Weak Lensing & Substructure

Dissertation

zur
Erlangung des Doktorgrades (Dr. rer. nat.)

der
Mathematisch-Naturwissenschaftlichen Fakultät

der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Emilio Pastor Mira

aus
Alcoy

Bonn Dez., 2010

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Peter Schneider
2. Gutachter: Prof. Dr. Cristiano Porciani

Tag der Promotion: 21. 04. 2011
Erscheinungsjahr: 2011

A tots i cadascú dels profesors he tingut; a totes les persones que de alguna manera m'han ensenyat alguna cosa en ma vida; a tots els que em van ajudar a poder tindre algún orgull en la meua ànima, a tots ells està dedicada aquesta tesi. I davant de totes aquestes persones anomenades, com els millors i els més antics, li dedique a la meua tesi a:

María Dolores Mira Conejero, i a
Emilio Pascual Pastor Martínez.

Contents

1. Introduction

This thesis is a feasibility study on applying galaxy-galaxy lensing to galaxies inside clusters. This method uses gravitational lensing to estimate the average mass profile of a sample of galaxies. Galaxy-galaxy lensing requires large cluster samples, which will be available in the near future.

The main goal of the thesis is to analyze what can be learned about this specific class of galaxies using future surveys. For our purpose, we use state-of-the-art cosmological and gravitational lensing simulations. This allows us to make predictions and arrive to conclusions in a perfectly controlled environment. We will use one of the largest cosmological simulations by the time this thesis is written, the Millennium Simulation, and ray-tracing simulations of the gravitational lensing effect within it.

We shall define a precise method to apply galaxy-galaxy lensing to galaxies inside clusters. Our intention is also to predict the expected signal-to-noise ratios, study how to optimize the information content and avoid any foreseeable systematic effect or bias. We also explore different possible analyses that could be compared to our theoretical predictions and models.

Galaxy clusters are the most massive bound structures known, with estimated masses of more than 10^{14} solar masses and hundreds of individual galaxies. They are so large that they are often assumed to be fair tracers of the composition of the whole Universe.

Their dynamics and composition offer a great amount of cosmological information. Also, the way in which large galaxy concentrations form and evolve is of great interest for current cosmological models. The method that we analyze in this thesis can shed some light on all these topics, and combined with galaxy evolution models, it can help as well to understand how the mass of galaxies affects their observables like luminosity, color etc.

The thesis is organized as follows. The first three chapters are a broad overview of the basic concepts necessary to develop our work. The rest contain the original work that we have produced.

- We start in Chapter 2 with an introduction to cosmology. We list the essential points of the current cosmological concordance model which are indispensable for any work in cosmology.

- Chapter 3 is a brief overview about structure growth in the Universe. We illustrate the theory, and we describe the cosmological simulation on which this thesis is based.

- In Chapter 4 we describe gravitational lensing. Gravitational lensing is the tool that we use to get the mass profiles of galaxies inside clusters. We also describe in here our gravitational lensing simulations.

- The original work and center of this thesis is presented in Chapter 5. Here we make a thorough description of the method we are going to test.

- We use simulations to make predictions. For a deeper understanding of our results, we study in detail the output of the simulations in Chapter 6. We also define the limits of our measurements.

- In Chapter 7 we use the concepts and conclusions developed before to compute the predictions we aim for. Here we analyze what is to be expected in different lensing surveys concerning weak lensing on galaxies in clusters. We also suggest different experiments to characterize the mass profiles of satellite galaxies.

Finally we conclude with the summary and conclusions.

1.1 Mathematical notation

To aid the reader and avoid ambiguity on the mathematical expressions, we present the common notation we maintain throughout the thesis.

- Important constants and units (Particle Data Group et al. 2008).
  - ⋆ Solar Mass: $1\,\mathrm{M}_\odot = (1.98842 \pm 0.00001) \times 10^{30}$ kg.
  - ⋆ Megaparsec: $1\,\mathrm{Mpc} = 3.0856776 \times 10^{22}$ m.
  - ⋆ The Hubble constant H_0 essential in the definition of cosmological quantities, is often written like $H_0 = 100\,h \quad \mathrm{km\,s^{-1}\,Mpc^{-1}}$. With the constant h we express the lack of a precise knowledge of its value.
  - ⋆ The speed of light $299\,792\,458\ \mathrm{m\,s^{-1}}$ (exact) is always represented with the letter c.
  - ⋆ The gravitational constant $(6.6742867 \pm 0.000010) \times 10^{-11}\ \ \mathrm{m^3\,kg^{-1}\,s^{-2}}$ is represented with the letter G.

- Vectorial quantities are noted in bold letters $\mathbf{x}$, both in two or three dimensions, otherwise the quantity is a scalar.

- Partial derivatives with respect to time, noted with t, will be abbreviated using a dot and double derivatives with a double dot, for instance:

$$\frac{\partial f(x)}{\partial t} = \dot{f}(x)\,, \quad \frac{\partial^2 f(x)}{\partial t^2} = \ddot{f}(x)\,. \tag{1.1}$$

  All other derivatives, including reparametrization of temporal derivatives, will be explicitly written.

- We use Einstein summation convention:

$$a_i \cdot b^i = \sum_i^N a_i \cdot b^i. \tag{1.2}$$

- The Fourier transform $\mathcal{F}(f)$ of a function $f(x)$, is denoted with a tilde $\sim$.

$$\mathcal{F}(f(x)) = \tilde{f}(k). \tag{1.3}$$

2. Basic concepts of cosmology

The problems treated in cosmology require to redefine the way we measure physical quantities such as time and distance. The dynamical evolution of the Universe is also parameterized in a specific way. This chapter is devoted to offer a general perspective on cosmology. It also summarizes a broad variety of concepts, offering an overview of the theoretical background. A deeper discussion about what is presented here can be found in Schneider (2006a), Peacock (1999), Weinberg (1972) and Misner et al. (1973).

Cosmology is the branch of physics that describes the whole Universe as a single entity. In order to deal with such a titanic task, we proceed in a distinct manner. First we assume that the physical laws are the same throughout the Universe. Second, in order to work on cosmological scales, we must use simplified phenomenological models for many small scale physics. And third, a principle, the *cosmological principle* is assumed: the Universe is homogeneous and isotropic; i.e. the Universe looks the same in any direction and from every point. Although this principle is an idealization, it suits our purposes well. The assumption of isotropy agrees, at least up to certain extent, with large scale observations (2dFGRS, Colless et al. 2001).

In cosmology, most of the times we can only observe and analyze. There is no experiment reproducing a studied phenomenon that we can set up. Our investigations are limited then to quantify the likelihood of the cosmological observables under different competing models. This fact imprints every idea in cosmology.

2.1 Gravitation

Over large distances, the only known interactions that have a measurable effect are electromagnetism and gravity, the others are short-ranged. However, on cosmological scales, the Universe is neutral. Except at early stages of the evolution of Universe, gravitation is the most relevant interaction at large scales and defines many characteristics of the Universe. The current accepted theory of gravitation is Einstein's theory of General Relativity (GR), which describes the gravitational interaction in terms of the space-time geometry.

We assume that the space-time of the Universe is well represented by a manifold. This manifold carries a metric which describes how distances and angles are related. The metric is represented by the tensor $g_{\alpha\beta}$, where α and β refer to one of the four possible space-time coordinates. The space-time distance element $\mathrm{d}s$ as a function of the coordinates x^α is therefore:

$$\mathrm{d}s^2 = g_{\alpha\beta}\,\mathrm{d}x^\alpha \mathrm{d}x^\beta. \tag{2.1}$$

In GR, the trajectories of particles are a consequence of the metric.

General Relativity links the geometry of the space-time (or metric) with its energy content, described by the energy-momentum tensor $T_{\alpha\beta}$. The tensor $T_{\alpha\beta}$ describes the

different energy-momentum fluxes α along the four directions β. The fundamental equation in GR relates the Einstein tensor $G_{\alpha\beta}$, which is a function of the metric tensor, its first and second derivatives with $T_{\alpha\beta}$. Einstein's field equation reads

$$G_{\alpha\beta} + \Lambda g_{\alpha\beta} = -\frac{8\pi G}{c^4} T_{\alpha\beta}. \tag{2.2}$$

The inclusion of the term $\Lambda g_{\alpha\beta}$ allows solutions for an Universe with an accelerated expansion. Our current observations favor such an expansion for the Universe at the present cosmological time. This last statement will be discussed further in Section 2.2.

2.2 Fundamental concepts

Having outlined the basic concepts of gravity we shall give a brief overview of the set of fundamental parameters with which we describe the metric and energy-momentum tensors. These parameters also allow a convenient description of distances and time.

In 1929 E. Hubble found that galaxies are receding from us, the further away the faster. This conclusion was derived by measuring the redshift of galaxies as a function of its distance from us. The relative motion between a light emitter and the observer changes the measured wavelength of the light. This effect is called redshift. If we write it as a function of the observed wavelength λ and the wavelength λ_0 measured in the frame of reference of the emitter, it reads

$$z = \frac{\lambda - \lambda_0}{\lambda_0}. \tag{2.3}$$

The assumption of the cosmological principle leads to interpret this phenomenon as a homogeneous isotropic expansion.

The framework of cosmology is therefore a homogeneous, isotropic Universe that expands with time. Due to this fact, for convenience, we define *comoving coordinates* $\boldsymbol{\chi}$. The relation between physical coordinates denoted with $\mathbf{r}$, and comoving coordinates is given by

$$\mathbf{r}(t) = a(t) \cdot \boldsymbol{\chi}, \tag{2.4}$$

where the *scale factor* $a(t)$ explicitly expresses the expansion. By convention, we fix the value of $a(t)$ at the present cosmological time t_0, such that $a(t_0) = 1$. The scale factor at the time the light was emitted and its redshift are related via

$$z = \frac{1}{a} - 1. \tag{2.5}$$

The scale factor and the redshift are indistinctly used. They are needed to compute distances and times in cosmology. General Relativity predicts that information cannot travel faster than light. As we observe an emission from a distant point, we are always looking backwards in time. Therefore, temporal and radial distance are linked. The redshift or the scale factor estimates the cosmological time at which an event occurred, and its distance from us. Since they are monotone functions of time, at least in the current cosmological model, they are often used directly as temporal parameters. The exact functional dependence of time with respect to the scale factor is presented in Sect. 2.4.

