta_{gg}(\vec{\theta}) + \delta_\mu(\vec{\theta})$, where $\delta_{gg}(\vec{\theta})$ is due to intrinsic galaxy clustering in absence of lensing. The overdensity $\delta_\mu(\vec{\theta})$ denotes the contribution from lensing to the variability of the observed number density of galaxies. Regarding Eq. (B.8), it is defined as

$$\delta_\mu(\vec{\theta},z) \approx 2(2.5s(m_{\rm lim}) - 1)\kappa(\vec{\theta},z), \tag{B.9}$$

so weak lensing magnification introduces a correlation between the matter density field along the line-of-sight and background galaxies. This effect can be measured by estimating the cross-correlation between the background number density of galaxy and a tracer of the foreground matter density field.

- *Deformation of source galaxy images* The deformation matrix $\mathcal{A}$ also describes the coherent deformation of source image. Considering a source galaxy image to be an ellipse with intrinsic ellipticity $\epsilon^{\rm int}$, the observed ellipticity $\epsilon^{\rm obs}$ is inferred from the second moments of the second moments Q_{ij} of its surface brightness $I_{\rm obs}$ such as

$$Q_{ij} = \frac{\int d^2\vec{\beta} I_{\rm obs}(\vec{\beta})(\beta_i - \bar{\beta}_i)(\beta_j - \bar{\beta}_j)}{\int d^2\vec{\beta} I_{\rm obs}(\vec{\beta})}, \tag{B.10}$$

where $\bar{\beta}$ is the centroid of the galaxy. The observed ellipticity $\epsilon^{\rm obs} = [(1-b/a)/(1+b/a)]e^{i2\phi}$ (a and b are respectively the major and minor axes, and ϕ is the orientation) is given by

$$\epsilon^{\rm obs} = \frac{Q_{11} - Q_{22} + iQ_{12}}{Q_{11} + Q_{22} + 2(Q_{11}Q_{22} - Q_{12}^2)^{1/2}} = \frac{\epsilon^{\rm int} + g}{1 + g^*\epsilon^{\rm int}}, \tag{B.11}$$

where $g = \gamma/(1-\kappa)$ is the reduced shear. We have obtained the last equality by considering the change of variable $\vec{\theta} \rightarrow \vec{\beta}(\vec{\theta})$ in the Q_{ij} integral, and applying the conservation of the surface brightness after the change of coordinates i.e. $I_{\rm true}(\vec{\theta}) = I_{\rm obs}[\vec{\beta}(\vec{\theta})]$ (called the Liouville theorem). Moreover, the apparent surface of the object relates to the original one following $S_{\rm obs} = S_{\rm true}/[(1-\kappa)^2 - |\gamma|^2]$. Then, the measurement of galaxy shapes can be directly related to the projected matter distribution, that induces coherent deformation of source images. In the weak lensing regime, we can relate the observed shape to the weak lensing shear even more directly to the shear via the weak lensing equation

$$\epsilon^{\rm obs} = \epsilon^{\rm int} + g + o(|g|^2) = \epsilon^{\rm int} + \gamma + o(|\gamma|^2, \kappa^2). \tag{B.12}$$

Making the average of the above equation at a given location in the sky, the intrinsic ellipticity term vanishes (in the absence of intrinsic alignment at large scales).

We show a pedagogical example in the Fig. B.1 about shape distortion and displacement due to weak gravitational lensing. We start by defining a luminosity map, calculated as the sum of individual luminosity maps (associated to individual galaxies) each of them characterized by an axis ratio, orientation and amplitude. The full luminosity map is represented in the first plot (left) in Fig. B.1. Then, we place a point-mass lens (for which the lens equation is easy to solve, e.g. a black hole) in the foreground, at the center of the image (represented yellow star). The second plot (right) shows the impact of the point mass distribution on the un-lensed luminosity map, represented in black. The observed luminosity map (here in red, obtained by applying the Liouville theorem) shows that the individual luminosity fluxes are "shifted" from the lens center (the center of the lensed and un-lensed galaxies are connected by a green segment). Moreover, we see that the observed individual shapes are distorted. For galaxies initially perfectly aligned with the lens center along the line-of-sight, the corresponding lensed image is "spread" over the Einstein ring (yellow dashed circle). We also see that each galaxy has a replica lying inside the Einstein ring, which is predicted by the lens equation for a point mass spherical lens (one solution is located outside the Einstein ring, the other is inside).

Appendix C

Discussions on the halo mass function

In this appendix, we discuss the formalism of the halo mass function, presented in Chapter 2. We also present the Fisher formalism to forecast the performance of the cluster abundance to constrain cosmology.

The halo mass function

A general form of the halo mass function is

$$\frac{dn(m,z)}{dm} = \frac{\rho_m(z)}{m}\left|\frac{d}{dm}\mathcal{F}(>m)\right|, \tag{C.1}$$

where $\mathcal{F}(>m)$ is the fraction of material in collapsed objects of mass $>m$ (or equivalently the probability that a particle at $\vec{x}$ is part of a collapsed object with mass $>m$). In a hierarchical model of structure formation, we recall that all the matter is assumed to be contained within halos. Then, we have that the mean matter density follows

$$\int dm\; m\frac{dn(m,z)}{dm} = \rho_m(z) \quad \text{such that} \quad 1 = \int dm\left|\frac{d}{dm}\mathcal{F}(>m)\right|. \tag{C.2}$$

The $d\mathcal{F}(>m)/dm$ term is usually given in term of the multiplicity function $f(\sigma)$,

$$\frac{d\mathcal{F}}{dm} = \frac{d\ln\sigma}{dm}\frac{d\mathcal{F}}{d\ln\sigma} \equiv \frac{d\ln\sigma}{dm}f(\sigma), \tag{C.3}$$

where $\sigma \equiv \sigma(m)$ is the variance of the density field, which is related to the matter power spectrum, filtered on a scale R enclosing an average mass $m = V_R\rho_m(z)$ as defined in Section 1.3.3. From a variety of simulation and theoretical approaches, the multiplicity function is found not to depend strongly on mass, redshift and cosmology, which is called the *universality* of the halo mass function. Press & Schechter (1974) first advanced a theory for the mass function, that we discuss now. They assumed that only the smooth linear Gaussian density field governs the structure formation. For Gaussian random field δ, it follows the probability density function

$$p(\delta) \propto \exp -\frac{\delta^2}{2\sigma^2}. \tag{C.4}$$

For a smoothed linear density perturbation within the radius R and with enclosed mean mass m, the probability for exceeding a critical threshold δ_c is given by

$$P(\delta_R > \delta_c) = \int_{\delta_c}^{\infty} p(\delta'_R)d\delta'_R. \tag{C.5}$$

They showed that fluctuations above the threshold $\delta_c \approx 1.69$ correspond to collapsed regions, and proposed that

$$\mathcal{F}(>m) = P(\delta_R > \delta_c) = \frac{1}{2}\left[1 - \mathrm{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma(R)}\right)\right]. \tag{C.6}$$

Taking the power spectrum $P(k) \propto k^n$ we can show that $\sigma^2(m) \propto m^{-\frac{(n+3)}{3}}$. For $m \to 0$, we get $\mathcal{F}(>0) = \frac{1}{2}$, suggesting that only half of the mass in the universe is part of collapsed objects of any mass. This is because for halos with $\delta < \delta_c$, collapse is possible on larger scales giving $\delta > \delta_c$. In other words, for Gaussian fields, half of the fluctuations are under dense, and these under dense regions will fall onto the massive objects. That is not accounted for in the Press & Schechter formalism (PS), and this is known as the "cloud-in-cloud" problem. Press & Schechter (1974), without a proper demonstration, adopted that $\mathcal{F}(>m) = 2P(\delta_R > \delta_c)$. The "cloud-and-cloud" problem was solved by Bond et al. (1991) (a pedagogical derivation of the "fudge" factor of 2 can be found in Peacock (1999), and summed up within the online lectures AST4320 of M. Dijkstra). The Press & Schechter multiplicity function $f_{\rm PS}$ is then given by

$$f_{\rm PS}(\nu) = \sqrt{\frac{2}{\pi}}\nu \exp\left(-\frac{\nu^2}{2}\right) \tag{C.7}$$

where $\nu = \delta_c/\sigma(m)$. The assumption of spherical collapse proposed by the PS formalism reproduces quite well the results of N-body simulations at high mass. Later, Sheth et al. (2001) proposed a multiplicity function that accounts for ellipsoidal collapse by introducing a mass-dependent collapse threshold, that is reacquired when the initial tidal field is elliptical. By applying the "excursion set approach" (Bond et al., 1991) to solve the "fudge" factor of 2 for ellipsoidal collapse, they find the multiplicity function to be well approximated by

$$f_{SMT}(\nu) = 0.322 \times \left(1 + \frac{1}{(0.84\nu)^{0.6}}\right) \times f_{PS}(0.84\nu). \tag{C.8}$$

The coefficients $0.322, 0.84$ and 0.6 provide a significantly better fit to the halo mass distributions that are observed in numerical simulations over a wide range of both mass and redshift, by notably predicting less low mass objects than the PS formalism, what is seen in simulations of hierarchical clustering. The actual dynamics of gravitational collapse and the process of dark matter halo formation is even more complex since it accounts for non-spherical collapse beyond elliptical as proposed by Sheth et al. (2001), but also interactions between halos, impact of baryon physics, feedback of Active Galactic Nuclei (AGN), etc.

Since then, Despali et al. (2015); Tinker et al. (2008); Tinker et al. (2010); Watson et al. (2013); Bocquet et al. (2015) have devised more accurate parameterizations than the PS and SMT formalisms. As an example, we show the Tinker et al. (2008) halo mass function in Fig. C.1. Other halo mass functions are shown in Fig. 2.2.

Forecasting the performance of cluster abundance cosmology

Fisher (1935) noted that some key ingredients of a cosmological probe can be combined to fairly describe its constraining power. We can first derive by the Fisher information associated to the parameter θ_α that is given by

$$F_{\rm CA} = \mathrm{E}_{\mathcal{L}_{\rm CA}}\left[\left(\frac{\partial \ln \mathcal{L}_{\rm CA}}{\partial \theta_\alpha}\right)^2\right] \equiv \sum_{i,j=1} \frac{\partial N_i}{\partial \theta_\alpha} \Sigma^{-1}_{ij} \frac{\partial N_j}{\partial \theta_\alpha}, \tag{C.9}$$

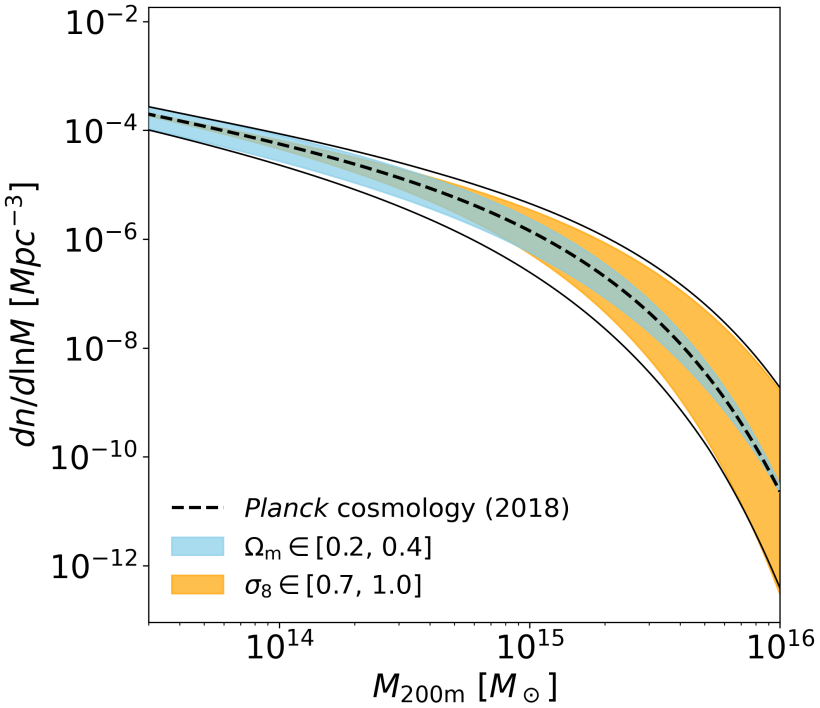

Figure C.1: Tinker et al. (2008) halo mass function at $z = 0$. The black dashed line corresponds to the predicted halo mass function for the Planck Collaboration et al. (2020a) cosmology ($\Omega_m = 0.315$ and $\sigma_8 = 0.811$). The region spanned by the halo mass function when Ω_m varies from 0.2 to 0.4 is represented in blue, and σ_8 ranging from 0.7 to 1 is represented in orange. These ranges correspond typically to priors that are used in cluster abundance analysis. The regions comprised between the two black lines corresponds to the variation of the halo mass function when Ω_m and σ_8 vary together and independently on their respective ranges. Varying Ω_m on its range induces roughly constant deviation of the *Planck* halo mass function (here $\sim 50\%$ variation) on the all mass range, whereas varying σ_8 impacts much more the abundance of very massive clusters (here, about $10^4\%$) rather than for low mass clusters (about $\sim 10\%$).

where the last term of the above equation is obtained when considering reasonable assumptions for the cluster abundance likelihood (again, we will discuss it in detail in Chapter 5). The Fisher information depends on $\partial N_i/\partial\theta_\alpha$ that is the first derivative of the predicted cluster count in bin i with respect to θ_α, and the inverse covariance matrix element Σ_{ij}^{-1}. Second, Cramer (1952) showed that the variance of each cosmological parameter derived from the posterior $\mathcal{P}$

$$\sigma^2(\theta_\alpha) = \mathrm{E}_{\mathcal{P}_{\mathrm{CA}}}[(\theta_\alpha - \mathrm{E}_{\mathcal{P}_{\mathrm{CA}}}[\theta_\alpha])^2] \tag{C.10}$$

is at least as high as the inverse of the Fisher information F_{CA}, i.e.

$$\sigma^2(\theta_\alpha) \geq F_{\mathrm{CA}}^{-1}. \tag{C.11}$$

Consequently, the Fisher information is a very important statistical quantity to quantify the overall precision of a cosmological probe, since the error is bounded lower by some intrinsic level of uncertainty; When F is large, the parameter uncertainty may be small and reversely (more information $\equiv$ less error). We can see two distinct features in Eq. (C.9), involving the *responsivity* and the *precision* of cluster abundance cosmology.