The evolution of the expansion at a given moment is quantified with the expansion rate, or *Hubble rate*

$$H(t) = \frac{\dot{a}(t)}{a(t)}. \tag{2.6}$$

The Hubble constant $H_0 = H(t_0)$, is the value of the Hubble rate for the local Universe at the present time t_0. It is one of the fundamental parameters of the currently accepted cosmological model.

As mentioned previously, the Universe is mainly driven by gravity, and gravity is defined by the energy-momentum content. Assuming the validity of the cosmological principle, the energy content of the Universe can be described in the same way as that of a homogeneous and isotropic fluid, which combines all possible energy components. The energy-momentum tensor in a static universe of such a fluid is in Cartesian coordinates at its rest frame

$$T_{\alpha\beta} = \begin{pmatrix} -\rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \tag{2.7}$$

We have denoted the energy density of the fluid with ρ and the pressure with p. The general form for the tensor, given the metric tensor $g_{\alpha\beta}$ is

$$T_{\alpha\beta} = \left(\rho + \frac{p}{c^2}\right) g_{\alpha\nu}\, g_{\beta\mu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} + p\, g_{\alpha\beta} \quad \text{where} \quad \mathrm{d}\tau = \sqrt{g_{\alpha\beta}\, \mathrm{d}x^\alpha \mathrm{d}x^\beta}. \tag{2.8}$$

The term $\frac{\mathrm{d}x^\gamma}{\mathrm{d}\tau}$ is the four dimensional velocity and the parameter τ is called the proper time. The three components considered in cosmology that contribute to the energy-momentum tensor are: matter, radiation, and vacuum energy. It is believed that nowadays around 70% of the energy budget of the Universe is vacuum energy, called *Dark Energy*; 25% corresponds to pressureless matter of yet unknown nature, not detected before in other scientific fields; and only 5% corresponds to matter as we know it (atomic nuclei and electrons). The current contribution of radiation is negligible, although it was important at early times in the evolution of the Universe.

The temporal dependence of pressure and energy density for each specie can be obtained as follows. We can write the first law of thermodynamics $\mathrm{d}U = -p\,\mathrm{d}V$, in terms of energy density and pressure, in comoving coordinates

$$\mathrm{d}(\rho c^2 a^3\, \chi^3) = -p\, \mathrm{d}(a^3\, \chi^3)\,, \tag{2.9}$$

and find solutions for $\rho(t)$ and $p(t)$. For pressureless matter, $p_\mathrm{m} = 0$, we obtain

$$\rho_\mathrm{m}(t) = \rho_{\mathrm{m}0}\; a^{-3}(t). \tag{2.10}$$

For radiation, out of relativistic considerations we know that $p_\mathrm{r} = \rho_\mathrm{r}\, c^2/3$, and therefore

$$\rho_\mathrm{r}(t) = \rho_{\mathrm{r}0}\; a^{-4}(t). \tag{2.11}$$

Finally, we consider vacuum energy density to be constant with time. This last assumption implies

$$-\rho_\Lambda\, c^2 = p_\Lambda. \tag{2.12}$$

The parameters $\rho_{\mathrm{x}0}$ are the different densities measured at the present epoch. These are rescaled into the *density parameters*:

$$\Omega_{\mathrm{r}} := \frac{\rho_{\mathrm{r}0}}{\rho_{\mathrm{crit}}}; \qquad \Omega_{\mathrm{m}} := \frac{\rho_{\mathrm{m}0}}{\rho_{\mathrm{crit}}}; \qquad \Omega_{\Lambda} := \frac{\rho_{\Lambda}}{\rho_{\mathrm{crit}}} = \frac{\Lambda}{3H_0^2}; \tag{2.13}$$

using the so-called *critical density*

$$\rho_{\mathrm{crit}} := \frac{3H_0^2}{8\pi G}. \tag{2.14}$$

Due to the different dependence with respect to the scale factor, the relative contribution of each species to the total energy density changes with time. In the history of the Universe, the energy density is dominated at different epochs by the different components.

Obtaining the values of the parameters described in this section with precision is currently a main goal of cosmology. Fixing them shall bring a quantitative description of the Universe and a much deeper understanding of physics.

2.3 Robertson-Walker metric & Friedmann equations

Once we specified the energy-momentum tensor, we can use GR to obtain dynamical equations for its evolution. We shall sketch how they are obtained. In principle, Einstein's field equation has no general analytical solution. In fact, only in a reduced number of cases the metric can be found given a energy-momentum tensor. For the case of an expanding homogeneous and isotropic Universe, however, there is one: the so-called Robertson-Walker metric. The Robertson-Walker metric reads

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)\left[\mathrm{d}\chi^2 + f_K^2(\chi)\Big(\mathrm{d}\theta^2 + \sin^2(\theta)\,\mathrm{d}\varphi^2\Big)\right], \tag{2.15}$$

where χ is the radial comoving distance and θ and φ are the polar angles. The parameter $K = (\Omega_{\mathrm{r}} + \Omega_{\mathrm{m}} + \Omega_{\Lambda} - 1)H_0^2/c^2$, the *curvature*, quantifies the deviation from a spatial Euclidean geometry ($K = 0$, flat curvature), and defines the *comoving angular diameter distance*

$$f_K(\chi) = \begin{cases} K^{-1/2}\sin\left(\sqrt{K}\,\chi\right) & \text{for} \quad K > 0, \\ \chi & \text{for} \quad K = 0, \\ (-K)^{-1/2}\sinh\left(\sqrt{-K}\,\chi\right) & \text{for} \quad K < 0. \end{cases} \tag{2.16}$$

If we insert the foregoing metric, and the energy-momentum tensor (Eq. 2.8), into Einstein's field equations (Eq. 2.2), we can solve for the relation between the expansion and the energy components. The result is the first and second Friedmann-Lemaître equations, concerning functions of the first and second derivatives of the scale factor $a(t)$ with respect of time. They read

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho(a) - \frac{Kc^2}{a^2} + \frac{\Lambda}{3}, \tag{2.17}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho(a) + \frac{3p(a)}{c^2}\right) + \frac{\Lambda}{3}. \tag{2.18}$$

With these relations, we link the observed evolution in the Universe with its composition. They are the dynamical equations that govern the evolution of the Universe. This equations are present in almost any cosmological problem.

2.4 Time & distances in cosmology

In cosmology one measures quantities like redshift, or angles on the sky. If we want to convert these into physical distances, or compute the age of the Universe when the radiation we see was emitted, we need to define our cosmological model (as we already did) and the values of its parameters.

The time between two epochs in the Universe can be obtained from the foregoing equations. Integrating Eq. (2.17) one can reach:

$$t(a_{\rm i}, a_{\rm f}) = \int_{a_{\rm i}}^{a_{\rm f}} {\rm d}a' \left(\frac{8\pi G}{3}\rho(a') - \frac{Kc^2}{a'^2} + \frac{\Lambda}{3} \right)^{-1/2} a'^{-1}, \tag{2.19}$$

where $a_{\rm i}$ and $a_{\rm f}$ are the scale factors at the initial and final moments respectively. The relation between scale factor and time, or redshift and time, can be obtained thereafter; we only need to invert last relation to get $a(t)$.

Due to the inclusion of a non-Euclidean metric, distances in cosmology are no longer uniquely defined. Moreover in GR they are also model dependent. For these reasons, we shall redefine the notion of distances. There are two main ways to define distances in cosmology, the *angular diameter distance* and the *luminosity distance*.

The angular diameter distance of an object is defined as the ratio between the subtended observed angle ${\rm d}\theta$ and its true physical transversal size ${\rm d}r$:

$$D_<(z) \equiv \frac{{\rm d}r}{{\rm d}\theta} = a(z)\, f_{\rm K}(\chi), \tag{2.20}$$

where χ denotes the radial comoving separation between the source at redshift z and the observer. We have denoted angle and transversal size as differentials, since in cosmology, the transversal scales we are treating are always very small compared to the distances. This allows us to work under the small angle approximation, i.e. $\sin(\theta) \simeq \tan(\theta) \simeq \theta$. The definition of the angular diameter distance is especially important in gravitational lensing, and we shall use it frequently.

The luminosity distance is defined in terms of the measured apparent magnitude and the absolute magnitude an object must have. If we consider the relation between flux S, the area $4\pi D_{\rm L}^2$ illuminated, and intrinsic luminosity L: $4\pi D_{\rm L}^2 S = L$; the luminosity distance yields

$$D_{\rm L}(z) \equiv \sqrt{\frac{L}{4\pi S}} = \frac{f_{\rm K}(\chi)}{a(z)}. \tag{2.21}$$

The last equality is obtained from the relation between luminosity and angular distance and is a result from the reciprocity theorem (Etherington 1933).

2.5 Brief history of the Universe

Over many years, the cosmological concepts and models have been extensively tested against observations of all kinds. In this section we want to give an overview of the current *concordance model*. The following picture is so far the one that conceals the agreement of a majority of researchers, and is consistent with past and present cosmological data.

The Universe is believed to have a finite age. Its starting point was a state of extremely high density and temperature, from which it began to expand and cool down. This paradigm is known as the Big Bang and it is supported by the following arguments. The current expansion implies that the Universe was denser in the past. Our knowledge about the energy content of the Universe (coming for example from dynamics in galaxy clusters), in combination with curvature constraints, leads to conclude from the Friedmann equations, that the Universe was once in a state of virtual infinite density at a moment around 13 Gyr ago.