- $\partial N_i/\partial\theta_\alpha$ is the *responsivity* of the cosmological probe. If N_i varies strongly with respect to cosmological parameters, then $(\partial N_i/\partial\theta_\alpha)^2$ increases, F_{CA} increases, and the parameter variance decreases; the constraints are then stronger, the cosmological probe has a "strong" response to a change in parameters. For cluster abundance, we have shown that it is highly sensitive to the growth of structure, depending on the mass and redshift domain used to measure the cluster distribution. For instance, the halo mass function is more sensitive to σ_8 at high mass and Ω_m roughly impacts all masses at similar level (see Fig. C.1).

- Σ_{ij}^{-1} is called the *precision* matrix, which intuitively is desired to be as large as possible. Let's consider the simplest case with uncorrelated bins. If the variance of each count N_i is large, F_{CA} decreases, and the error of parameters increases, so the precision of a cosmological parameter is directly linked to the inverse variance of the data. For a variety of cosmological inference setups, the major contribution to the cluster abundance variance is the Poisson noise, but we discuss it more in detail in Chapter 5.

Appendix D

Forecasting the parameter ensemble covariance

This appendix addresses the derivation of the parameter ensemble covariance matrix from binned cluster abundance, as introduced in Section 5.1.

Formalism

We consider the parameter estimator $\widehat{\theta}_\alpha(\vec{N}) = \widehat{\theta}_\alpha(\vec{\lambda}_0 + \delta\vec{N})$. Where $\vec{\lambda}_0$ is the average true model prediction (i.e. the observed data are sampled according to $\vec{\lambda}_0$ as the mean). The Taylor expansion of the parameter estimator at $\vec{N} = \vec{\lambda}_0$ is given by

$$\widehat{\theta}_\alpha - \theta_\alpha = D_k^{(\alpha)}\delta N_k + D_{lm}^{(\alpha)}\delta N_l \delta N_m + \odot(\delta N^3) \tag{D.1}$$

Here $D_k^{(\alpha)}$, $D_{kl}^{(\alpha)}$ and $D_{klm}^{(\alpha)}$ correspond to the first, second and third order coefficients of the Taylor expansion. The ensemble covariance matrix is given by

$$\mathcal{C}_{\alpha\beta}^{\text{ens}} = \langle(\widehat{\theta}_\alpha - \theta_\alpha)(\widehat{\theta}_\beta - \theta_\beta)\rangle, \tag{D.2}$$

where the average is performed considering the *true* underlying likelihood. We get

$$\begin{aligned}
\mathcal{C}_{\alpha\beta}^{\text{ens}} \approx & [\Sigma_X]_{kl} D_k^{(\alpha)} D_l^{(\beta)} && \text{(D.3)}\\
& + [\Sigma_X]_{klm}[D_k^{(\alpha)} D_{lm}^{(\beta)} + D_k^{(\beta)} D_{lm}^{(\alpha)}] && \text{(D.4)}\\
& + [\Sigma_X]_{klmn}[D_{kl}^{(\alpha)} D_{mn}^{(\beta)} + D_k^{(\alpha)} D_{lmn}^{(\beta)} + D_k^{(\beta)} D_{lmn}^{(\alpha)}], && \text{(D.5)}
\end{aligned}$$

where $[\Sigma^X]_{klm}, [\Sigma^X]_{klmn}$ are respectively the 3^{rd} and 4^{th} order correlation matrix of the data. By the definition of the Taylor expansion, we have

$$D_{i\ldots k}^{(\alpha)} = \frac{1}{n!}(\widehat{\theta}_\alpha)_{,i\ldots k}, \tag{D.6}$$

with n being the order of the derivative, and $_{,i\ldots k} = .\frac{\partial^n}{\partial_i \ldots \partial k}$. So we have derived a link between the ensemble parameter covariance matrix and the data covariance matrix Σ_X, as well as the properties of the estimator encoded in the $D^{(\alpha)}$ factors. We have, from Efron (2015) that

$$D_i^{(\alpha)} = \text{Cov}_{\mathcal{P}_Y}[\theta_\alpha, (\ln \mathcal{L}_Y(\vec{N}|\theta))_{,i}]. \tag{D.7}$$

We obtain, by the quotient rule that

$$D_{ij}^{(\alpha)} = \frac{1}{2}\frac{A_{,ij} - B_{,ij}\widehat{\theta}_\alpha - B_{,j}(\widehat{\theta}_\alpha)_{,i} - B_{,i}(\widehat{\theta}_\alpha)_{,j}}{B}. \tag{D.8}$$

Also, by using the notation used in Efron (2015), we have

$$\frac{A_{,ij} - B_{,ij}\widehat{\theta}_\alpha}{B} = \mathrm{Cov}_{\mathcal{P}_Y}(\{\theta_\alpha, \frac{\mathcal{L}_Y(\vec{N}|\theta)_{,ij}}{\mathcal{L}_Y(\vec{N}|\theta)}\}), \tag{D.9}$$

and

$$\frac{B_{,i}}{B} = \mathrm{E}_{\mathcal{P}_Y}[(\ln \mathcal{L}_Y)_{,i}]. \tag{D.10}$$

Then, we get the general result of the second derivative of the parameter covariance

$$D^{(\alpha)}_{ij} = \frac{1}{2}\left[\mathrm{Cov}_{\mathcal{P}_Y}(\{\theta_\alpha, \frac{\mathcal{L}_Y(\vec{N}|\theta)_{,ij}}{\mathcal{L}_Y(\vec{N}|\theta)}\}) - D^{(\alpha)}_{[i}\mathrm{E}_{\mathcal{P}_Y}[(\ln \mathcal{L}_Y)_{,j]}]\right]. \tag{D.11}$$

We have by the quotient rule that

$$D^{(\alpha)}_{ijk} = \frac{1}{6}\left[\mathrm{Cov}_{\mathcal{P}_Y}(\{\theta_\alpha, \frac{\mathcal{L}_Y(\vec{N}|\theta)_{,ijk}}{\mathcal{L}_Y(\vec{N}|\theta)}\}) - D^{(\alpha)}_{[i}\mathbb{E}_{\mathcal{P}_Y}\left[\frac{(\mathcal{L}_Y)_{,jk]}}{\mathcal{L}_Y}\right] - D^{(\alpha)}_{[ij}\mathbb{E}_{\mathcal{P}_Y}\left[\frac{(\mathcal{L}_Y)_{,k]}}{\mathcal{L}_Y}\right]\right]. \tag{D.12}$$

More generally, for the n-th derivative, we have

$$D^{(\alpha)}_{ijk\ldots p} = \frac{1}{n!}\left[\mathrm{Cov}_{\mathcal{P}_Y}(\{\theta_\alpha, \frac{\mathcal{L}_Y(\vec{N}|\theta)_{,ij\ldots p}}{\mathcal{L}_Y(\vec{N}|\theta)}\}) - \sum_{k'_i \in \{ijk\ldots n\}} D^{(\alpha)}_{[ij\ldots k'_i}\mathrm{E}_{\mathcal{P}_Y}\left[\frac{(\mathcal{L}_Y)_{,k'_{i+1}\ldots n]}}{\mathcal{L}_Y}\right]\right] \tag{D.13}$$

The case of Gaussian likelihood

This calculation refers to the results presented in Section 5.1.2 and is taken from my paper Payerne et al. (2023). The gradient $\nabla \ln \mathcal{L}_Y$ can be expressed for the particular case of a Gaussian likelihood $\mathcal{L}_Y$ with an analysis covariance matrix Σ_Y, and we have that

$$\nabla \ln \mathcal{L}_Y(\vec{N}|\theta) = \Sigma_Y^{-1}[\vec{\lambda}(\theta) - \vec{N}]. \tag{D.14}$$

Using Eq. (D.14) in Eq. (D.7), we get the general expression, which is valid for an *input* Gaussian likelihood $\mathcal{L}_Y$ that

$$[\nabla\widehat{\theta}_\alpha]_i = [\Sigma_Y^{-1}]_{ij}\mathrm{Cov}_{\mathcal{P}_Y}[\{\theta_\alpha, \lambda_j(\vec{\theta})\}|\vec{N}]. \tag{D.15}$$

Second, we consider the parameter estimator $\widehat{\theta}(\vec{N})$, at fixed observed cluster count $\vec{N}$. We now consider that the observed cluster count $\vec{N}$ corresponds to the *latent* average cluster count λ_0 with small additive deviation, i.e. we can write $\vec{N} = \vec{\lambda}_0 + \delta\vec{N}$. In this formalism, we state that $\vec{\theta}_0$ are the *latent* cosmological parameters giving $\vec{\lambda}_0 = \lambda(\vec{\theta}_0)$, and we state that $\vec{\lambda}_0 = \mathbb{E}_{\mathcal{L}_Y}(\vec{N}|\vec{\theta}_0) = \mathbb{E}_{\mathcal{L}_X}(\vec{N}|\vec{\theta}_0)$. We consider the Taylor expansion of the estimator $\widehat{\theta}(\vec{\lambda}_0 + \delta\vec{N})$ considering the deviation $\delta\vec{N} = \vec{N} - \vec{\lambda}_0$, so we get for the first order that

$$\widehat{\theta}_\alpha(\vec{N}) - \widehat{\theta}_\alpha(\lambda_0) \approx [\nabla\widehat{\theta}_\alpha]_i \cdot [\vec{N} - \vec{\lambda}_0]_i, \tag{D.16}$$

where in Eq. (D.16), $[\nabla\widehat{\theta}_\alpha]_i$ corresponds to the gradient defined in Eq. (D.7) evaluated at $\vec{N} = \vec{\lambda}_0$. For Eq. (D.16) to be valid, it is assumed that the estimator $\widehat{\theta}(\vec{N})$ is *sufficiently* linear around $\vec{\lambda}_0$ with respect to the data, as it is considered for linear error propagation from data to estimated parameter (Ku, 1996). We now assume that the model is well approximated by a linear dependence on the parameter of interest, e.g. Raveri & Hu (2019), and taking the first order in the Taylor expansion of $\lambda(\theta)$ at θ_0, we get that

$$[\nabla\widehat{\theta}_\alpha]_i = [\Sigma_Y^{-1}]_{ij}\mathcal{C}^{\mathrm{ind}}_{\alpha\gamma}(\vec{\lambda}_{,\gamma})_j = [\Sigma_Y^{-1}(\mathcal{C}^{\mathrm{ind}}_{\alpha\gamma}\vec{\lambda}_{,\gamma})]_i, \tag{D.17}$$

where $\mathcal{C}^{\rm ind}_{\alpha\gamma} = {\rm Cov}_{\mathcal{P}_Y}[\{\theta_\alpha, \theta_\beta\}|\vec{\lambda}_0]$ is the individual posterior covariance matrix, and $\vec{\lambda}_{,\gamma}$ is the derivative of the model prediction evaluated at $\theta = \theta_0$. Note that we subsequently conserved the first-order moments of the two Taylor expansions respectively in the data and parameter space, that need to be valid to follow the proof. We will consider these approximations correct for the following steps. The ensemble covariance is given by

$$\mathcal{C}^{\rm ens}_{\alpha\beta} = {\rm Cov}_{\mathcal{L}_X}[\{\widehat{\theta}_\alpha(\vec{N}), \widehat{\theta}_\beta(\vec{N})\}|\vec{\theta}_0] \tag{D.18}$$

$$= (\mathcal{C}^{\rm ind}\vec{\lambda}_{,})^T_\alpha \Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}(\mathcal{C}^{\rm ind}\vec{\lambda}_{,})_\beta. \tag{D.19}$$