Since the Universe has a finite age, and according to GR information cannot travel faster than the speed of light, parts of the Universe should have never been in causal contact. Light emitted shortly after the Big Bang reaches the earth today from distant parts of the Universe. Some parts are so apart that they cannot be in causal contact. However, these regions are still statistically identical, which leads to conclude that shortly after the Big Bang, the Universe expanded at a very high rate. The expansion was so fast that regions in tight causal contact back then, are today no longer connected. This period known as *Inflation* (see Linde 2005 for a review), also naturally explains the measured almost zero curvature. No matter what the curvature of the Universe was, due to the enlargement, we are going to see just a local region, and so it will appear to be flat. Inflation also gives an origin for the initial perturbations of the matter field which initiated the current mass structures in the Universe. Initial quantum fluctuations, during inflation were enlarged to macroscopic scales, seeding the structure growth that eventually formed galaxies, groups of galaxies and galaxy clusters.

After Inflation ended, the Universe was in a dense and hot state; since then it has been cooling down. As the temperature dropped, eventually atomic nuclei bound from free protons and neutrons, defining the primordial chemical composition of the Universe in a process we call Big Bang Nucleosynthesis. Based on our confirmed nuclear physics models and spectroscopic observations of the intergalactic medium, we are confident that the initial chemical composition of baryonic matter was 25% Helium and 75% Hydrogen, with traces of Lithium and other light elements (Tytler et al. 2000).

During Big Bang Nucleosynthesis, radiation was the dominant component of the energy density, but slowly matter took over. At a certain point nuclei and electrons formed atoms, while radiation was set free. This radiation has traveled ever since, from their original locations, suffering almost only the effects of the expansion of the Universe. As the temperature decreased, atomic Hydrogen was favored against free electrons and protons. The formation of neutral Hydrogen happened in a short period of time. The so-called Cosmic Microwave Background (CMB) radiation carries information from that moment. The most recent analysis of this radiation, has been done with the satellite mission WMAP (Bennett et al. 2003). This experiment offers us one of the most valuable ways to constrain the cosmological parameters. The sole existence of CMB is a proof of the expansion of the Universe and the Big Bang. From the typical scale of CMB anisotropies the ratio baryonic-dark matter can be computed. In combination with independent measures of

H_0 (Freedman et al. 2001) it is the best indicator of the curvature of the Universe.

The radiation-matter decoupling phenomenon is also imprinted on the formation of matter structure. One can detect today into the galaxy spatial distribution, a feature which is the result of the size of the sound horizon at this moment. This phenomenon is called Baryonic Acoustic Oscillations (BAOs).

Then a period known as the Dark Age began, where the matter overdensities grew, while first stars were formed. Their light most likely reionized the intergalactic medium as we see it today. Matter overdensities have been growing ever since, driving galaxy formation. First small objects like galaxies formed, then galaxy groups and clusters. The comparison of the structure we detect to our structure evolution simulations shows that dark matter is cold, that is, mass is far from a relativistic regime (White et al. 1983, White et al. 1984). We also know that objects of stellar origin do not account for it (Renault et al. 1997). The most popular theory is that dark matter is formed by weakly interacting massive subatomic particles.

Nowadays, we believe that since recent cosmological times expansion is dominated by the vacuum energy. The existence of vacuum energy was revealed studying Type Ia supernovae which have an almost constant luminosity regardless of their environment. Their luminosity distance as a function of redshift (Riess et al. 1998, Schmidt et al. 1998) reveals an expansion history which support the existence of vacuum energy.

2.6 Random fields

The Universe is homogeneous when averaged over large scales but obviously there are small-scale inhomogeneities. Its evolution and origin are treated in depth in the next chapter. However, in order to be able to characterize them properly we have to talk first about random fields.

The nature of many cosmological results is statistical. We assume that the observables that we can measure have an inherent randomness. In order to compare our models with reality we shall study the underlying probability distribution functions.

A *random field* is the set of possible realizations of a field plus the likelihood of each one. In mathematical terms, it is a map from the set of functions $\rho : \mathbb{R}^n \mapsto \mathbb{R}$ to $\mathbb{R}_0^+$. In cosmology, we often have a space identified with the $\mathbb{R}^3$ manifold, and we can assign to every point $\mathbf{x}$ a value by the mass density function $\rho(\mathbf{x})$. A given random field defines how probable each possible $\rho(\mathbf{x})$ function is. However, for simplicity we shall use the *density contrast* $\delta(\mathbf{x}) = (\rho(\mathbf{x}) - \bar{\rho})/\bar{\rho}$, where $\bar{\rho}$ is the spatial average of $\rho(\mathbf{x})$.

In reality, we only observe one Universe, and we must infer the underlying random field just from one realization. To overcome this problem we invoke the ergodic principle. This means that we assume that the expectation values over an ensemble of different small regions in the Universe is the same as over an ensemble of different possible universes at a small region.

In order to proceed further we must analyze the properties of a *Gaussian* random field. If we use the density contrast as example, a Gaussian field is characterized by the following conditions:

- all values of the function in Fourier space $\tilde{\delta}(\mathbf{k})$ are statistically mutually independent;
- the probability density for $\tilde{\delta}(\mathbf{k})$ is described by a Gaussian distribution.

The importance of Gaussian random fields can be explained by the Central Limit Theorem. This theorem states that the sum of independent random variables is asymptotically described by a Gaussian distribution as the number of variables increases. If we accept that the phenomena in the Universe come from a combination of independent random processes, the Gaussian field is the best starting point to describe a random field. A Gaussian random field is fully determined by the variance of each Fourier mode $\sigma^2\big(\tilde{\delta}(\mathbf{k})\big)$. The probability of a set of $\{\tilde{\delta}(\mathbf{k})\}$ is then

$$p\left(\left\{\tilde{\delta}(\mathbf{k})\right\}\right) = \prod_{\mathbf{k}} \frac{1}{\sqrt{2\pi\sigma^2(|\mathbf{k}|)}} \exp\left(-\frac{\tilde{\delta}^2(\mathbf{k})}{2\sigma^2(|\mathbf{k}|)}\right), \tag{2.22}$$

where the modulus of $\mathbf{k}$ appears since we assume an isotropic Universe. The variance of each mode is the so-called *power spectrum* $P(|\mathbf{k}|) \propto \sigma^2(|\mathbf{k}|)$, which can be easily estimated from the ensemble average of the product of two modes. In the most general case, where we have an infinite and continuous set of $\{\tilde{\delta}(\mathbf{k})\}$, we have

$$\left\langle \tilde{\delta}(\mathbf{k})\, \tilde{\delta}^*(\mathbf{k}') \right\rangle = (2\pi)^3 \delta_{\mathrm{D}}(\mathbf{k}-\mathbf{k}') P(|\mathbf{k}|), \tag{2.23}$$

where $\delta_{\mathrm{D}}(\mathbf{k}-\mathbf{k}')$ is Dirac's delta. Dirac's delta appears because we assume statistical homogeneity in the random field. In configuration space we can specify the random fields using the Fourier transform of the power spectrum, the *correlation function*. This can also be estimated from the two point correlator:

$$\langle \delta(\mathbf{x}) \delta^*(\mathbf{x}') \rangle = C(|\mathbf{x}-\mathbf{x}'|). \tag{2.24}$$

Due to the assumed homogeneity and isotropy, the correlator depends only on $|\mathbf{x}-\mathbf{x}'|$.

In the case of non-Gaussian random fields we shall need high order correlators for a full description. Nonetheless, the power spectrum and the correlation function are always the first step concerning random fields studies.

2.7 List of cosmological parameters

Finally we present a table with values for the most important cosmological parameters. The source used is the Legacy Archive for Microwave Background Data Analysis (LAMBDA, *http://lambda.gsfc.nasa.gov*), from NASA. They are the parameters for the concordance model ΛCDM, that is, a model with cosmological constant Λ and cold dark matter. The combined data are CMB data from WMAP7 (Komatsu et al. 2010), BAOs (Percival et al. 2010), and H_0 measurements (Riess et al. 2009).

From the list of parameters, the only one which has not been defined is σ_8, with which we parameterize the power spectrum normalization. The σ_8 parameter is the dispersion of the matter density contrast averaged over spheres of radius $8\,\mathrm{Mpc}/h$ at $z = 0$, and its definition has historical reasons. The power spectrum normalization can be written in terms of σ_8 which is easy to measure.

Cosmological Parameters for the concordance model.

Hubble constant H_0 ($\mathrm{km\,s^{-1}\,Mpc^{-1}}$)	$70.4^{+1.3}_{-1.4}$
Baryonic matter Ω_b	0.0449 ± 0.0028
Dark Matter Ω_m	0.222 ± 0.026
Dark Energy Ω_Λ	0.734 ± 0.029
Power Spectrum normalization parameter σ_8	0.809 ± 0.024
Mean redshift of the CMB	1088.2 ± 1.1
Age of the Universe (Gyr)	13.75 ± 0.11

3. Structure formation: theory and computational techniques

We assume that the Universe is homogeneous, at least on large scales. However inhomogeneities exist. The way the inhomogeneities appear has been discussed previously. Initial quantum fluctuations seeded matter inhomogeneities, and these grew with time due to self-gravitation.

This chapter lists the most important concepts concerning structure growth, and how they relate to a model of matter structure formation. For simplicity, most of the times the equations are presented without a detailed derivation. A more thorough discussion can be found in Dodelson (2003).

After reviewing the analytical treatment, we give an overview of the computational methods used to produce the cosmological simulations from which we derive our results. The technical details of the cosmological simulations will also be necessary in order to interpret properly our results in the following chapters.

3.1 Boltzmann equations

All assumptions and principles that we stated in the previous chapter, make cosmology a science where predictions and conclusions can only be expressed in terms of statistical results. We cannot (and should not aim to) explain the precise characteristics of a certain part of the Universe, only how likely it is. The best way to proceed therefore is to work with distribution functions. For each particle type (photons, baryons, etc.) in the Universe, we consider its distribution function f and how it changes with time t. The distribution function $f(\mathbf{x}, \mathbf{p}, t)$, describes for a given particle type, the occupation number for each phase-space volume element defined by the position $\mathbf{x}$ and the momentum $\mathbf{p}$. The total number of particles must not be constant, some species can transform into another, for instance matter into radiation. This is taken into account by a collision term $C(f)$. We encode all these considerations in the Boltzmann equation

$$\frac{\mathrm{d}f}{\mathrm{d}t} := \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}\frac{\mathrm{d}x^i}{\mathrm{d}t} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial f}{\partial \hat{p}^i}\frac{\mathrm{d}\hat{p}^i}{\mathrm{d}t} = C(f). \tag{3.1}$$

For convenience, the dependence in $\mathbf{p}$ has been divided into the dependence on the modulus $p = |\mathbf{p}|$, and the components of the unit vector $\hat{p}^i = p^i/|\mathbf{p}|$. The Boltzmann equation describes the evolution of the particle's distribution in the phase-space. When $C(f) = 0$, the phase-space volume is conserved.

In order to proceed further, we need to introduce the metric that defines space-time and so characterizes dynamics. In the following, for simplicity, we use comoving coordinates and also natural units i.e. we set the speed c of light and Planck's constant $\hbar$ to

unity. We consider a homogeneous expanding universe with some small perturbations. The perturbations to the Friedmann-Robertson metric can be decomposed into scalar, vectorial, or tensor perturbations (Bertschinger 2001). The different possibilities do not couple directly, whereas scalar perturbations are connected to the density field. Henceforth we only consider scalar perturbations to study structure growth. For the cases we treat in the following, we can parameterize scalar perturbation in the metric with the potential $\Phi(\mathbf{x}, t)$, which is considered to be small at any time. As stated in the previous chapter, we assume a flat space geometry for the Universe. In our particular choice of coordinates the metric tensor is then approximated by

$$\begin{aligned} g_{00} &= -1 - 2\,\Phi(\mathbf{x}, t), \\ g_{ii} &= a^2\,[1 - 2\,\Phi(\mathbf{x}, t)]\,, \end{aligned} \tag{3.2}$$

with all non-diagonal elements being zero. With the metric, we can rewrite $\mathrm{d}p^i/\mathrm{d}t$, $\mathrm{d}x^i/\mathrm{d}t$ and $\partial/\partial p$, as functions of the metric perturbations and the energy E. After some algebraic manipulations, inserting the metric into Eq. (3.1) yields:

$$\frac{\partial f}{\partial t} + \frac{\hat{p}^i\,p}{a\,E}\frac{\partial f}{\partial x^i} - \frac{\partial f}{\partial E}\left(\frac{p^2}{E}\frac{\dot{a}}{a} + \frac{p^2}{E}\frac{\partial \Phi}{\partial t} + \frac{\hat{p}^i p}{a}\frac{\partial \Phi}{\partial x^i}\right) + \frac{\partial f}{\partial \hat{p}^i}\frac{\mathrm{d}\hat{p}^i}{\mathrm{d}t} = C(f). \tag{3.3}$$

In order to understand better the physics behind the Boltzmann equation, we must solve the problem piecewise. We subdivide the problem according to the scale of the perturbation we want to study, and the epoch we want to describe. We define this subdivision by considering the expansion of the Universe as follows.

Information can only travel at a finite velocity, which means that any point in the universe is in causal contact only with a finite region. The *causal horizon* defines how far information can travel since the Big Bang, and therefore how big this region is. The horizon grows with time, but at any period it separates between scales where there is causal contact from those where there is not.

Depending on the point in time we want to treat, the expansion is different; at early times it was driven by radiation, later by matter, and nowadays by Dark Energy. The physical conditions such as pressure or temperature that affect perturbations inside the horizon, are different as well. The transition periods, especially between radiation and matter dominance, are used as milestones.

3.2 Linear Theory

The Boltzmann equation as presented here is not very usable due to its complexity. We may not find a solution $f(\mathbf{x}, \mathbf{p}, t)$ for the full equation but for simpler derived quantities. The aim of this section is to illustrate how analytical solutions can be found at least under certain conditions. For early times and for large scales at late times, in the so-called linear regime, the Boltzmann equation is linear in $f(\mathbf{x}, \mathbf{p}, t)$. Therefore it can be worked out, producing solvable expressions.

We shall focus on describing the evolution of matter perturbations when the Universe is matter dominated, on scales much smaller than the horizon. For this case, we can consider dark matter alone, which is collisionless ($C(f) = 0$). The solution of the equation for other times, scales, or even species is in many ways similar.

The standard way to proceed is to derive equations for the spatial density field and the velocity field from Eq. (3.3). The first useful equation arises as we multiply it by $\mathrm{d}^3p/(2\pi)^3$ and we integrate over momentum. The factor $(2\pi)^3$ ($(2\pi\hbar)^3$ in reality with $\hbar = 1$), is the phase-space volume element. This leads to define the dark matter density as

$$\rho_{\rm dm}(\mathbf{x}, t) = \int \frac{\mathrm{d}^3p}{(2\pi)^3}\, f_{\rm dm}(\mathbf{x}, \mathbf{p}, t), \tag{3.4}$$

and the first-order correction to homogeneity, the density contrast,

$$\delta_{\rm dm}(\mathbf{x}, t) = \frac{\rho_{\rm dm}(\mathbf{x}, t) - \bar{\rho}_{\rm dm}(t)}{\bar{\rho}_{\rm dm}(t)} \quad \text{such that} \quad \rho_{\rm dm}(\mathbf{x}, t) = \bar{\rho}_{\rm dm}(t)\big(1 + \delta_{\rm dm}(\mathbf{x}, t)\big), \tag{3.5}$$

with $\bar{\rho}(t)$ as the spatial average of $\rho(\mathbf{x}, t)$. We also need to define the velocity of each volume element. The velocity reads, component-wise like

$$v^i_{\rm dm}(\mathbf{x}, t) := \frac{1}{\rho_{\rm dm}(\mathbf{x}, t)} \int \frac{\mathrm{d}^3p}{(2\pi)^3}\, f_{\rm dm}(\mathbf{x}, \mathbf{p}, t)\, \frac{p\,\hat{p}^i}{E}. \tag{3.6}$$

The remaining task is to distinguish between dominant and higher order correction terms. We have assumed at zeroth-order an isotropic and homogeneous distribution, henceforth dependencies on the momentum direction in $f_{\rm dm}(\mathbf{x}, \mathbf{p}, t)$ are first-order corrections. The temporal variation of the momentum direction is a first-order correction in momentum. Therefore, the term $\frac{\partial f}{\partial \hat{p}^i}\frac{\mathrm{d}\hat{p}^i}{\mathrm{d}t}$ in Eq. (3.3), is a product of two first-order corrections and we can neglect it. From isotropy, as we integrate out momentum, the term $\hat{p}^i p \frac{\partial f}{\partial E}$, should give a null contribution at zeroth-order in $f_{\rm dm}$. The spatial derivative of Φ is first-order in the metric, so the term with these two last quantities drops out. The surviving terms at zeroth and first-order are then

$$\frac{\partial \rho_{\rm dm}}{\partial t} + \frac{1}{a}\frac{\partial(\rho_{\rm dm} v^i_{\rm dm})}{\partial x^i} + 3\left(\frac{\dot{a}}{a} + \frac{\partial \Phi}{\partial t}\right)\rho_{\rm dm} = 0. \tag{3.7}$$

Now, we can collect the zeroth-order terms into the expression

$$\frac{\partial \bar{\rho}_{\rm dm}}{\partial t} + 3\,\frac{\dot{a}}{a}\,\bar{\rho}_{\rm dm} = \frac{\mathrm{d}\,(\bar{\rho}_{\rm dm}\, a^3)}{\mathrm{d}t} = 0, \quad \text{with the solution} \;\; \bar{\rho}_{\rm dm} \propto a^{-3}; \tag{3.8}$$

and the first-order corrections into the equation:

$$\frac{\partial \delta_{\rm dm}}{\partial t} + \frac{1}{a}\frac{\partial v^i_{\rm dm}}{\partial x^i} + 3\frac{\partial \Phi}{\partial t} = 0. \tag{3.9}$$

The second equation necessary to solve for both $\delta_{\rm dm}$ and $v^i_{\rm dm}$ is obtained by computing the expectation value of $(\hat{p}^i p/E)$. To do so, we multiply the Boltzmann equation by $\mathrm{d}^3p\,\hat{p}^i p/(2\pi)^3/E$ and we integrate over momentum again. Now we must carefully distinguish between first-order and second-order correction terms. The critical point is to remember that we are dealing with non-relativistic particles and therefore $p/E \ll 1$. Leaving aside the details, the resulting expression for the first-order terms is

$$\frac{\partial \rho_{\rm dm} v^j_{\rm dm}}{\partial t} + 4\,\frac{\dot{a}}{a}\rho_{\rm dm} v^j_{\rm dm} - \frac{\rho_{\rm dm}}{a}\frac{\partial \Phi}{\partial x^i} = 0.$$

If we apply the zeroth-order approximation $\rho_{\rm dm} \simeq \bar{\rho}_{\rm dm}$ and the solution from Eq. (3.8) we obtain:

$$\frac{\partial v^j_{\rm dm}}{\partial t} + \frac{\dot{a}}{a} v^j_{\rm dm} - \frac{1}{a}\frac{\partial \Phi}{\partial x^i} = 0. \tag{3.10}$$

The last step is to Fourier transform the equations. In this way we get rid of spatial partial derivatives as they transform like $\partial/\partial x^i \mapsto \mathrm{i}\, k^i$. Now we have functions of the wave vector $\mathbf{k}$ instead of position $\mathbf{x}$. The velocity field is assumed to be irrotational, which is equivalent to say that its Fourier transform $\tilde{\mathbf{v}}$ and $\mathbf{k}$ are parallel. This leads to a further simplification; we can replace in our equations $\tilde{\mathbf{v}}$ and $\mathbf{k}$ by their moduli k and v.

$$\dot{\tilde{\delta}}_{\rm dm} + \frac{\mathrm{i}\, k}{a}\, \tilde{v}_{\rm dm} + 3\, \dot{\tilde{\Phi}} = 0, \tag{3.11}$$

$$\dot{\tilde{v}}_{\rm dm} + \frac{\dot{a}}{a}\tilde{v}_{\rm dm} - \frac{\mathrm{i}\, k}{a}\, \tilde{\Phi} = 0. \tag{3.12}$$

The set of equations that describes our problem is completed with the relation between the potential perturbations and the energy perturbations, which is the relativistic analogy of the Poisson equation in classical dynamics:

$$k^2\, \tilde{\Phi} = \frac{3\, H_0^2\, \Omega_{\rm m}}{2\, a} \tilde{\delta}_{\rm dm}. \tag{3.13}$$

Note that our choice of the metric convention is responsible for the lack of a minus sign, which appears in other derivations.