In Eq. (D.19), the statistical properties of the data assumed by choosing $\mathcal{L}_Y$ appear in Σ_Y^{-1}, as well as in the individual parameter covariance computed using $\mathcal{P}_Y$. Whereas the *latent* statistical properties of data are included in Σ_X. The Eq. (D.19) is an application of the *delta method approximation* method for error propagation (Efron (2015), see their Eq. (2.8)) for the Gaussian case, but also a generalization since we take account of covariance between parameters and different *latent* and *analysis* data covariances respectively Σ_X and Σ_Y. We can explore the solution when, in this formalism, we get $\mathcal{C}^{\rm ens} = \mathcal{C}^{\rm ind} = \mathcal{C}$, i.e. when the covariance of the estimator when considering an infinite realization of the data $\vec{N}$ is consistent with the individual parameter covariance obtained from a single dataset $\vec{N}$. Following the Gaussian approximation of individual posteriors in the parameter space, we can use the Fisher forecast for $\mathcal{C}^{\rm ind}$, since the Eq. (D.19) is evaluated at $\vec{N} = \vec{\lambda}_0$. Using this, we can simplify the problem by trying to find the relation between the data covariances Σ_X and Σ_Y to show that

$$\mathcal{C}_{\alpha\beta} = (\mathcal{C}\vec{\lambda}_{,})^T_\alpha \Sigma^{-1}(\mathcal{C}\vec{\lambda}_{,})_\beta \tag{D.20}$$

$$\mathcal{C}^{-1}_{\alpha\beta} = \vec{\lambda}^T_{,\alpha}\Sigma_X^{-1}\vec{\lambda}_{,\beta} \tag{D.21}$$

where we defined the modified covariance matrix $\Sigma^{-1} = \Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}$ in Eq. (D.20), and Eq. (D.21) is given by Fisher forecasting the individual parameter covariance in Eq. (5.25), that is always true for Gaussian and Poisson likelihoods. Multiplying Eq. (D.20) by $\mathcal{C}^{-1}_{\eta\alpha}$ and summing over α, we get

$$\delta^K_{\eta\beta} = \vec{\lambda}^T_{,\eta}\Sigma^{-1}(\mathcal{C}\vec{\lambda}_{,})_\beta = \vec{\lambda}^T_{,\eta}\Sigma^{-1}\vec{\lambda}_{,\delta}\mathcal{C}_{\delta\beta}. \tag{D.22}$$

By re-multiplying the above equation by $\mathcal{C}^{-1}_{\alpha\beta}$, we get the result

$$\mathcal{C}^{-1}_{\eta\alpha} = \vec{\lambda}^T_{,\eta}\Sigma^{-1}\vec{\lambda}_{,\alpha}, \tag{D.23}$$

which is the same as Eq. (D.21) when setting $\Sigma_X = \Sigma_Y$. So we found that for the ensemble parameter covariance to be equal to the individual parameter covariance, this formalism requires that the data covariance of the analysis and latent likelihoods are the same. In this appendix, we considered only the first order in the Taylor expansion in Eq. (D.16). When posteriors are not Gaussian, this approximation is not valid.

More information on the ensemble covariance can be found using the Cauchy-Schwarz's inequality given by

$$|{\rm Cov}_{\mathcal{L}_X}(A, B)|^2 \leq {\rm Var}_{\mathcal{L}_X}(A){\rm Var}_{\mathcal{L}_X}(B). \tag{D.24}$$

When considering

$$A = \widehat{\theta}_\alpha \quad \text{and} \quad B = (\ln\mathcal{L})_{,\beta}, \tag{D.25}$$

We have

$${\rm Cov}_{\mathcal{L}_X}(A, B) = (\mathcal{C}^{\rm ind}\vec{\lambda}_{,})^T_\alpha\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{\lambda}_{,\beta} \quad \text{and} \quad {\rm Var}_{\mathcal{L}_X}(B) = \vec{\lambda}^T_{,\beta}\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{\lambda}_{,\beta}. \tag{D.26}$$

Then, by multiplying each side of the equation by $\delta^K_{\alpha\beta}$ and summing over β, we get that the ensemble error of the parameter α verifies the *modified* Cramer-Rau lower bound

$$\mathcal{C}^{\rm ens}_{\alpha\alpha} \geq \frac{[(\mathcal{C}^{\rm ind}\vec{\lambda}_{,})^T_\alpha \Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{\lambda}_{,\alpha}]^2}{\vec{\lambda}^T_{,\alpha}\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{\lambda}_{,\alpha}}. \tag{D.27}$$

This is an even more general expression than previously found, since we do not need to make the Taylor expansion of the estimator around its mean.

The case of Poisson likelihood

We can now explore the case of a Poisson likelihood, where we will successively compute the mean and variance from a parameter posterior given a data vector, and then we compute the spread of mean known as the Frequentist variance of a Bayesian estimator. We first have the parameter posterior

$$P(\lambda|\{n_i\}) = \frac{\pi(\lambda)}{P(\{n_i\})}\prod_{k=1}^{N}\frac{\lambda^{n_k}}{n_k!}e^{-\lambda} = \frac{\pi(\lambda)}{P(\{n_i\})}e^{-N\lambda}\lambda^{S_N}\prod_{k=1}^{N}\frac{1}{n_k!}, \tag{D.28}$$

where

$$S_N = \sum_{k=1}^{N} n_k. \tag{D.29}$$

Considering positive flat prior, i.e. $\pi(\lambda) = \text{cst}$ that is large enough to contain the variation of the likelihood with respect to the parameter λ, we have the mean given by

$$\langle\lambda\rangle(\{n_i\}) = \frac{\int_0^{+\infty}\lambda P(\lambda|\{n_i\})d\lambda}{\int_0^{+\infty}P(\lambda|\{n_i\})d\lambda} = \frac{N^{-(2+S_N)}\Gamma(2+S_N)}{N^{-(1+S_N)}\Gamma(1+S_N)}. \tag{D.30}$$

$$= N^{-1}(1+S_N) \tag{D.31}$$

We have

$$\mathrm{E}[\langle\lambda\rangle(\{n_i\})] = N^{-1}(1+N\lambda) \to \lambda \text{ as } N \to +\infty. \tag{D.32}$$

The mean of the squared parameter is obtained by

$$\langle\lambda^2\rangle(\{n_i\}) = \frac{\int_0^{+\infty}\lambda^2 P(\lambda|\{n_i\})d\lambda}{\int_0^{+\infty}P(\lambda|\{n_i\})d\lambda} = \frac{N^{-(3+S_N)}\Gamma(3+S_N)}{N^{-(1+S_N)}\Gamma(1+S_N)} \tag{D.33}$$

$$= N^{-2}(S_N+2)(S_N+1) \tag{D.34}$$

$$= N^{-2}(S_N+1)^2 + N^{-2}(S_N+1) \tag{D.35}$$

Then, the posterior variance is given by

$$\sigma^2_{\rm ind} = N^{-2}(S_N+1) = \langle\lambda\rangle(\{n_i\})/N. \tag{D.36}$$

Taking its average over the data, we get that $\mathrm{E}[\sigma^2_{\rm ind}] = N^{-2}(1+N\lambda) \to \lambda/N$ as $N \to +\infty$. Once we compute the posterior variance, we get that the Frequentist variance of posterior means is given by

$$\sigma^2_{\rm ens} = \mathrm{E}[\langle\lambda\rangle^2] - \mathrm{E}[\langle\lambda\rangle]^2, \tag{D.37}$$

where

$$\mathrm{E}[\langle\lambda\rangle^2] = N^{-2}\langle(1+S_N)(1+S_N)\rangle = N^{-2}(1+3N\lambda+N^2\lambda^2), \tag{D.38}$$

$$\mathrm{E}[\langle\lambda\rangle]^2 = N^{-2}(1+N\lambda)^2 = N^{-2}(1+2N\lambda+N^2\lambda^2), \tag{D.39}$$

Then, we have that

$$\sigma^2_{\rm ens} = N^{-2} N\lambda = \lambda/N. \tag{D.40}$$

So,

$$\sigma^2_{\rm ens} = \sigma^2_{\rm ind} + \frac{1}{N^2} \to \sigma^2_{\rm ind} \text{ as } N \to +\infty. \tag{D.41}$$

Let's explore the case of a parameteric model. In the previous section, we found that the ensemble parameter and individual parameter covariances are equal when $\Sigma_X = \Sigma_Y$, when the analysis likelihood $\mathcal{L}_Y$ is a multivariate Gaussian distribution with data covariance Σ_Y, and where the latent likelihood is not specified, but with data covariance matrix Σ_X. In this section, we consider the case where the analysis likelihood $\mathcal{L}_Y$ is the Poisson likelihood. To compute the ensemble forecast, we first need to evaluate $[\nabla\widehat{\theta}_\alpha]_i$ in Eq. (D.7). First, we have that the Poisson log-likelihood is,

$$\ln \mathcal{L}_Y(\vec{N}|\vec{\theta}) = -\lambda_{\rm tot}(\vec{\theta}) + \sum_{k=1}^{c} \vec{N}_k \ln \lambda_k(\vec{\theta}) - \ln(\vec{N}_k!), \tag{D.42}$$

where $\lambda_{\rm tot}(\vec{\theta})$ is the total predicted abundance over all the mass-redshift bins. As given in Eq. (D.14), we take the first derivative of the log-likelihood with respect to the data vector component $\vec{N}_i$, giving

$$\nabla_i \ln \mathcal{L}_Y(\vec{N}|\vec{\theta}) = \ln \lambda_i(\vec{\theta}) - \psi(\vec{N}_i), \tag{D.43}$$

where the ψ is known as the digamma function and is defined as

$$\psi(\vec{N}_i) = -\gamma + \sum_{k'=1}^{\vec{N}_i} \frac{1}{k'}, \tag{D.44}$$

where $\gamma \approx 0.577$ is the Euler–Mascheroni constant, and the sum on the right side is the m-th harmonic number. We note that the quantity $\ln(\vec{\lambda}_0)_i - \psi(\vec{N}_i)$ is independent of $\vec{\theta}$, so it can be removed from the calculation of the covariance in Eq. (D.7). $[\nabla\widehat{\theta}_\alpha]_i$ now verifies the general form

$$[\nabla\widehat{\theta}_\alpha]_i = {\rm Cov}_{\mathcal{P}_Y}[\{\theta_\alpha, \ln\lambda_i(\vec{\theta})\}|\vec{N}]. \tag{D.45}$$

Next, taking the first order of the Taylor expansion of $\lambda_i(\vec{\theta})$ with respect to $\vec{\theta}$, we get $\lambda_i(\vec{\theta}) = (\lambda_0)_i + (\lambda_{,\gamma})_i(\vec{\theta} - \vec{\theta}_0)_\gamma$. Using $\ln(1+x) \approx x + o(x^2)$, we get finally that

$$[\nabla\widehat{\theta}_\alpha]_i = \frac{1}{(\vec{\lambda}_0)_i} \mathcal{C}^{\rm ind}_{\alpha\gamma}(\lambda_{,\gamma})_i = [\Sigma_Y^{-1}]_{ij} \mathcal{C}^{\rm ind}_{\alpha\gamma}(\lambda_{,\gamma})_j, \tag{D.46}$$

where $\Sigma_Y = {\rm diag}[\{(\vec{\lambda}_0)_i\}_{1\leq i\leq c}]$ is the covariance matrix of the data considering the Poisson distribution (in that case, abundances are not correlated and the variance of each k-th cluster count is the shot noise $(\vec{\lambda}_0)_k$), and $\mathcal{C}^{\rm ind}_{\alpha\gamma} = {\rm Cov}_{\mathcal{P}_Y}[\{\theta_\alpha, \theta_\beta\}|\vec{\lambda}_0]$ is the posterior covariance matrix. We find the same expression as for the Gaussian case in Eq. (D.17). To test the condition $\mathcal{C}^{\rm ens} = \mathcal{C}^{\rm ind}$ for the Poisson analysis likelihood, we first need to compute the fisher forecast for the Poisson case. From the Poisson likelihood in Eq. (D.42), we have that the second derivative of the Poisson log-likelihood writes

$$(\ln\mathcal{L})_{,\alpha\beta} = -(\lambda_{\rm tot})_{,\alpha\beta} + \sum_{k=1} \widehat{N}_k \left(\frac{(\lambda_{,\alpha\beta})_k}{\lambda_k} - \frac{(\lambda_{,\alpha})_k (\lambda_{,\beta})_k}{\lambda_k^2} \right), \tag{D.47}$$

Following the Fisher matrix general formula in Eq. (5.24) and using the equality

$$(\ln \mathcal{L})_{,\alpha\beta} = \frac{\mathcal{L}_{,\alpha\beta}}{\mathcal{L}} - (\ln \mathcal{L})_{,\alpha}(\ln \mathcal{L})_{,\beta}, \tag{D.48}$$

we get after a few calculations the Fisher information matrix for the Poisson likelihood that writes

$$\mathrm{F}_{\alpha\beta} = \sum_{k=1}^{c} \frac{(\lambda_{,\alpha})_k(\lambda_{,\beta})_k}{(\lambda_0)_k} = \vec{\lambda}^T_{,\alpha}\Sigma_Y^{-1}\vec{\lambda}_{,\beta}, \tag{D.49}$$

where $\mathcal{C}^{\mathrm{ind}}$ can be forecasted as F^{-1}. We see that the forecasted ensemble and individual parameter covariances for the Poisson analysis likelihood are obtained using the diagonal data covariance Σ_Y with only shot noise. The equality $\mathcal{C}^{\mathrm{ind}} = \mathcal{C}^{\mathrm{ens}} = \mathcal{C}$ is then verified when $\Sigma_X = \Sigma_Y$ (result obtained following the derivation from Eq. (D.20) to Eq. (D.23)).