3.3 Transfer function & growth function

For the work in this thesis we only need predictions for structures smaller than the horizon well after the matter domination started. To study how these structures evolve we have derived in the previous section the necessary equations, and now we shall give a solution for them.

From equations (3.11) and (3.12), changing the derivatives in t to derivatives in a for convenience, we obtain

$$\frac{\mathrm{d}\tilde{\delta}}{\mathrm{d}a} + \frac{\mathrm{i}\, k\, \tilde{v}}{a^2\, H(a)} = -3\frac{\mathrm{d}\tilde{\Phi}}{\mathrm{d}a}, \tag{3.14}$$

$$\frac{\mathrm{d}\tilde{v}}{\mathrm{d}a} + \frac{\tilde{v}}{a} = \frac{\mathrm{i}\, k\, \tilde{\Phi}}{a^2\, H(a)}. \tag{3.15}$$

Since we only treat sub-horizon scales, which implies that $k/[a\, H(a)] \gg 1$, we can neglect the second term in the l.h.s. of Eq. (3.14), and do the same with the r.h.s. in Eq. (3.15). With a little bit of algebra we can combine these last equations with Eq. (3.13) and get:

$$\frac{\mathrm{d}^2\tilde{\delta}}{\mathrm{d}a^2} + \left(\frac{\mathrm{d}\ln(H)}{\mathrm{d}a} + \frac{3}{a}\right)\frac{\mathrm{d}\tilde{\delta}}{\mathrm{d}a} - \frac{3\,\Omega_{\rm m}\, H_0^2}{2\, a^5\, H^2(a)}\tilde{\delta} = 0. \tag{3.16}$$

A first analysis of the equation shows us that there is no dependence on k, so the solution can be factorized like

$$\tilde{\delta}(k, a) = D_+(a)T(k) + D_-(a)T(k), \tag{3.17}$$

where we wrote a growing D_+T and a decaying D_-T partial solution. It is easy to prove that in our case, with a constant Dark Energy density, the Hubble rate is one of the two partial solutions of the equation. As the Hubble rate decreases in the regime that we describe: $D_-(a) = H(a)$. Given one partial solution, the other can be obtained using the Wronskian theorem. In this way we obtain the *growth function*, which yields:

$$D_+(a) \propto H(a)\, H_0^2 \int_0^a \frac{\mathrm{d}a'}{\left[a'\, H(a')\right]^3}. \tag{3.18}$$

We normalize the growth function $D_+(a)$ such that it is unity today $D_+(a_0) = 1$, the proportionality constant is then $5\,\Omega_\mathrm{m}/2$. This solution is valid for sub-horizon scales after matter dominance started.

The whole evolution undergone at earlier times can be condensed into the so-called *transfer function* $T(k)$. With the transfer function we describe the distinct evolution that perturbations of different wavenumber k suffered before. Perturbations of different scales entered the causal horizon at different moments, and this affected their evolution differently.

The transfer function parameterizes the growth of a perturbation of a given mode k with respect to a fiducial one. The fiducial perturbation has a small wavenumber k_fid (large scale), small enough that it entered the horizon today. Let us denote a perturbation today by $\delta_0(k) = \delta(k, a_0)$; and a perturbation at a moment in the past where the growth function description starts to be valid like $\delta_i(k) = \delta(k, a_i)$. Then the definition of the transfer function is:

$$T(k) := \frac{\delta_0(k)/\delta_0(k_\mathrm{fid})}{\delta_i(k)/\delta_i(k_\mathrm{fid})}. \tag{3.19}$$

Small perturbations that entered the horizon at radiation dominance era were suppressed. Large perturbations that entered the horizon recently on the other hand, had a different evolution. The transfer function reflects this relative change.

Nowadays the optimal way to deal with the Boltzmann equation is to compute a a solution with numerical methods. The transfer function is normally described by a fitting formula for the resulting output (e.g. Eisenstein & Hu 1998).

As mentioned in Sect. 2.6, we normally characterize the inhomogeneities distribution using their power spectrum. The foregoing analysis of the Boltzmann equation allows us compute in the linear regime, how the power spectrum evolves with time:

$$P(k, a) = A\, k^{\mathrm{n_s}}\, T_k^2\, D_+^2(a). \tag{3.20}$$

The initial power spectrum is assumed to be the power law $A\, k^{n_s}$, so it does not single out any scale. Moreover, for $n_\mathrm{s} = 1$ (known as the *Harrison-Zel'dovich spectrum*), the power spectrum is scale invariant. This means that no moment in time is singled out neither. This power spectrum does not carry any information of its causes. It was postulated as a sensible ansatz, and measurements today indicate that it is a good approximation.

3.4 Non-linear evolution

To tackle the problem beyond the linear regime, one can use perturbation theory or phenomenological models. However, the most successful approach is to rely on numerical methods.

The full solution of the Boltzmann equation however cannot be obtained by a direct integration. The high dimensionality of the problem makes a brute force approach impossible. The most successful way to overcome this problem are dark matter simulations. Since we cannot find the full solution for the distribution $f_{\rm dm}(\mathbf{x}, \mathbf{p}, t)$, we use the Boltzmann equation to sample it. The statistics over this sample can be used to reconstruct $f_{\rm dm}(\mathbf{x}, \mathbf{p}, t)$.

In a typical cosmological numerical simulation, we try to simulate a large region of the universe. In the simulations that we use, for example, the matter distribution is represented with discrete collisionless mass particles, which mimic the behavior of dark matter. We compute how they interact due to gravitation with a background expansion, and we store the position and velocity of the particles at a set of instants for future analysis. An important feature of this method is that the sampling changes from region to region. Underdense regions are sparsely sampled and therefore much worse represented than overdense regions. How the particles cluster is compared to our cosmological observations.

The inclusion of full hydrodynamics in cosmological simulations is still a computational challenge today. However, baryons have an impact mainly on small scales as they form galaxies.

Since we still need galaxies for our study, the simulation used in this thesis is populated with galaxies using semi-analytical galaxy models. These galaxies are obtained from an already finished dark matter only cosmological simulation. We describe the method in Sect. 3.7.

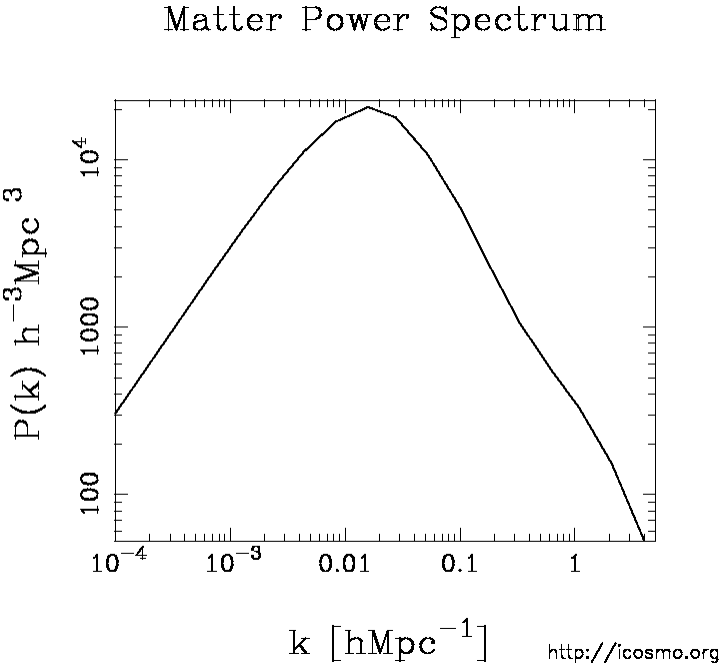

Figure 3.1: Matter power spectrum computed in the linear regime with Eisenstein & Hu (1998) and in the non-linear regime following Smith et al. (2003). The plot was obtained from the web page http://www.icosmo.org.

The combination of the analysis of the linear regime and the simulations on the non-linear evolution allows us to compute the matter power spectrum for any cosmology. A formula can be derived thereafter (Smith et al. 2003). We plot in Fig. 3.1 the power spectrum for a ΛCDM cosmology with the parameters described in Sect. 2.7.

3.5 Millennium Run

This work is based on the Millennium Simulation (Springel et al. 2005). The Millennium Simulation currently ranks among the largest cosmological simulations ever made. The simulation represents a periodic box of $500\,h^{-1}$ comoving Mpc on each side, containing 2160^3 collisionless particles with a mass of 8.6×10^8 $\mathrm{M}_\odot\,h^{-1}$. The output consist of 64 snapshots logarithmically spaced in redshift. Both velocity and position of each particle are stored. The cosmological parameters that define the simulation are: $\Omega_\mathrm{m} = 0.25$, $\Omega_\Lambda = 0.75$, $h = 0.73$, consistent with the ΛCDM model; and $\sigma_8 = 0.9$ which is slightly higher than the currently accepted value. The value used for σ_8 shall increase the number density of clusters in comparison to reality. However we estimate that it can only have a mild effect on our final measurements. The simulation was started at a redshift $z = 127$. The initial conditions were created by perturbing a homogeneous quasi-random particle distribution, with a realization of a Gaussian random field. The Gaussian field had the ΛCDM linear power spectrum, computed with the code CMBFAST (Seljak & Zaldarriaga 1996). Finally, periodic boundaries are used which produces homogeneity at large scales.

To illustrate some aspects of the cosmological simulation, we shall describe the main points of GADGET2 (Springel 2005). This code was used to conduct the Millennium Run. We remark that velocities and potentials in cosmological simulations are always small enough so Newtonian approximations can be made. That is, we work with instant interactions.

In order to avoid point masses, which do not reflect reality and lead to unphysical strong scattering events between two particles, the mass of a particle is distributed over a sphere. The size of the sphere is parameterized with the so-called softening length ϵ. In our case the sphere has a radius of $r \lesssim 2.8\epsilon$, ($\epsilon = 5\,h^{-1}$ kiloparsecs) with a specific function $\delta_\epsilon(\mathbf{x})$. The simulation has therefore a spatial resolution limited to scales of $2.8\epsilon = 14\,h^{-1}$ kiloparsecs.

The gravitational potential produced by a particle of mass m at a distance $\mathbf{x}$ from its center follows the Poisson equation

$$\nabla^2\varphi(\mathbf{x}) = 4\pi G\,m\left(-\frac{1}{L^3} + \sum_{\mathbf{n}} \delta_\epsilon(\mathbf{x} - \mathbf{n}L)\right), \tag{3.21}$$

where the sum over $\mathbf{n}$, represents a sum over all periodic replicas of the box, present due to the periodic boundary conditions. The parameter L is the size of the box. The first term on the r.h.s. of the equation subtract the mean density, which is not a source for the potential. We obtain the final gravitational potential at a given location $\mathbf{r}$ by summing over the N simulation particles:

$$\nabla^2\phi(\mathbf{r}) = \sum_j^N \nabla^2\varphi(\mathbf{r} - \mathbf{r}_j), \tag{3.22}$$

where we denote the location of the jth particle by $\mathbf{r}_j$. Dark energy appears in the background expansion, and not directly in the gravitational potential.

The computation of the potential at any given moment, an essential part of the simulation, is done combining two methods. The potential is split into a long-range and a short-range contribution. For long distances, the particles are grided into a coarse mesh and Fast Fourier Transform methods are used to solve the Poisson equation. FFT is time efficient but the memory consumption is high, and the grid imposes a geometry that is not isotropic on small scales. For this reason, the short-range contribution is computed using a tree code. A tree code avoids computations with an adaptative spatial grid. The required precision defines the spatial resolution in each case. If at a given location there is a compact group of particles, for large enough separations from the group center, the tree code computes the potential contribution of the group's total mass located at its average position. In order to compute the interaction between the particles of the group, the tree code considers them individually. The grid is changed depending on the scales and the densities. Eventually long-range and short-range contributions are combined to achieve the gravitational force exerted at each particle.

The integration of the dynamical equations is done in a scheme referred to as *leapfrog*. This scheme has been proven to be more efficient that other integration rules. Under this scheme, the ith particle at its nth integration step, is assigned a time step $\Delta t_i^{(n)}$, a current time $t_i^{(n)}$, and with position, velocity and acceleration $\left(\mathbf{r}_i^{(n)}, \mathbf{v}_i^{(n)}, \mathbf{a}_i^{(n)}\right)$.

The states of the particles are computed following an order. In each moment the next particle to be advanced in time, the particle k, is that with the minimal $\tau_k^{(n)} = t_k^{(n)} + 0.5\,\Delta t_k^{(n)}$. The scheme computes the position of all particles at an intermediate position $\tilde{\mathbf{r}}_i$ from the current state n,

$$\tilde{\mathbf{r}}_i = \mathbf{r}_i^{(n)} + \mathbf{v}_i^{(n)}(\tau_k - t_i). \tag{3.23}$$

With the intermediate positions, the acceleration for the integration step $n+1$ of the k particle is computed

$$\mathbf{a}_k^{(n+1)} = -\nabla\phi(\tilde{\mathbf{r}})\Big|_{\tilde{\mathbf{r}}_k}, \tag{3.24}$$

and once the acceleration is computed, the velocity and the position are updated:

$$\mathbf{v}_k^{(n+1)} = \mathbf{v}_k^{(n)} + \mathbf{a}_k^{(n+1)}(\tau_k - t_k) \tag{3.25}$$

$$\mathbf{r}_k^{(n+1)} = \mathbf{r}_k^{(n)} + \left(\mathbf{v}_k^{(n+1)} + \mathbf{v}_k^{(n)}\right)(\tau_k - t_k). \tag{3.26}$$

Finally the time is also updated

$$t_k^{(n+1)} = 2\tau - t_k^{(n)}, \tag{3.27}$$

and the new time step is derived from the position of the particle and the density of the particle environment.

The time-line is common to all particles, but each one advances along it in steps of independent size. The particles in higher density regions have shorter time steps, and are updated more frequently.

In Fig. 3.2 one can see a slice of the simulation with a series of zooms of an overdensity. The color encodes the density. On the background picture one can see the large-scale filamentary structure, the so-called cosmic web. One can also notice that the simulation

is homogeneous on large scales. In the first zoom, it is possible to see how the filaments cross into a particular overdensity, and the presence of voids. The last zoom also shows how within the large mass clump, there are small local overdensities. This substructure or sub-halos is the object of our study.

3.6 Substructure in the Millennium Simulation

Previously we have presented a general way of characterizing inhomogeneities, through random fields. We can also study the statistics of large matter overdensities, referred to as halos, which are represented in the simulation by large groups of particles. Moreover, we find small inhomogeneities embedded into large ones, called sub-halos. The statistics of individual halos and sub-halos is a complementary analysis to that conducted with a random field study. In this section we describe how structure and substructure is defined in the Millennium Simulation.

Independent groups of particles are grouped with the *Friends-of-Friends* (FOF) algorithm as the simulation runs. All particles which are closer than 0.2 times the mean particle separation are linked together. These FOF groups are identified as main halos, and should represent fairly the dark matter component of galaxy clusters, groups or isolated galaxies. This method only studies the spatial relation. Some structure will be linked although they are not gravitationally bound or even coming apart. Nonetheless, this caveat is also present on the definition of membership in galaxy groups on real observations, where we only detect angular positions and redshift. For every FOF group a mass $M_{\rm vir}$ is estimated. The value given is the mass inside a sphere of radius $R_{\rm vir}$, centered on the most bound particle. The radius is computed such that the density inside the sphere is 200 times the mean density of the simulation.

Substructure is determined, once the simulation is finished, by the SUBFIND algorithm (Springel et al. 2001). Each FOF group is subjected to study and a truly gravitational analysis is performed. In principle a full hierarchy of sub-halos could be found, that is, sub-halos within sub-halos are possible. However only one degree of substructure is considered here.

The determination is done as follows. First a local density is estimated considering the $N \simeq 10$ nearest neighbors at each particle position of the group. Then the particles are ordered according to the density at their positions. Each time a particle has no neighbor with higher density, it is considered the most bound particle of a sub-halo candidate. The sub-halo candidates grow as more particles are assigned going from the most dense positions to the least dense. Whenever a particle can be place into two sub-halo candidates the candidates stop to grow and are joined to form a new sub-halo candidate. Then the candidates are filtered. Each sub-halo candidate is peeled of particles with positive binding energy until none is left. The position of the sub-halo is the position of its most bound particle. Every group of at least 20 gravitationally bound particles forms a sub-halo, and its characteristics are stored. The mass of the sub-halos is not defined as for the FOF halos. Instead, it is the total mass from the particles which have been assigned by SUBFIND. The comparison of radial extent between sub-halos and main halos of the same nominal mass is not well defined as different mass definitions are considered.

It must be noted that the technique used neglects the fact that the host halo helps to bind the structure. Also there are particles which surround the sub-halo and therefore are related to it but they are not detected to be gravitationally bound. With all this some