The impact of non-flat prior

If a non-flat *prior* is considered in the Bayesian inference, then the Fisher information matrix can be decomposed in two terms: the first one is the standard Fisher matrix, and the second is given by the shape of the prior. We can write the individual parameter covariance with non-flat prior

$$(\mathcal{C}_p^{\mathrm{ind}})^{-1} = (\mathcal{C}^{\mathrm{ind}})^{-1} + C_0^{-1} \tag{D.50}$$

where C_0 is the covariance matrix of prior in the parameter space, namely a flat prior is given by $C_0 \to +\infty$. $(\mathcal{C}^{\mathrm{ind}})^{-1}$ is the standard Fisher matrix. Then, we have

$$[\mathcal{C}^{\mathrm{ens}}(\mathcal{C}^{\mathrm{ind}})^{-1}]_{\alpha\beta} = \begin{cases} [C_0]_{\alpha\gamma}\vec{\lambda}^T_{,\gamma}\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{\lambda}_{,\beta} & \text{if } \det C_0 \to 0 \\ \mathcal{C}^{\mathrm{ind}}_{\alpha\gamma}\vec{\lambda}^T_{,\gamma}\Sigma_Y^{-1}\Sigma_X\Sigma_Y^{-1}\vec{\lambda}_{,\beta} & \text{if } \det C_0 \to +\infty \end{cases} \tag{D.51}$$

Higher order forecasts

We want to derive $(C^{\mathrm{ens}}_{\alpha\beta})^{(1)} = [\Sigma_X]_{klm}[D_k^{(\alpha)}D_{lm}^{(\beta)} + D_k^{(\beta)}D_{lm}^{(\alpha)}]$ in Eq. (D.5). Taking the expression $D_{ij}^{(\alpha)}$, we can write the first term using

$$\frac{\mathcal{L}_{Y,ij}}{\mathcal{L}_Y} = (\ln \mathcal{L}_Y)_{,ij} + (\ln \mathcal{L}_Y)_{,i}(\ln \mathcal{L}_Y)_{,j}. \tag{D.52}$$

We get, after a few calculations,

$$\frac{1}{2}\mathrm{Cov}_{\mathcal{P}_Y}(\{\theta_\alpha, \frac{\mathcal{L}_{Y,ij}}{\mathcal{L}_Y}\}) = \frac{1}{2}[\Sigma_Y^{-1}\vec{\lambda}_{,\gamma}]_i[\Sigma_Y^{-1}\vec{\lambda}_{,\delta}]_j\mathcal{C}^{\mathrm{ind}}_{\alpha\gamma\delta}. \tag{D.53}$$

For the second term in $D_{ij}^{(\alpha)}$, we have that

$$\mathrm{E}_{\mathcal{P}_Y}[(\ln \mathcal{L}_Y)_{,j}] = (\Sigma_Y)^{-1}_{jk}\mathrm{E}_{\mathcal{P}_Y}[(\lambda(\theta) - \lambda_0)_k] = (\Sigma_Y)^{-1}_{jk}[(\mathrm{E}_{\mathcal{P}_Y}[\vec{\lambda}(\vec{\theta})] - \vec{\lambda}_0)_k]. \tag{D.54}$$

We have that $\mathbb{E}_{\mathcal{P}_Y}[\vec{\lambda}(\vec{\theta})]$ is not necessarily equal to $\vec{\lambda}_0$, when evaluating $\vec{N} = \vec{\lambda}_0$, because in the above equation $\vec{\lambda}_0 = \mathbb{E}_{\mathcal{L}_Y}[\vec{N}|\vec{\theta}_0]$. If we now consider the linear expansion at $\vec{\theta} = \vec{\theta}_0$, then

$$\mathbb{E}_{\mathcal{P}_Y}[(\vec{\lambda}(\vec{\theta}))_i] \approx (\vec{\lambda}_0)_i + (\lambda_i)_{,\gamma}\mathbb{E}_{\mathcal{P}_Y}[\theta_\gamma - (\theta_0)_\gamma]. \tag{D.55}$$

Then, if we have an unbiased estimator, then by definition $\mathbb{E}_{\mathcal{P}_Y}[\theta_\gamma - (\theta_0)_\gamma] = 0$ so we will consider the case where $\mathrm{E}_{\mathcal{P}_Y}[(\ln \mathcal{L}_Y)_{,j}] = 0$. We see that $D_{ij}^{(\alpha)}$ depends on the three-point correlation matrix of the individual posterior (given by $\mathcal{C}^{\rm ind}_{\alpha\gamma\delta}$), which is non-zero for non-Gaussian posterior (which is almost always the case). See work on higher order Fisher forecasts to predict $\mathcal{C}^{\rm ind}_{\alpha\gamma\delta}$ (Sellentin et al., 2014). Finally, we get that

$$D_{ij}^{(\alpha)} = \frac{1}{2}[\Sigma_Y^{-1}\vec{\lambda}_{,\gamma}]_i[\Sigma_Y^{-1}\vec{\lambda}_{,\delta}]_j\mathcal{C}^{\rm ind}_{\alpha\gamma\delta}, \tag{D.56}$$

and we have for a Gaussian likelihood, that

$$D_{ijk}^{(\alpha)} = -\frac{1}{6}\left[D^{(\alpha)}_{[i}\mathbb{E}_{\mathcal{P}_Y}\left[\frac{(\mathcal{L}_Y)_{,jk]}}{\mathcal{L}_Y}\right]\right], \tag{D.57}$$

$$= -\frac{1}{6}[\Sigma_Y^{-1}C^{\rm ind}_{\alpha\gamma}(\vec{\lambda}_{,\gamma})]_{[i]}\left([\Sigma_Y^{-1}]_{[jk]} + \mathcal{C}^{\rm ind}_{\beta\gamma}[\Sigma_Y^{-1}\vec{\lambda}_{,\beta}]_{[j]}[\Sigma_Y^{-1}\vec{\lambda}_{,\gamma}]_{[k]}\right). \tag{D.58}$$

Appendix E

Forecasting composite likelihoods

This appendix presents an attempt to forecast the Gauss-Poisson Compound likelihood introduced in Chapter 5. Forecasting the accuracy of a product of Gaussian likelihoods (which may be the case when performing joint analysis using multiple cosmological probes) can be done using the widely known Fisher formalism, namely if the total likelihood is

$$\mathcal{L} = \prod_{k=1}^{c} \mathcal{L}_k, \tag{E.1}$$

then we have that

$$F_{\alpha\beta} = \sum_{k=1}^{c} F^k_{\alpha\beta}, \tag{E.2}$$

where $F^k_{\alpha\beta}$ are individual fisher matrices. Let us define the composite likelihood $\mathcal{L}$, as the weighted average of c different likelihoods $\mathcal{L}_i$ such as

$$\mathcal{L} = \sum_{i=1}^{c} w_i \mathcal{L}_i = \langle \mathcal{L} \rangle_W. \tag{E.3}$$

To compute the Fisher matrix for the composite likelihood, we use that $F_{\alpha,\beta} = \mathrm{E}[(\ln \mathcal{L})_{,\alpha}(\ln \mathcal{L})_{,\beta}]$. Then, we have

$$F_{\alpha\beta} = \sum_{i,j=1}^{c} w_i w_j \mathrm{E}_{\mathcal{L}}[\frac{(\mathcal{L}_i)_{,\alpha}(\mathcal{L}_j)_{,\beta}}{\mathcal{L}^2}], \tag{E.4}$$

$$= \sum_{i,j,k=1}^{c} w_i w_j w_k \mathrm{E}_{\mathcal{L}_k}[\frac{(\mathcal{L}_i)_{,\alpha}(\mathcal{L}_j)_{,\beta}}{\mathcal{L}^2}]. \tag{E.5}$$

We want to calculate the f_{ijk} terms given by

$$f_{ijk} = \mathrm{E}_{\mathcal{L}_k}[\frac{(\mathcal{L}_i)_{,\alpha}(\mathcal{L}_j)_{,\beta}}{\mathcal{L}^2}] = \mathrm{E}_{\mathcal{L}_k}[\frac{\mathcal{L}_i\mathcal{L}_j}{\mathcal{L}^2}(\ln \mathcal{L}_i)_{,\alpha}(\ln \mathcal{L}_j)_{,\beta}], \tag{E.6}$$

when each individual likelihood is a Gaussian function, thus defined by the covariance matrix Σ_i and mean $\lambda_i = \lambda$ (we consider the simplest case where each likelihood shares the same mean). We have

$$(\ln \mathcal{L}_i)_{,\alpha} = -(\lambda_i)_{m,\alpha}[\Sigma_i^{-1}]_{mn}(\widehat{N} - \lambda)_n. \tag{E.7}$$

Now, the ratio $\frac{\mathcal{L}_i\mathcal{L}_j}{\mathcal{L}^2}$ is a non-trivial function of the data. We can consider the Taylor expansion

$$\frac{\mathcal{L}_i\mathcal{L}_j}{\mathcal{L}^2} = \frac{\mathcal{L}_i\mathcal{L}_j}{\mathcal{L}^2}|_0 + (d_1)_n(\widehat{N} - \lambda)_n. \tag{E.8}$$

We define the ratio

$$\frac{\mathcal{L}_i\mathcal{L}_j}{\mathcal{L}^2}|_0 = \frac{|\Sigma_i|^{-1/2}|\Sigma_j|^{-1/2}}{(\sum_{m=1}^{c} w_m|\Sigma_m|^{-1/2})^2} = r_i r_j, \tag{E.9}$$

where $|\Sigma|$ is the determinant of Σ, and $r_i = |\Sigma_i|^{-1/2}/\sum_{m=1}^{c} w_m|\Sigma_m|^{-1/2}$. The r_i terms follow the summation rule $\sum w_m r_m = 1$, so $w_n r_n$ are normalized weights. We can define the corresponding average $\langle\cdot\rangle_{WR}$. So we have that

$$f_{ijk} = \frac{\mathcal{L}_i\mathcal{L}_j}{\mathcal{L}^2}|_0 \times (\lambda)_{,\alpha}^T \Sigma_i^{-1}\Sigma_k\Sigma_j^{-1}(\lambda)_{,\beta}. \tag{E.10}$$

The terms corresponding to the first derivative of the ratio $\frac{\mathcal{L}_i\mathcal{L}_j}{\mathcal{L}^2}$ vanishes since it is proportional to the third correlation function of the data, which is 0 for Gaussian likelihoods. So we have that

$$F_{\alpha\beta} = \sum_{i,j,k=1}^{c} w_i w_j w_k f_{ijk}, \tag{E.11}$$

$$= \sum_{i,j=1}^{c} w_i r_i w_j r_j \lambda_{,\alpha}^T \Sigma_i^{-1}\langle\Sigma\rangle_W \Sigma_j^{-1}\lambda_{,\beta}, \tag{E.12}$$

$$= \lambda_{,\alpha}^T \langle\Sigma^{-1}\rangle_{WR}\langle\Sigma\rangle_W\langle\Sigma^{-1}\rangle_{WR}\lambda_{,\beta} = \lambda_{,\alpha}^T \tilde{\Sigma}^{-1}\lambda_{,\beta}. \tag{E.13}$$

Appendix F

Unbinned likelihood with Super-Sample Covariance

This appendix presents the calculation details to derive the unbinned likelihood with super-sample covariance presented in Section 5.2.