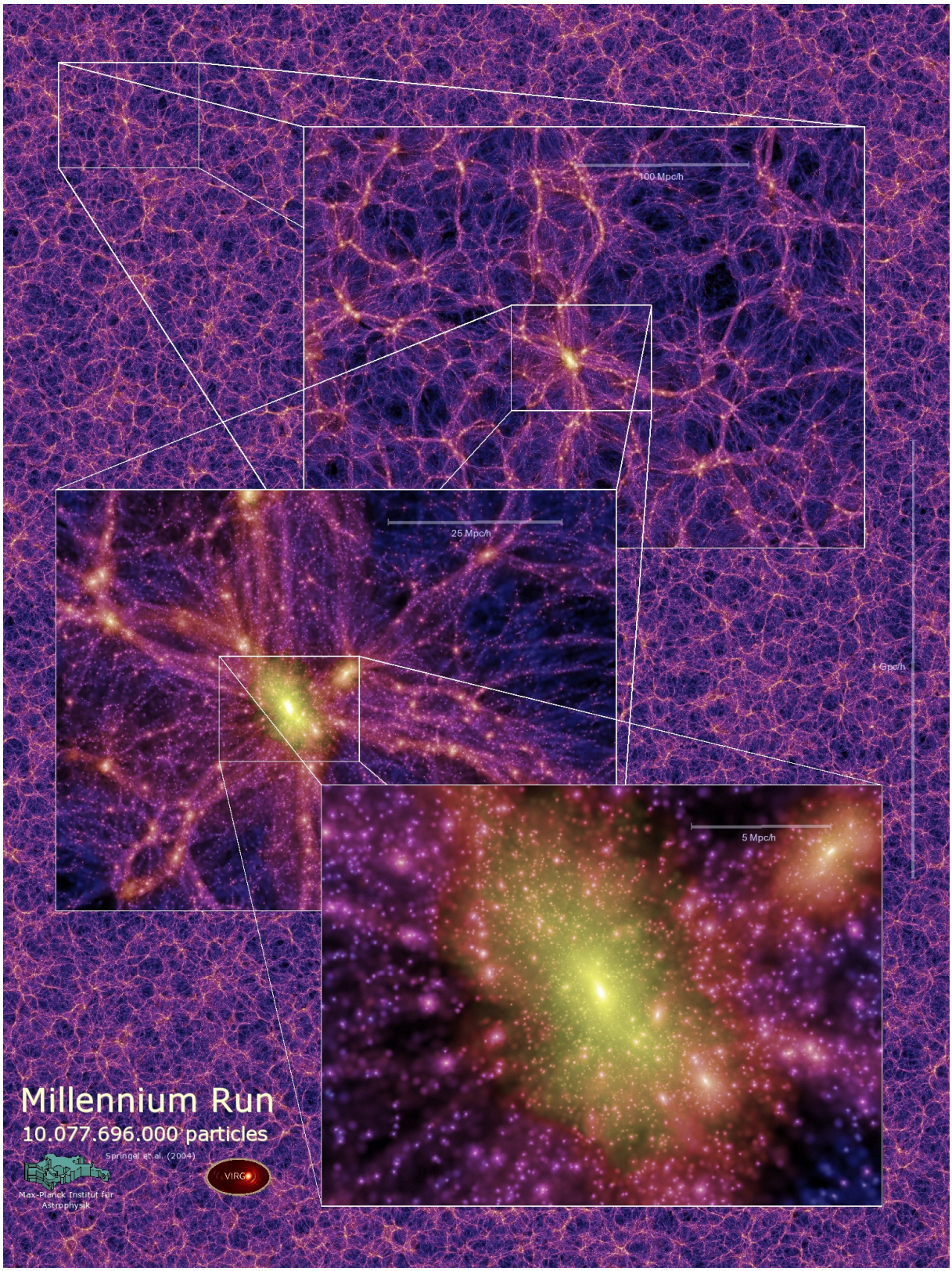

Figure 3.2: Slice of the Millennium Simulation with a series zooms of an overdensity. From the Millennium Simulation web page http://www.mpa-garching.mpg.de/galform/presse/

ambiguity remains, but this does not come from the algorithm; it is an intrinsic problem to the substructure definition. Other algorithms assign a different amount of particles, and the mass assigned is also a function of the sub-halo position inside the main halo (Muldrew et al. 2010). We consider that there are many valid ways to define the mass of a sub-halo. In order to be coherent, for the rest of the work we use the mass from SUBFIND $M_{\rm SUBF}$ as the sub-halo mass, which we obtain from the simulation catalogues.

3.7 Semi-analytic catalogues

Since the Millennium Simulation contains only dark matter, in order to include galaxies, semi-analytic models of galaxy formation are used. The semi-analytic models are crucial for this thesis. Our method requires to classify galaxies into homogeneous subsets, as we obtain high signal-to-noise ratios by averaging over a set of them. Within the simulation, we can classify the sub-halos, with quantities like mass or time spent inside the main halo. However, in real data one does not have this quantities and one must rely on observables of the hosted galaxy. The semi-analytical catalogues give us a set of galaxy properties which we can attach to the halos, and the possibility of comparing our results to real surveys. The correlation between the quantity we want to infer and the observable may not be strong enough. As a result from this, we may arrive at a wrong conclusion. With the help of the semi-analytic catalogues we can study this effect.

Once the dark matter simulation is conducted, these models assign a galaxy to each dark matter halo based on its history and characteristics. In this section we sketch the recipes used. However the technical details are outside the scope of this thesis. An overall description can be found in Croton et al. (2006), and most of the critical issues are developed in De Lucia et al. (2004) and Springel et al. (2001). In this work we use the semi-analytic catalogues by De Lucia & Blaizot (2007).

3.7.1 Basic concept

The starting point of the semi-analytical model is to consider, at high redshift, a fraction of the halo mass as baryonic mass. In the current model, it is 17% of the mass. This value is reduced for low mass halos, since dwarf galaxies contain a relatively small fraction of all condensed baryons according to observations (Kravtsov et al. 2004). Initially the baryons are in the form of a diffuse gas with primordial composition.

Galaxies form through the evolution of the gas. A disk and a bulge shall form under the appropriate conditions, as well as stars. The gas also changes its chemical composition, as the metalicity increases. These processes lead to form galaxies with a broad variety of properties.

Once the baryons are assigned to the halos at high redshift, the halos are followed through time to determine the evolution of the gas. We shall divide the causes that drive baryon evolution into two types: passive evolution defined by the past and present conditions of the halo, and evolution through merging of different halos.

The mergers are crucial in the galaxy evolution and the backbone of the semi-analytic models. In fact a *merger tree* is defined, storing for each halo at each redshift what merging events it suffered. The merging events are divided into *minor* mergers, where the mass ratio is smaller than 0.3, and into *major* mergers otherwise. The different components of the two galaxies (disk, bulge, stellar mass ...) hosted by the merging halos are added

into the resulting one, but in a different way in each case. When a small galaxy merges into a larger one, the stars of the satellite are added to the bulge of the central one. The masses of both cold gas disks are summed up, and a minor star burst is induced. When two galaxies of comparable size merge, a significant star burst occurs, both disks are destroyed, and a bulge is created out of them where all stars are placed.

The recipes used to obtain the baryon evolution have a series of free parameters. These free parameters are fixed in such a way that the final galaxy catalogues follow the observed properties of galaxy populations, such as luminosity scaling relations. In the following we offer a list of the most important recipes for the semi-analytical models.

3.7.2 Gas accretion and cooling

The gas collapses into the potential wells due to gravity, releasing thermal radiation. Under the appropriate conditions this gas condenses and a galaxy is formed. The rate at which gas falls into the halo center depends on the halo mass and the redshift.

In most of the cases the gas shocks to the virial temperature and forms a quasi-static hot atmosphere. The inner part of this atmosphere can collapse into a central object which is assumed to be a cold gas disk. In order to describe the mass cooling rate $\dot{m}_{\rm cool}$, a *cooling time* is defined for each halo radius. The cooling time is modeled in terms of the specific thermal energy of the hot gas and a cooling function which depends on the chemical composition and density. Then for a given halo, the radius at which the cooling time equals the dynamical time $\left(0.1\,H^{-1}(z)\right)$ defines the *cooling radius* $r_{\rm cool}$. The cooling radius encloses the gas that is accreted from the halo atmosphere. The mass cooling rate is then given by the equation

$$\dot{m}_{\rm cool} = 4\pi\rho_{\rm g}(r_{\rm cool})\, r_{\rm cool}^2\, \dot{r}_{\rm cool}, \tag{3.28}$$

where $\rho_{\rm g}(r_{\rm cool})$ is the hot gas density at the cooling radius. If the cooling radius is too small, the galaxy will not accrete much material. In low-mass halos the formal cooling radius is larger than the halo radius, and basically all infall material is accreted immediately onto the central object.

Some corrections are made to the initial model of gas accretion. Supermassive black holes, which we believe to exist at the center of each galaxy, can accrete gas. This gas accretion produces radiation that will suppress cooling. This is also present in the semi-analytic models. The accretion of hot gas by the central black hole $\dot{m}_{\rm BH}$ is described by a phenomenological model which depends on the black hole mass, the hot gas fraction, and the virial radius and mass of the dark matter halo. The luminosity produced by this accretion is then $L_{\rm BH} = 0.1\,\dot{m}_{\rm BH}\,c^2$. The factor 0.1 is the standard efficiency with which mass is assumed to produce energy near the event horizon. This luminosity produces a radiation pressure that expels mass at the rate

$$\dot{m}_{\rm exp} = \frac{2L_{\rm BH}}{V_{\rm vir}^2} \tag{3.29}$$

where $V_{\rm vir}$ is the virial velocity of the halo. The virial velocity can be expressed as a function of the virial mass and virial radius: $V_{\rm vir} = \sqrt{G\,M_{\rm vir}/R_{\rm vir}}$. The final modified cooling rate is the difference between the gas cooled and expelled by the black hole radiation:

$$\dot{m}_{\rm cool} = 4\pi\rho_{\rm g}(r_{\rm cool})\, r_{\rm cool}^2\, \dot{r}_{\rm cool} - \frac{2L_{\rm BH}R_{\rm vir}}{GM_{\rm vir}}. \tag{3.30}$$

The black hole mass grows also through mergers. In a merging event the two central black holes will merge, producing a larger one. At the same time the merging produces instabilities in the cold gas disk component, and part of this cold gas is accreted onto the new black hole.

3.7.3 Star formation & feedback

Star formation can happen through mergers that will induce star bursts, and in a quiescent way in the cold gas disk. In the case of mergers, the fraction of the combined cold gas from the two galaxies that is turned into stars is:

$$f_{*\mathrm{burst}} = 0.56 \cdot (m_{\mathrm{sat}}/m_{\mathrm{central}})^{0.7}. \tag{3.31}$$

We have denoted the mass of the heaviest galaxy as m_{central} and the lightest one by m_{sat}; the power law amplitude and index are empirically chosen to match observations.