As follows in Takada & Spergel (2014), each x_k can be written as $x_k = N_k(1 + b_k\delta_k)$, where δ_k is the overdensity in the redshift bin k. Using the Taylor expansion of the individual Poisson distribution, according to both conditions $b_k\delta_k \ll 1$ and $b_k N_k \delta_k \ll 1$. We have

$$f(x+\delta x) = f(x)\left(1 + (\ln f)'(x)\delta x + \frac{f''(x)}{2f(x)}\delta x^2 + ...\right). \tag{F.1}$$

Each individual Poisson distribution $\mathcal{P}(\widehat{N}_k|x_k)$ can be written

$$\mathcal{P}(\widehat{N}_k|x_k) = A_{1k}A_{2k}\mathcal{P}(\widehat{N}_k|N_k), \tag{F.2}$$

where $A_{ik} = 1 + a_{ik}\delta_k + b_{ik}\delta_k^2 + o(\delta^3)$, and the coefficients a_{ik} are given by

$$\{a_{1k}, a_{2k}\} = \{-N_k b_k,\ \widehat{N}_k b_k\}, \tag{F.3}$$

$$\{b_{1k}, b_{2k}\} = \{\frac{1}{2}N_k^2 b_k^2,\ b_k^2\frac{\widehat{N}_k(\widehat{N}_k - 1)}{2}\}. \tag{F.4}$$

Then, the full likelihood at fixed overdensity is given by the form

$$\mathcal{L}(\{\widehat{N}_i\}|\{x_k\}) = A_1 A_2 \prod_{k=1}^{c}\mathcal{P}(\widehat{N}_k|N_k), \tag{F.5}$$

where the coefficient A_i are

$$A_i = \prod_{k=1}^{c} A_{ik} = 1 + \sum_{k=1}^{c} a_{ik}\delta_k + \sum_{l>k}^{c} a_{ik}a_{il}\delta_k\delta_l + \sum_{k=1}^{c} b_{ik}\delta_k^2 + o(\delta^3). \tag{F.6}$$

We can express the local product A_1A_2 in terms of power of δ, such as

$$A_1A_2 = 1 + \sum_{k,l=1}^{c}(a_{1k} + a_{2k})\delta_k + \sum_{k,l=1}^{c} a_{1k}a_{2l}\delta_k\delta_l. \tag{F.7}$$

$$+ \sum_{l>k}^{c}(a_{1k}a_{1l} + a_{2k}a_{2l})\delta_k\delta_l + \sum_{k=1}^{c}(b_{1k} + b_{2k})\delta_k^2 + o(\delta^3). \tag{F.8}$$

Now, we consider that the over-densities $\{\delta_m\}$ follow the Gaussian multivariate distribution

$$p(\{\delta_m\}) \propto \exp\left(-\frac{1}{2}\vec{\delta}^T \mathrm{S}\vec{\delta}\right). \tag{F.9}$$

where $S_{kl} = \langle\delta_k\delta_l\rangle$ is the covariance between the over-densities within the redshift bins k and l (Lacasa et al., 2018). Using the summation rule $\sum_{l>k} = \frac{1}{2}\left(\sum_{k,l=1} - \sum_{k=1}\right)$ and neglecting terms $\propto o(\delta^3)$, we have

$$\langle A_1 A_2\rangle \approx 1 + \frac{1}{2}\sum_{k,l=1}^{c}(a_{1k}a_{1l} + 2a_{1k}a_{2l} + a_{2k}a_{2l})S_{kl} + \frac{1}{2}\sum_{k=1}^{c}(2b_{1k} + 2b_{2k} - a_{1k}^2 - a_{2k}^2)S_{kk}. \tag{F.10}$$

By replacing the coefficients a_{ik} and b_{ik} in Eq. (F.10), we get

$$\langle A_1 A_2\rangle = 1 + \frac{1}{2}\sum_{k,l=1}^{c} b_k b_l[N_k N_l - 2N_k\widehat{N}_l + \widehat{N}_k\widehat{N}_l - \widehat{N}_k\delta_{kl}^K]S_{kl}. \tag{F.11}$$

We define

$$\mathcal{Q}_{kl} = b_k b_l S_{kl}[N_k N_l - 2N_k\widehat{N}_l + \widehat{N}_k\widehat{N}_l - \widehat{N}_k\delta_{kl}^K], \tag{F.12}$$

such as the complete likelihood is given by

$$\mathcal{L}(\{\widehat{N}_i\}) = \prod_{k=1}^{c}\mathcal{P}(\widehat{N}_k|N_k)\left(1 + \frac{1}{2}\sum_{k,l=1}^{c}\mathcal{Q}_{kl}\right). \tag{F.13}$$

We find that the linearized likelihood is correctly normalized, using Eq. (F.12) and defining $\langle\cdot\rangle_{\mathcal{P}}$ the average over the *standard* Poisson likelihood, then we have that the sum

$$\sum_{k=1}^{c}\sum_{n_k=0}^{+\infty}\mathcal{L}(\{\widehat{N}_k = n_k\}_{1\leq k\leq c}) \tag{F.14}$$

is given by

$$1 + \frac{1}{2}\sum_{k,l=1}^{c} b_k b_l S_{kl}[N_k N_l - 2N_k\langle\widehat{N}_l\rangle_{\mathcal{P}} + \langle\widehat{N}_k\widehat{N}_l\rangle_{\mathcal{P}} - \langle\widehat{N}_k\rangle_{\mathcal{P}}\delta_{kl}^K] \tag{F.15}$$

$$= 1 + \frac{1}{2}\sum_{k,l=1}^{c} b_k b_l S_{kl}[N_k N_l - 2N_k N_l + N_k\delta_{kl}^K + N_k N_l - N_k\delta_{kl}^K] \tag{F.16}$$

$$= 1. \tag{F.17}$$

We also have, after a few calculations, that

$$\langle\widehat{N}_i\rangle = N_i, \tag{F.18}$$

and the average product between 2 counts i and j is given by

$$\langle\widehat{N}_i\widehat{N}_j\rangle = N_i\delta_{ij}^K + N_i N_j(1 + b_i b_j S_{ij}), \tag{F.19}$$

and the average third product is given by

$$\langle\widehat{N}_i\widehat{N}_j\widehat{N}_k\rangle = N_i N_j N_k(1 + b_i b_j S_{ij} + b_j b_k S_{jk} + b_k b_i S_{ki}) \text{ for } i\neq j,\ j\neq k \text{ and } k\neq i. \tag{F.20}$$

Doing so, the variance of a single cluster count is given by $\sigma^2(\widehat{N}_i) = N_i + N_i^2 b_i^2 S_{ii}$, and the covariance between two different mass-redshift bins writes $\mathrm{Cov}(\widehat{N}_i, \widehat{N}_j) = N_i N_j b_i b_j S_{ij}$. We first consider a simpler version of Eq. (F.13) where all clusters are in a single redshift bin, which was already explored by Takada & Spergel (2014). In particular, we consider the same matter density fluctuation amplitude S_0 in a narrow redshift bin, namely setting $S_{kl} = \delta^K_{kl} S_0$ in the expression of $\mathcal{Q}_{kl}$. We get the result (see Eq. (8) in Takada & Spergel (2014))

$$\sum_{k,l=1}^{c} \mathcal{Q}_{kl} = \left[\left(\sum_{k=1}^{c} b_k (N_k - \widehat{N}_k) \right)^2 - \sum_{k=1}^{c} b_k^2 \widehat{N}_k \right] S_0. \tag{F.21}$$

We add specific indexes for both mass and redshift bins such that $[k]$ that previously denoted the k-th mass-redshift bins are now given by two indexes $[ik]$ such as the indexes ik refer to the i-th mass and k-th redshift bins. So the total number of bins c can be decomposed in two terms, namely $c = c_z \times c_m$, where c_z is the total number of redshift bins, and c_m is the total number of mass bins. Then, we have

$$\mathcal{Q}_{kl} \to \mathcal{Q}_{[im][jn]} = S_{mn} b_{im} b_{jn} [N_{im} N_{jn} - 2 N_{im} \widehat{N}_{jn} + \widehat{N}_{im} \widehat{N}_{jn} - \widehat{N}_{im} \delta^K_{imjn}]. \tag{F.22}$$

The sum of Q_{kl} terms is now given by

$$\sum_{m,n=1}^{c_z} S_{mn} \sum_{i,j=1}^{c_m} b_{im} b_{jn} [N_{im} N_{jn} - 2 N_{im} \widehat{N}_{jn} + \widehat{N}_{im} \widehat{N}_{jn}] - \sum_{m=1}^{c_z} S_{mm} \sum_{i=1}^{c_m} b_{im} b_{im} \widehat{N}_{im}. \tag{F.23}$$

We have that Eq. (F.23) can be decomposed in 4 different contributions, such that

$$Eq.\ (\mathrm{F}.23) = [1] - 2 \times [2] + [3] - [4] \tag{F.24}$$

where the coefficients [1], [2], [3] and [4] are given by

$$[1] : \sum_{m,n=1}^{c_z} S_{mn} \sum_{i,j=1}^{c_m} b_{im} b_{jn} N_{im} N_{jn} = N_{\mathrm{th}}^2 \langle Sbb \rangle_{\mathrm{th}}, \tag{F.25}$$

$$[2] : \sum_{m,n=1}^{c_z} S_{mn} \sum_{i,j=1}^{c_m} b_{im} b_{jn} N_{im} \widehat{N}_{jn} = N_{\mathrm{th}} \widehat{N}_{\mathrm{obs}} \langle Sbb \rangle_{\mathrm{th/obs}}, \tag{F.26}$$

$$[3] : \sum_{m,n=1}^{c_z} S_{mn} \sum_{i,j=1}^{c_m} b_{im} b_{jn} \widehat{N}_{im} \widehat{N}_{jn} = \widehat{N}_{\mathrm{obs}}^2 \langle Sbb \rangle_{\mathrm{obs}}, \tag{F.27}$$

$$[4] : \sum_{m=1}^{c_z} S_{mm} \sum_{i=1}^{c_m} b_{im} b_{im} \widehat{N}_{im} = \widehat{N}_{\mathrm{obs}} \langle Sb^2 \rangle_{\mathrm{obs}}, \tag{F.28}$$

where $\langle \cdot \rangle_{\mathrm{th}}$ means the theoretical average over the 4 dimensions (2 mass dimensions, 2 redshift dimensions), accounting predicted abundances N_k, $\langle \cdot \rangle_{\mathrm{obs}}$ in [3] is the 4-dimension average over the observed counts $\widehat{N}_k$ (only for 1 mass dimension and 1 redshift dimension in [4]) and $\langle \cdot \rangle_{\mathrm{th/obs}}$ for averaging 1+1 mass-redshift dimensions with N_k and 1+1 dimension with $\widehat{N}_k$.

In this paragraph, we try to make some strong approximation to capture the trend of the SSC contribution with respect to a variation in cosmological parameter. We set $\langle Sbb \rangle_{\mathrm{obs}} = \langle Sbb \rangle_{\mathrm{th}} = \langle Sbb \rangle_{\mathrm{th/obs}} = \langle Sbb \rangle$, but also that $\langle Sb^2 \rangle_{\mathrm{obs}}$ is negligible compared to the other terms (which is the case for large N_{tot}. Choosing that the total number of clusters is given $N_{\mathrm{th}} = N_0 (\Omega_m / \Omega_m^0)^\alpha$ (Bolliet et al., 2020), we get that

$$[1] - 2 \times [2] + [3] - [4] \approx \langle Sbb \rangle N_0^2 \left[\left(\frac{\Omega_m}{\Omega_m^0} \right)^\alpha - \frac{\widehat{N}_{\mathrm{obs}}}{N_0} \right]^2, \tag{F.29}$$

which corresponds to a distorted parabola $f(x) \sim (x^\alpha - 1)^2$, since we have that $\frac{\widehat{N}_{\rm obs}}{N_0} \sim 1$. The width of this parabola is given by $1/\langle Sbb\rangle N_0^2$.

Note on the Takada & Spergel approximation

We have considered previously the condition $\delta_k b_k \ll 1$ and $N_k\delta_k b_k \ll 1$ in the Taylor expansion of the Poisson distribution. Using these approximations, we found the coefficients a_{ik} in Eq. (F.3) and b_{ik} in Eq. (F.4). When the spatial region becomes smaller, the amplitude of the smoothed density perturbations increases, so $\delta_k b_k \ll 1$ may not be valid in the unbinned regime. Releasing $\delta_k b_k \ll 1$ but keeping $N_k\delta_k b_k \ll 1$, the Taylor expansion of the individual Poisson likelihood is given by

$$\mathcal{P}(\widehat{N}_k|x_k) = \mathcal{P}(\widehat{N}_k|N_k)(1+b_k\delta_k)^{\widehat{N}_k}(1 - b_k\widehat{N}_k\delta_k + \frac{1}{2}b_k^2\widehat{N}_k^2\delta_k^2). \tag{F.30}$$

Compared to Eq. (F.2) we do not expand the term $(1+b_k\delta_k)^{\widehat{N}_k}$. When considering the unbinned regime, such as $\widehat{N}_k = \{0,1\}$, the a_{ik} and b_{ik} coefficients turn to

$$\{a_{1k}, a_{2k}\} = \{-N_k b_k, (1-\delta^K_{\widehat{N}_k})b_k = \widehat{N}_k b_k\}, \tag{F.31}$$

and

$$\{b_{1k}, b_{2k}\} = \{\frac{1}{2}N_k^2 b_k^2, 0\}. \tag{F.32}$$

By replacing the coefficients a_{ik} and b_{ik} in the A_{ik} coefficients, we get

$$\langle A_1 A_2\rangle \approx 1 + \frac{1}{2}\sum_{k,l=1}^{c}(a_{1k}a_{1l} + 2a_{1k}a_{2l} + a_{2k}a_{2l})S_{kl} + \frac{1}{2}\sum_{k=1}^{c}(2b_{1k} - a_{1k}^2 - a_{2k}^2)S_{kk} \tag{F.33}$$

$$= 1 + \frac{1}{2}\sum_{k,l=1}^{c} S_{kl}b_k b_l(N_k N_l - 2N_k\widehat{N}_l + \widehat{N}_k\widehat{N}_l) + \frac{1}{2}\sum_{k=1}^{c} S_{kk}b_k^2(N_k^2 - N_k^2 - \widehat{N}_k^2) \tag{F.34}$$

$$= 1 + \frac{1}{2}\sum_{k,l=1}^{c} b_k b_l[N_k N_l - 2N_k\widehat{N}_l + \widehat{N}_k\widehat{N}_l - \widehat{N}_k^2\delta^K_{kl}]S_{kl}. \tag{F.35}$$

Compared to the Eq. (F.11), we find that the last term in the sum $-\widehat{N}_k\delta^K_{kl}$ is now $-\widehat{N}_k^2\delta^K_{kl}$. The derivation above is only valid in the unbinned regime, i.e., when $\widehat{N}_k = \{0,1\}$, that verifies $\widehat{N}_k = \widehat{N}_k^2$. In the unbinned regime, the SSC likelihood we have derived is still valid when releasing the hypothesis $\delta_k b_k \ll 1$.

Appendix G

Forecast the unbinned cluster likelihood

In this appendix, we address the Fisher forecast of the unbinned likelihood with SSC, as introduced in Section 5.2. For a general likelihood $\mathcal{L}$, we get that the Fisher information is given by

$$\mathrm{F}_{\alpha\beta} = -\mathrm{E}\left[\frac{\partial^2 \ln \mathcal{L}}{\partial\theta_\alpha \partial\theta_\beta}\right]. \tag{G.1}$$

Standard unbinned likelihood

Taking the standard Poisson unbinned likelihood $\mathcal{L}_{\rm ULC}$, we get that

$$(\ln \mathcal{L}_{\rm ULC})_{,\alpha\beta} = -(N_{\rm tot})_{,\alpha\beta} + \sum_{k=1}^{c} \widehat{N}_k (\ln \mu_k)_{,\alpha\beta} \;\; \rightarrow \;\; \mathrm{F}_{\alpha\beta} = (N_{\rm tot})_{,\alpha\beta} - \sum_{k=1}^{c} \mu_k (\ln n_h)_{,\alpha\beta} \tag{G.2}$$

where $\mu_{,\alpha\beta}$ is the second partial derivative of the total number of halos with respect to θ_α and θ_β. We took the average $\langle N_k \rangle = \mu_k$, where μ_k is the small abundance approximation. It can be re-written

$$(F_{\rm ULC})_{\alpha\beta} = (N_{\rm tot})_{,\alpha\beta} - N_{\rm tot} \langle (\ln n_h)_{,\alpha\beta} \rangle, \tag{G.3}$$

where

$$\langle (\ln n_h)_{,\alpha\beta} \rangle = \frac{1}{N_{\rm tot}} \int_{z_{\rm min}}^{z_{\rm max}} dz \int_{m_{\rm min}}^{m_{\rm max}} dm \; n_h(m,z) \frac{\partial^2 \ln n_h(m,z)}{\partial\theta_\alpha \partial\theta_\beta}. \tag{G.4}$$

If we take

$$(\ln \mu_k)_{,\alpha\beta} = \left(\frac{(\mu_k)_{,\alpha\beta}}{\mu_k} - \frac{(\mu_k)_{,\alpha} (\mu_k)_{,\beta}}{\mu_k^2} \right), \tag{G.5}$$

then we have another form of the Fisher matrix $(F_{\rm ULC})_{\alpha\beta} = \langle (\ln n_h)_{,\alpha} (\ln n_h)_{,\beta} \rangle$ (that is more similar to the Fisher matrix of the binned Poisson likelihood).

Unbinned likelihood with super sample covariance

Now, $\mathcal{L}_{\rm ULC}^{\rm SSC} = \mathcal{L}_{\rm ULC}(1 + f_{\rm SSC})$, so we get after some calculation that

$$F_{\rm ULC}^{\rm SSC} = F_{\rm ULC} + \int d^d N \mathcal{L}_{\rm ULC}^{\rm SSC} \left\{ \frac{\partial \ln \mathcal{L}_{\rm ULC}}{\partial\theta_\alpha} \frac{\partial \ln(1+f_{\rm SSC})}{\partial\theta_\beta} + \frac{\partial \ln \mathcal{L}_{\rm ULC}}{\partial\theta_\beta} \frac{\partial \ln(1+f_{\rm SSC})}{\partial\theta_\alpha} \right\} \tag{G.6}$$

$$+ \int d^d N \mathcal{L}_{\rm ULC}^{\rm SSC} \frac{\partial \ln(1+f_{\rm SSC})}{\partial\theta_\alpha} \frac{\partial \ln(1+f_{\rm SSC})}{\partial\theta_\beta}, \tag{G.7}$$

that leads to

$$F_{\rm ULC}^{\rm SSC} = F_{\rm ULC} + \int d^d N \mathcal{L}_{\rm ULC}[(\ln \mathcal{L}_{\rm ULC})_{,\alpha}(f_{\rm SSC})_{,\beta} + (\ln \mathcal{L}_{\rm ULC})_{,\beta}(f_{\rm SSC})_{,\alpha}] \tag{G.8}$$

$$+ \int d^d N \mathcal{L}_{\rm ULC}(f_{\rm SSC})_{,\alpha}(\ln(1+f_{\rm SSC}))_{,\beta}. \tag{G.9}$$

Then,

$$F_{\rm ULC}^{\rm SSC} = F_{\rm ULC} + {\rm E}_{\mathcal{L}_{\rm ULC}}[(\ln \mathcal{L}_{\rm ULC})_{,\alpha}(f_{\rm SSC})_{,\beta} + (\ln \mathcal{L}_{\rm ULC})_{,\beta}(f_{\rm SSC})_{,\alpha} + (f_{\rm SSC})_{,\alpha}(\ln(1+f_{\rm SSC}))_{,\beta}]. \tag{G.10}$$

From this formula, we find that the unbinned SSC fisher forecast requires to compute integrals with as many dimensions as clusters in the sample.

Appendix H

Study cases for the unbinned SSC likelihood

This appendix presents two study cases for the unbinned SSC likelihood presented in Chapter 5. We first study the impact of the SSC contribution on the unbinned constraints in Fig. H.1. We see that the contours are smaller as we increase the number of clusters in the analysis (from blue to red). The posteriors are slightly shifted between the standard unbinned (dashed lines) and unbinned SSC (full lines) likelihoods. The errors on parameter are affected by at most 10%.

In Fig. H.2 we test the impact of redshift ranges, by considering narrow bins, and then by considering the full redshift range (in green). We see (top figure) that the correlation is different depending on the redshifts spanned by the analysis. The joint analysis gives smaller error bars since it benefits from the information of the whole cluster sample. However, we do not see any particular impact on the recovered cosmological parameters (bottom figure).

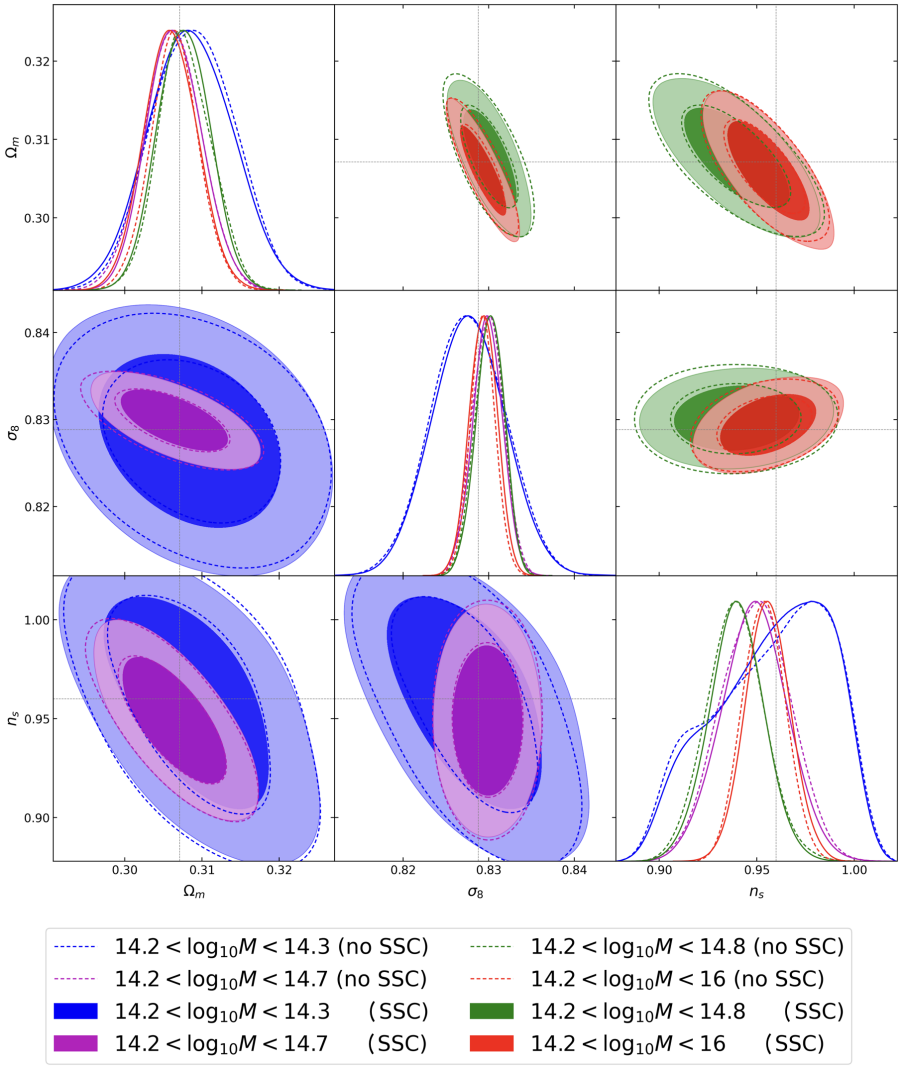

Figure H.1: Joint posterior distribution of Ω_m, σ_8 and n_s considering the unbinned standard and unbinned SSC likelihoods for different mass samples. The blue, magenta, green and red samples differ from the higher mass of the sample (starting from $\log_{10}(M) = 14.2$), that increases from blue to red.

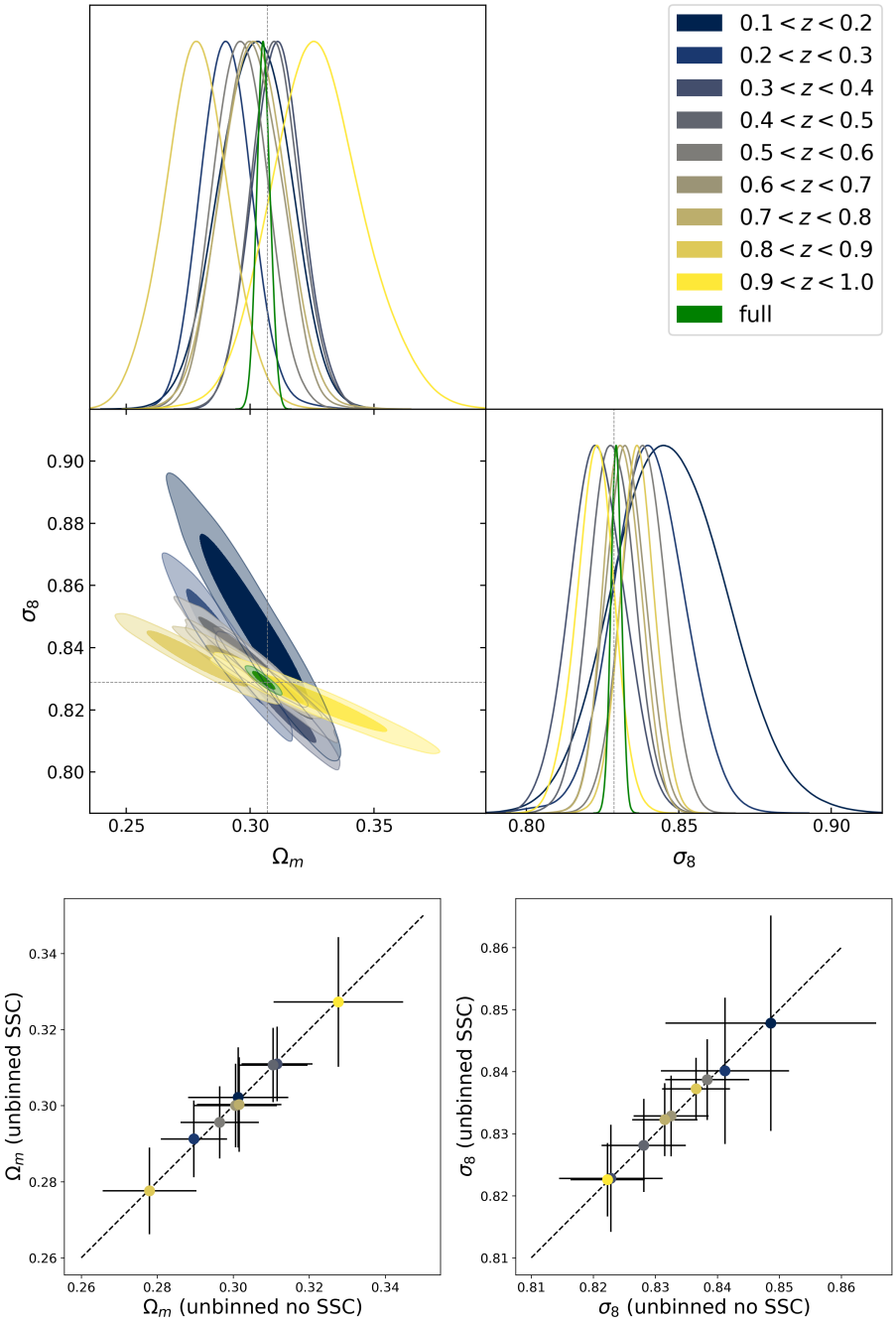

Figure H.2: Top: Joint posterior distribution of Ω_m, σ_8 considering the unbinned standard for different redshift samples. The full analysis (that comprises all redshift bins) is shown in green. Bottom: standard unbinned posterior means on Ω_m and σ_8 with respect to the unbinned SSC posterior means. The error bars correspond to the posterior variance for each case.

Appendix I

Shear multipole formalism

This appendix presents additional information to the Section 6.2 especially about the calculations of the shear multipole formalism.

Formalism

The multipole decomposition of the tangential and cross shears can be defined following the formalism developed in Bernstein & Nakajima (2009). In this paper, the *rotated* shear (in the frame of the cluster center) is defined as $\Gamma = \gamma_+ + i\gamma_\times$. It can be written as

$$\Gamma = \frac{1}{2}\sum_{m=-\infty}^{m=+\infty}(1+\delta_m)\Gamma_+^{(m)}(R)e^{im\varphi}, \tag{I.1}$$

with each multipole $\Gamma^{(m)}(R)$ verifying

$$\Gamma^{(m)}(R) = \frac{1}{(1+\delta_m)\pi}\int_0^{2\pi}\Gamma(R,\varphi)e^{-im\varphi}d\varphi. \tag{I.2}$$

From Bernstein & Nakajima (2009), the multipoles of Γ can be determined using the Poisson equation of the gravitational potential such as

$$\Gamma^{(m\geq 0)}(R) = -\kappa^{(m)} + 2(m+1)R^{-(m+2)}\alpha_m^{\rm in}, \tag{I.3}$$
$$\Gamma^{(m<0)}(R) = -\kappa^{(m)} - 2(m+1)R^{-(m+2)}\alpha_m^{\rm out}, \tag{I.4}$$

such as

$$\alpha_m^{\rm in} = \int_0^R r'^{m+1}\kappa^{(m)}(r')dr' \quad \text{and} \quad \alpha_m^{\rm out} = \int_R^{+\infty} r'^{m+1}\kappa^{(m)}(r')dr'. \tag{I.5}$$

Then, based on Bernstein & Nakajima (2009) results, we define the tangential shear γ_+ as the real part of Γ, so the tangential and cross shear multipoles follow

$$\gamma_+^{(m)} = \frac{1}{2}[\Gamma^{(m)} + (\Gamma^{(-m)})^*] \quad \text{and} \quad \gamma_\times^{(m)} = \frac{1}{2i}[\Gamma^{(m)} - (\Gamma^{(-m)})^*]. \tag{I.6}$$

For any real field f (such as the convergence and the two components of the shear field), we have that $\Im[f] = 0$, such as $(f^{(m)})^* = (f^{(m)})$. We can verify that $\gamma_+ \in \mathcal{R}$, since $(\gamma_+^{(m)})^* =$

$\frac{1}{2}[(\Gamma^{(m)})^* + (\Gamma^{(-m)})] = (\gamma_+^{(-m)})$. We define

$$\int \frac{d\varphi}{\pi}\gamma_+(r,\varphi)\cos(-m\varphi) = \mathrm{Re}[\gamma_+^{(m)}], \tag{I.7}$$

$$= \frac{1}{4}[\Gamma^{(m)} + (\Gamma^{(-m)})^* + (\Gamma^{(m)})^* + \Gamma^{(-m)}], \tag{I.8}$$

$$\text{(by identity)} = \frac{1}{2}\mathrm{Re}[\Gamma^{(m)} + \Gamma^{(-m)}] \tag{I.9}$$

$$\int \frac{d\varphi}{\pi}\gamma_+(r,\varphi)\sin(-m\varphi) = \mathrm{Im}[\gamma_+^{(m)}], \tag{I.10}$$

$$= \frac{1}{4i}[\Gamma^{(m)} + (\Gamma^{(-m)})^* - (\Gamma^{(m)})^* - \Gamma^{(-m)}], \tag{I.11}$$

$$\text{(by identity)} = \frac{1}{2}\mathrm{Im}[\Gamma^{(m)} - \Gamma^{(-m)}]. \tag{I.12}$$

The cross shear multipoles are given by

$$\int \frac{d\varphi}{\pi}\gamma_\times(r,\varphi)\cos(-m\varphi) = \mathrm{Re}[\gamma_\times^{(m)}] \tag{I.13}$$

$$= \frac{1}{2}\left(\frac{1}{2i}[\Gamma^{(m)} - (\Gamma^{(-m)})^*] - \frac{1}{2i}[(\Gamma^{(m)})^* - \Gamma^{(-m)}]\right), \tag{I.14}$$

$$= \frac{1}{4i}[\Gamma^{(m)} - \Gamma^{(-m)*} - \Gamma^{(m)*} + \Gamma^{(-m)}], \tag{I.15}$$

$$\text{(by identity)} = \frac{1}{2}\mathrm{Im}[\Gamma^{(m)} + \Gamma^{(-m)}], \tag{I.16}$$

$$\int \frac{d\varphi}{\pi}\gamma_\times(r,\varphi)\sin(-m\varphi) = \mathrm{Im}[\gamma_\times^{(m)}] \tag{I.17}$$

$$= \frac{1}{2i}\left(\frac{1}{2i}[\Gamma^{(m)} - (\Gamma^{(-m)})^*] + \frac{1}{2i}[(\Gamma^{(m)})^* - \Gamma^{(-m)}]\right), \tag{I.18}$$

$$= -\frac{1}{4}[\Gamma^{(m)} - (\Gamma^{(-m)})^* + (\Gamma^{(m)})^* - \Gamma^{(-m)}], \tag{I.19}$$

$$\text{(by identity)} = -\frac{1}{2}\mathrm{Re}[\Gamma^{(m)} - \Gamma^{(-m)}]. \tag{I.20}$$

We see that for the cross shear, $\mathrm{Re}[\gamma_\times^{(0)}] = \mathrm{Im}[\Gamma^{(0)}] = 0$, since $\mathrm{Im}[\kappa^{(0)}] = 0$ (as explored in Appendix K).

Approximation of the $m = 2$ and $m = 4$ convergence multipole

From Adhikari et al. (2015), the convergence is approximated such as it follows a power law $\kappa = AR_q^{-\eta_0}$ where the log-slope is given by

$$\eta_0 = -\frac{d\ln\kappa^{(0)}}{d\ln R}, \tag{I.21}$$

and R_q given by

$$R_q = R\left(\frac{\cos\varphi^2}{a^2} + \frac{\sin\varphi^2}{b^2}\right). \tag{I.22}$$

The term $\kappa^{(0)}$ is the monopole of the convergence. So we get that

$$\kappa = AR^{-\eta_0}\left(\frac{\cos\varphi^2}{a^2}+\frac{\sin\varphi^2}{b^2}\right)^{-\eta_0/2}, \tag{I.23}$$

$$= A(R/a)^{-\eta_0}(\frac{1+q^2}{2q^2})^{-\eta_0}\left(1-\epsilon\cos 2\varphi\right)^{-\eta_0/2}, \tag{I.24}$$

$$= \kappa^{(0)}+\kappa^{(2)}\cos 2\varphi + o(\epsilon^2), \tag{I.25}$$

where we define $\epsilon = \frac{1-q^2}{1+q^2}$ and $\kappa^{(2)} = \frac{\epsilon}{2}\eta_0\kappa^{(0)}$. When considering an offset angle φ_0, the $m=2$ multipole is given by $\kappa^{(2)'} = \kappa^{(2)}(\varphi_0 = 0)\cos(2\varphi_0)$. For the $m=4$ multipole, we have that

$$\kappa \propto 1+\frac{\epsilon}{2}\eta_0\cos 2\varphi + 2\epsilon^2\frac{\eta_0}{2}\left(\frac{\eta_0}{2}+1\right)(\cos 2\varphi)^2, \tag{I.26}$$

$$\propto 1+\frac{\epsilon}{2}\eta_0\cos 2\varphi + 2\epsilon^2\delta\frac{\cos 4\varphi+1}{2}, \tag{I.27}$$

$$\propto 1+\epsilon^2\delta+\frac{\epsilon}{2}\eta_0\cos 2\varphi+\epsilon^2\delta\cos 4\varphi, \tag{I.28}$$

$$= \kappa^{(0)}+\kappa^{(2)}\cos 2\varphi+\kappa^{(4)}\cos 4\varphi + o(\epsilon^2). \tag{I.29}$$

After identifying the multipoles, we get

$$\kappa^{(0)} = A_R(1+\epsilon^2\delta), \tag{I.30}$$

$$\kappa^{(2)} = A_R\epsilon\eta_0, \tag{I.31}$$

$$\kappa^{(4)} = A_R\epsilon^2\delta, \tag{I.32}$$

where

$$A_R = A(R/a)^{-\eta_0}(\frac{1+q^2}{2q^2})^{-\eta_0} \quad \text{and} \quad \delta = \frac{\eta_0}{2}\left(\frac{\eta_0}{2}+1\right). \tag{I.33}$$

Non-weak lensing corrections

The galaxy ellipticities probe the reduced shear that is given by $g = \gamma/(1-\kappa)$. When making the Taylor expansion at first order in κ we get $g_{+/\times} = \gamma_{+/\times}+\gamma_{+/\times}\kappa+o(\kappa^2)$. Then, the multipole $g^{(m)}_{+/\times}$ writes

$$g^{(m)}_{+/\times} = \gamma^{(m)}_{+/\times} + (\gamma_{+/\times}\kappa)^{(m)}. \tag{I.34}$$

We have that

$$(\gamma_{+/\times}\kappa)^{(m)} = \frac{1}{2(1+\delta_m)}\sum_{p=-\infty}^{p=+\infty}(1+\delta_p)(1+\delta_{p-m})\gamma^{(p)}_{+/\times}\kappa^{(m-p)}. \tag{I.35}$$

We can simplify the above equation by considering only the terms where the monopole of each component appears, thus neglecting the product of multipoles where each one has simultaneously $m \neq 0$. We get

$$(\gamma_{+/\times}\kappa)^{(m)} \approx \frac{1}{2}\left(\gamma^{(m)}_{+/\times}\kappa^{(0)}+\gamma^{(0)}_{+/\times}\kappa^{(m)}\right). \tag{I.36}$$

For the real part of the multipole, we have for instance

$$\mathrm{Re}[(\gamma_{+/\times}\kappa)^{(m)}] \approx \frac{1}{2}(\mathrm{Re}[\gamma^{(m)}_{+/\times}]\mathrm{Re}[\kappa^{(0)}]+\mathrm{Re}[\gamma^{(0)}_{+/\times}]\mathrm{Re}[\kappa^{(m)}]-\mathrm{Im}[\gamma^{(0)}_{+/\times}]\mathrm{Im}[\kappa^{(m)}]), \tag{I.37}$$

$$= \frac{1}{2}(\mathrm{Re}[\gamma^{(m)}_{+/\times}]\mathrm{Re}[\kappa^{(0)}]+\mathrm{Re}[\gamma^{(0)}_{+/\times}]\mathrm{Re}[\kappa^{(m)}]). \tag{I.38}$$

Let us note that for spherical mass distribution, the shear and convergence multipoles with $m \neq 0$ vanish, so we have that the $m = 0$ weak lensing correction is given $\mathrm{Re}[\gamma^{(0)}_{+/\times}]\mathrm{Re}[\kappa^{(0)}]$ as it is the case in Mandelbaum et al. (2006); Johnston et al. (2007).

Appendix J

Triaxial ellipsoid

This appendix presents a few calculation details associated with the projected matter distribution of a triaxial ellipsoid cluster.

Monopole of an elliptical projected matter distribution

The three dimensional matter density around the cluster is given by

$$\rho_{abc}(\vec{r}) = \rho_{\rm sph}(|\mathcal{A}\cdot\vec{r}|), \tag{J.1}$$

so it is defined as a 3D deformation of a spherical matter density distribution. We define the deformation matrix

$$\mathcal{A} = \begin{pmatrix} 1/a_x & 0 & 0 \\ 0 & 1/a_y & 0 \\ 0 & 0 & 1/a_z \end{pmatrix}. \tag{J.2}$$

By integrating along the line of sight (along the z axis) and by mass conservation, we have that the projected surface density is given in terms of the two-dimensional halo-matter power spectrum of the spherical projected mass density $P_{\rm hm}(k_{2d})$ such as

$$\Sigma(R,\varphi) = \frac{(a_x a_y)^{-1}\rho_m^0}{(2\pi)^2}\int d^2\vec{k}_{2d} P_{\rm hm}(k_{2d}) e^{-i\vec{k}_{2d}\mathcal{A}\cdot\vec{R}}, \tag{J.3}$$

with

$$\vec{k}_{2d}\mathcal{A}\cdot\vec{R} = k_{2d}R\left[\frac{1}{a_x}\cos\phi\cos\varphi + \frac{1}{a_y}\sin\phi\sin\varphi\right], \tag{J.4}$$

$$= k_{2d}R[\delta\cos(\phi-\varphi) + \mu\cos(\phi+\varphi)] \tag{J.5}$$

and φ and ϕ are respectively the polar angles of $\vec{k}_{\rm 2d}$ and $\vec{R}$. We define δ and μ such as

$$\delta = \frac{1/a_x - 1/a_y}{2} \quad \text{and} \quad \mu = \frac{1/a_x + 1/a_y}{2}. \tag{J.6}$$

We get, after making the azimuth average over the polar angle ϕ that the projected mass density is given by

$$\Sigma_{a_x a_y}(R) = \frac{(a_x a_y)^{-1}\rho_m^0}{(2\pi)}\int k_{2d}\ dk_{2d}\ P_{\rm hm}(k_{2d}) I[k_{2d}R,\delta,\mu], \tag{J.7}$$

where we have defined

$$I[x,\delta,\mu] = \int \frac{d\phi d\varphi}{(2\pi)^2}\exp\{-ix[\delta\cos(\phi-\varphi)+\mu\cos(\phi+\varphi)]\} = J_0(x\delta)J_0(x\mu). \tag{J.8}$$

The monopole of the elliptical projected mass density is given by

$$\Sigma_{a_x a_y}(R) = \frac{(a_x a_y)^{-1}\rho_m^0}{f_K(w)^2}\int \frac{ldl}{(2\pi)} P_{\rm hm}(k_l) J_0(l\theta\delta) J_0(l\theta\mu). \tag{J.9}$$

In the most simple case where the cluster is oblate/prolate shaped with major/minor axis aligned along the line of sight, the 2 factors a_x and a_y are equal but $\neq 1$. Then, using that $J_0(0) = 1$, we get the simple form

$$\Sigma_{a_x}(R) = \frac{a_x^{-2}\rho_m^0}{f_k(w)^2}\int \frac{ldl}{(2\pi)} P_{\rm hm}(k_l) J_0(l\theta a_x^{-1}) = a_x^{-2}\Sigma_{\rm sph}(a_x^{-1}R). \tag{J.10}$$

In this particular case, we can derive a simple form for the excess surface density that writes $\Delta\Sigma_{a_x}(R) = a_x^{-2}\Delta\Sigma(a_x^{-1}R)$.

Projection of an elliptical three-dimensional matter distribution

We studied above the projected matter density of an elliptical 3D density distribution with no rotation angle (we did not account for any rotation angle of the ellipsoid in the three-dimensional system of coordinates). In the more general case, we have

$$\rho(\vec{r}) = \rho_{\rm sph}(|\Lambda_{\rm 3d} R_{\rm 3d}\vec{r}|) = \rho_{\rm sph}[(\vec{r}^T R_{\rm 3d}^T \Lambda_{\rm 3d}^2 R_{\rm 3d}\vec{r})^{1/2}] \tag{J.11}$$

where we define the matrix $S = \Lambda_{\rm 3d} R_{\rm 3d}$ where $\Lambda_{\rm 3d} = {\rm diag}[1/a_x, 1/a_y, 1/a_z]$ and R is the rotation matrix with the 3 Euler angles in the three-dimensional reference frame. The projected matter density is given by

$$\Sigma(R,\varphi) = \Sigma_s[(\vec{r}^T R_{\rm 2d}^T \Lambda_{\rm 2d}^2 R_{\rm 2d}\vec{r})^{1/2}] = \Sigma_s[|\Lambda_{\rm 2d} R_{\rm 2d}\vec{r}|], \tag{J.12}$$

i.e. a two-dimensional rotation and deformation of the spherical projected matter density with effective rotation/deformation matrix $\Lambda_{\rm 2d}$ in a 2D space.

Proof:

The projected surface density writes

$$\Sigma(R,\varphi) = \int dy \rho(x,y,z), \tag{J.13}$$

$$= \int dy \int d^3k P(k) e^{-ik_j S_{jk} r_k}, \tag{J.14}$$

$$= \int d^3k P(k) \int dy e^{-i(k_j S_{jk'} r_{k'} + k_y R_{jy} y)} \text{ with } k' \in [1,2], \tag{J.15}$$

$$= \int d^3k P(k) e^{-ik_j S_{jk'} r_{k'}} \delta^{(1)}(S_{yj}^T k_j). \tag{J.16}$$

Giving $u = S^T k$, we get

$$\Sigma(R,\varphi) = \frac{1}{|(S^T)^{-1}|}\int d^3u \delta^{(1)}(u_y) P[S^{(T,-1)}u] \exp\{-iS_{jl}^{(T,-1)} u_l S_{jk'} r_{k'}\}, \tag{J.17}$$

$$= \frac{1}{|(S^T)^{-1}|}\int d^3u \delta^{(1)}(u_y) P[S^{(T,-1)}u] \exp\{-iu_l S_{lj}^{(T,-1,T)} S_{jk'} r_{k'}\}. \tag{J.18}$$

We define $\widehat{S}_{lk'} = S_{lj}^{(T,-1,T)} S_{jk'}$, so we get

$$\Sigma(R,\varphi) = \frac{1}{|S^{(T,-1)}|}\int d^3u \delta^{(1)}(u_y) P[|S^{(T,-1)}u|] \exp\{-iu_l \widehat{S}_{lk'} r_{k'}\}, \tag{J.19}$$

$$= \frac{1}{|S^{(T,-1)}|}\int d^2u' P[|S^{(T,-1)}u'|] \exp\{-iu_{l'} \widehat{S}_{l'k'} r_{k'}\}, \tag{J.20}$$

where $u' = [u_1, u_2, 0]$. Setting $|S^{(T,-1)}u'| = |Fu'| = |F^{2\mathrm{d}}u'|$, we get

$$\Sigma(R,\varphi) = \frac{1}{|S^{(T,-1)}|}\int d^2u' P[|F^{2\mathrm{d}}u'|] e^{-iu_{l'}\widehat{S}_{l'k'}r_{k'}}. \tag{J.21}$$

Setting $x = F_{2\mathrm{d}}u'$, we get

$$\Sigma(R,\varphi) = \frac{1}{|S^{(T,-1)}|}\frac{1}{|(F_{2\mathrm{d}})^{-1}|}\int d^2x P[|x|] e^{-ix_{m'}[F_{2\mathrm{d}}^{-1}]_{l'm'}\widehat{S}_{l'k'}r_{k'}}. \tag{J.22}$$

Using the Singular value decomposition, the 2-dimensional matrix $(F_{2\mathrm{d}}^{-1})^T\widehat{S}$ is given by

$$([F^{2\mathrm{d}}]^{-1})^T\widehat{S} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \frac{1}{(a')^2} & 0 \\ 0 & \frac{1}{(b')^2} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{J.23}$$

$$= R_{2\mathrm{d}}^T \Lambda_{2\mathrm{d}}^2 R_{2\mathrm{d}}. \tag{J.24}$$

So we get that

$$\Sigma(R,\varphi) \propto \Sigma_{\mathrm{sph}}[(\vec{r}^T R_{2\mathrm{d}}^T \Lambda_{2\mathrm{d}}^2 R_{2\mathrm{d}} \vec{r})^{1/2}] = \Sigma_{\mathrm{sph}}[|\Lambda_{2\mathrm{d}} R_{2\mathrm{d}} \vec{r}|]. \tag{J.25}$$

Appendix K

Cross shear monopole

In this appendix, we address the calculation of the cross shear monopole, as introduced in Chapter 6.

We present a derivation of the cross shear monopole in real space for an off-centered observer measuring the shear signal around a single lens. Let (θ, φ) be the 2 polar coordinates in the reference frame $\mathcal{R}$, and (θ', φ') the position in $\mathcal{R}'$, shifted to the right from $\mathcal{R}$ by an angle $\theta_{\rm off}$ (then $\varphi_{\rm off} = \varphi'_{\rm off} = 0$). Let be an axisymmetric mass distribution at $\theta = 0$ in $\mathcal{R}$, generating a complex shear at (θ, φ) given by $\gamma(\theta, \varphi) = \gamma(\theta)e^{2i\varphi}$ with $\gamma(\theta)$ real. The complex shear at (θ, φ) will be the same seen in $\mathcal{R}'$ at a different location (θ', φ'). Then

$$\gamma(\theta, \varphi) = \gamma(\theta', \varphi'). \tag{K.1}$$

Let be an observer in $(\theta', \varphi') = (0, 0)$ that computes the cross shear relative to their position, then

$$\begin{aligned} \gamma_\times(\theta', \varphi') &= -\mathrm{Im}[\gamma(\theta', \varphi')e^{-2i\varphi'}], &\text{(K.2)}\\ &= -\sin(2[\varphi - \varphi'])\gamma(\theta), &\text{(K.3)} \end{aligned}$$

due to an axisymmetric lens. We need to calculate (θ, φ) in terms of the shifted position (θ', φ') and the off-centering angle $\theta_{\rm off}$. We get, after a few calculations

$$\begin{aligned} \theta &= |\vec{\theta}_{\rm off} + \vec{\theta'}| &\text{(K.4)}\\ &= (\theta'^2 + \theta^2_{\rm off} + 2\theta'\theta_{\rm off}\cos[\varphi'])^{1/2}, &\text{(K.5)} \end{aligned}$$

and

$$\tan\varphi = \frac{\theta'\sin(\varphi')}{([\theta_{\rm off} + \theta'\cos(\varphi')]^2)^{1/2}}. \tag{K.6}$$

To evaluate the averaged cross shear for the shifted reference frame, we compute the integral

$$\begin{aligned} \bar{\gamma}_\times(\theta') &= \int_{-\pi}^{\pi} \frac{d\varphi'}{2\pi}\gamma_\times(\theta', \varphi'), &\text{(K.7)}\\ &= -\int_{-\pi}^{\pi} \frac{d\varphi'}{2\pi}\sin(2[\varphi - \varphi'])\gamma(\theta). &\text{(K.8)} \end{aligned}$$

It's the integral of a function over the period 2π. the sin is an odd function. Let's evaluate the parity of $\gamma(\theta)$ and $\sin(2[\varphi - \varphi'])$ relative to φ'. Let's do the transformation

$$\varphi' \to -\varphi', \tag{K.9}$$

giving

$$\sin(2[\varphi(\varphi') - \varphi']) \to \sin(2[\varphi(-\varphi') + \varphi']). \tag{K.10}$$

We see that

$$\varphi(\varphi') = \arctan\left(\frac{\theta' \sin(\varphi')}{([\theta_{\rm off} + \theta' \cos(\varphi')]^2)^{1/2}}\right). \tag{K.11}$$

Then

$$\varphi(-\varphi') = \arctan\left(\frac{\theta' \sin(-\varphi')}{([\theta_{\rm off} + \theta' \cos(-\varphi')]^2)^{1/2}}\right), \tag{K.12}$$

$$= -\arctan\left(\frac{\theta' \sin(\varphi')}{([\theta_{\rm off} + \theta' \cos(\varphi')]^2)^{1/2}}\right), \tag{K.13}$$

$$= -\varphi(\varphi'), \tag{K.14}$$

since arctan is an odd function. Then, we have

$$\sin(2[\varphi(-\varphi') + \varphi']) = \sin(2[-\varphi(\varphi') + \varphi']), \tag{K.15}$$

$$= -\sin(2[\varphi(\varphi') - \varphi']). \tag{K.16}$$

We find that $\sin(\varphi' - \varphi)$ is an odd function. For $\gamma(\theta(\varphi'))$, we get the result

$$\gamma(\theta(\varphi')) \to \gamma(\theta(-\varphi')) = \gamma(\theta(\varphi')). \tag{K.17}$$

Then, the product of the 2 functions is an odd function. The integration from $-\pi$ to π gives 0, implying that the cross shear is still 0. Then

$$\bar{\gamma}_\times(\theta'|\theta_{\rm off}) = 0 \ \forall \ \theta_{\rm off}. \tag{K.18}$$

This result does not depend on the polar angle of $\vec{\theta}_{\rm off}$, since the derivation can be performed in any rotated reference frame letting $\varphi_{\rm off} = 0$, and the polar averaging vanishes this dependence. It is also the case for the cross shear for a set of multiple masses distributed in the sky, where the computation of the cross shear is performed relative to any point.

We can do similar calculations in Fourier space, we can show that the two components $\gamma_{1/2}$ can be written in terms of the Fourier transform of the effective convergence such as (Jeong et al., 2009)

$$\gamma_1(\vec{\theta}) = \int \frac{d^2\vec{l}}{(2\pi)^2} \kappa(\vec{l}) \cos(2\beta) e^{i\vec{l}.\vec{\theta}}, \tag{K.19}$$

$$\gamma_2(\vec{\theta}) = \int \frac{d^2\vec{l}}{(2\pi)^2} \kappa(\vec{l}) \sin(2\beta) e^{i\vec{l}.\vec{\theta}}, \tag{K.20}$$

where β is the angle between $\vec{l}$ and the unit vector $\vec{e}_1$ in real space, i.e. $\vec{l} = l(\cos\varphi, \sin\varphi)$. Using Eq. (K.19) and Eq. (K.20), the tangential shear expresses as

$$\gamma_+(\vec{\theta}) = -\int \frac{d^2\vec{l}}{(2\pi)^2} \kappa(\vec{l}) \cos[2(\beta - \phi)] e^{i\vec{l}.\vec{\theta}}. \tag{K.21}$$

Taking the average over the polar angle ϕ of $\vec{\theta} - \vec{\theta}_0$, the mean tangential shear is given by

$$\bar{\gamma_+}(\theta) = \int_0^{2\pi} \frac{d\phi}{2\pi} \gamma_t(\vec{\theta}), \tag{K.22}$$

$$= \int \frac{d^2\vec{l}}{(2\pi)^2} \kappa(\vec{l}) J_2(l\theta), \tag{K.23}$$

where J_2 is the second order Bessel function of the first kind. The cross shear can be obtained by replacing "cos" with "sin" in Eq. (K.21), and then vanishes due to parity. The averaged cross shear is always 0, for any matter distribution.