The quiescent star formation depends on the parameters of the cold gas disk. The cold gas disk is described by its mass m_{cold}, and the halo M_{vir}, R_{vir} and spin parameter λ. The spin parameter is a dimensionless quantity derived from the angular momentum modulus J of the mass inside R_{vir}:

$$\lambda = \frac{J}{\sqrt{2} M_{\mathrm{vir}} V_{\mathrm{vir}} R_{\mathrm{vir}}}, \tag{3.32}$$

The model defines a critical mass m_{crit} for each halo like:

$$m_{\mathrm{crit}} = 3.8 \times 10^9 \frac{V_{\mathrm{vir}}}{200\,\mathrm{km\,s^{-1}}} \frac{3\lambda\, R_{\mathrm{vir}}}{\sqrt{2}\,10\,\mathrm{kpc}} \mathrm{M}_\odot. \tag{3.33}$$

For disks with mass below m_{crit} no stars will be produced. Also a dynamical time for the disk $t_{\mathrm{dyn,disk}}$ is defined. The final star formation rate coming from all these assumptions, is governed by the equation:

$$\dot{m}_* = \alpha(m_{\mathrm{cold}} - m_{\mathrm{crit}})/t_{\mathrm{dyn,disk}} = \alpha(m_{\mathrm{cold}} - m_{\mathrm{crit}}) \frac{V_{\mathrm{vir}}\sqrt{2}}{3\lambda\, R_{\mathrm{vir}}}. \tag{3.34}$$

The efficiency parameter α is set so that 5 to 15 percent of the gas is converted into star in a disk dynamical time.

The processes that lead to star formation implies also a chemical enrichment of the medium and an energy release when massive stars end in supernovae. The chemical enrichment in the gas is computed assuming a constant ratio between stellar mass formed and heavy elements returned. At the same time, a portion of gas is ejected outside of the halo, in order to consider the effect of supernovae. This ejected gas may be re-captured later on. Each one of these processes have a free parameter which is adjusted to produce realistic galaxy populations.

The way in which the star evolution affects the characteristics of the host galaxy is a field of research in development. The use of empirical approaches such as the one referred here, is therefore one of the most efficient ways to simulate galaxies.

4. Gravitational lensing

Light deflection by gravity or *gravitational lensing* is the cosmological probe we are going to use for our research. Its analysis allows us to measure the mass of objects at astronomical and cosmological distances.

Depending on the feature studied it is divided into the following sub-categories or regimes. With the aid of gravitational lensing we can detect planets and other compact objects of stellar mass, in what is called *microlensing*. If such a compact massive object crosses the line-of-sight to a background star, its mass can focus the light of the star, changing the observed brightness over time. The detection requires an extremely good alignment between source, lens and observer, and astronomical separations to maximize the effect. Nonetheless experiments have been conducted and successfully revealed that the dark matter of our galaxy cannot be (at least completely) composed of such objects (Alcock et al. 2000, Tisserand et al. 2007).

The large mass of galaxies or galaxy clusters, and the distance between them, enhance the probability of a detectable lensing event with respect to the previous case. For good alignments, these massive objects can distort the image of a background galaxy into giant arcs, rings, or even produce multiple images. This is referred to as *strong lensing*.

If the mass is not large enough or the alignment between background galaxies and the foreground lens is not good enough, the only effect is a small distortion of the background images. As the image distortion cannot be separated from the intrinsic shape, it is only measurable statistically. This effect is called *weak lensing* and has been applied for different purposes. It can be used to infer the mass of galaxy clusters, by correlating the image distortion to the position of the cluster. In the case of isolated galaxies, where the effect is smaller, we can stack many galaxies together and measure profiles of a galaxy type. This type of weak lensing is called *galaxy-galaxy lensing*.

The large-scale structure also imprints a weak lensing signature on the images of background galaxies, referred to as *cosmic shear*. Cosmic shear is one of the few probes that we can use to characterize the matter density field in the Universe.

This is only a brief list of lensing studies. For a better review about all lensing possibilities we recommend to consult Meylan et al. (2006). Along this chapter we overview basic concepts in gravitational lensing, and present the measure that we want to use, namely galaxy-galaxy lensing. We derive the necessary equations, starting from the deflection of a single light ray. A more detailed discussion over the topics presented in this Chapter can be found in Schneider (2006b).

We also address the description of the ray-tracing simulation from which we derive all our results. Ray-tracing simulations describe gravitational lensing inside a cosmological simulation, which in our case is the Millennium Simulation. The goal of this thesis is to use this simulated data to forecast the signals that one could measure in upcoming surveys.

4.1 Basic theory

4.1.1 Point mass deflection & thin lens approximation

The best starting point to derive the fundamental lensing equations is to consider the deflection angle $\hat{\boldsymbol{\alpha}}$ of a point mass. The deflection angle is the difference in the direction vector between the incoming and outgoing trajectory of a light ray. In terms of the mass M and the impact vector $\boldsymbol{\xi}$, the deflection angle reads

$$\hat{\boldsymbol{\alpha}} = \frac{4\,G\,M}{c^2}\frac{\boldsymbol{\xi}}{\xi^2}. \tag{4.1}$$

In Fig. 4.1 we sketch the paradigmatic situation. The distances separating observer, lens and source, are many orders of magnitude larger than the extension of the lensing mass. For that reason an excellent approximation is to assume that the lensing mass is contained within a *lens plane*. Under this approximation the trajectory of the light ray changes as it crosses the lens plane. This is called the *geometrically-thin lens* approximation.

Now we only need to describe the deflection angle at the lens plane at any angular separation $\boldsymbol{\xi}$ from the lens. Considering the mass elements $\mathrm{d}m$ at a radial distance r_i of the observer and at position $\boldsymbol{\xi'}$ in the plane, the deflection angle reads:

$$\begin{aligned}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) &= \frac{4\,G}{c^2}\sum_i \mathrm{d}m(\boldsymbol{\xi}'_i, r_i)\frac{\boldsymbol{\xi}-\boldsymbol{\xi}'_i}{|\boldsymbol{\xi}-\boldsymbol{\xi}'_i|^2} \qquad (4.2)\\ &= \frac{4\,G}{c^2}\int \mathrm{d}^2\xi' \int \mathrm{d}r\ \rho(\boldsymbol{\xi'}, r)\frac{\boldsymbol{\xi}-\boldsymbol{\xi'}}{|\boldsymbol{\xi}-\boldsymbol{\xi'}|^2},\end{aligned}$$

where in the second step we took the continuum limit, and defined a density function $\rho(\boldsymbol{\xi'}, r)$.

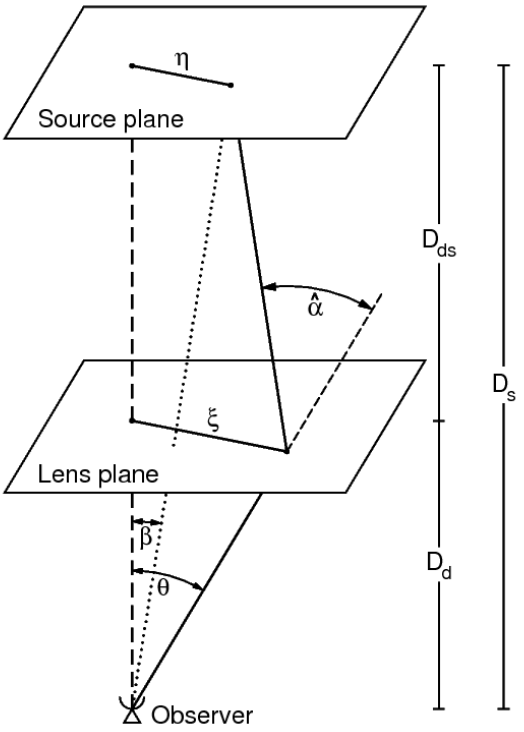

Figure 4.1: Paradigmatic situation of a lensing phenomenon. In this sketch the proportions are not realistic, as the deflection angles are normally on the range of arcminutes. From Bartelmann & Schneider (2001).

The integration along r can be carried out. Using the previous approximations we can define the *surface mass density* as

$$\Sigma(\boldsymbol{\xi}) \equiv \int \mathrm{d}r\,\rho(\boldsymbol{\xi}, r). \tag{4.3}$$

It is customary at this point to reparameterize the surface mass density. We define the *convergence* $\kappa(\boldsymbol{\xi})$ with the help of the *critical surface mass density* Σ_{crit}

$$\kappa(\boldsymbol{\xi}) = \frac{\Sigma(\boldsymbol{\xi})}{\Sigma_{\mathrm{crit}}}, \qquad \Sigma_{\mathrm{crit}} = \frac{c^2}{4\pi G}\frac{D_{\mathrm{s}}}{D_{\mathrm{ds}}D_{\mathrm{d}}}, \tag{4.4}$$

where the distances D_s, D_d and D_ds are angular diameter distances to the source, to the lens and between lens and source. A mass distribution with $\Sigma \geq \Sigma_\mathrm{crit}$, i.e. $\kappa \geq 1$ somewhere, can produce multiple images of a source. The deflection angle produced by the convergence on the lens plane reads

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{1}{\pi}\frac{D_\mathrm{s}}{D_\mathrm{d}\, D_\mathrm{ds}} \int \mathrm{d}^2\xi' \;\kappa(\boldsymbol{\xi}')\, \frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|^2}. \tag{4.5}$$

4.1.2 The lens equation

From Fig. 4.1, we can derive the relation between image and source position. The relation between source position in the source plane $\boldsymbol{\eta}$, lens position in the lens plane $\boldsymbol{\xi}$ and the deflection angle is

$$\boldsymbol{\eta} = \frac{D_\mathrm{s}}{D_\mathrm{d}}\boldsymbol{\xi} - D_\mathrm{ds}\,\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{4.6}$$

Note that the typical angles involved are extremely small, and we can use the approximation $\sin\vartheta \approx \tan\vartheta \approx \vartheta$. Now we convert the distance vectors to angular positions, which are more commonly used:

$$\boldsymbol{\eta} = D_\mathrm{s}\,\boldsymbol{\beta}, \qquad\qquad \boldsymbol{\xi} = D_\mathrm{d}\,\boldsymbol{\theta}, \tag{4.7}$$

and thereafter we define also the *scaled deflection angle*:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\boldsymbol{\alpha}}(D_\mathrm{d}\boldsymbol{\theta}). \tag{4.8}$$

The *lens equation* relates the apparent angular position $\boldsymbol{\theta}$ of the image, and the angular position $\boldsymbol{\beta}$ that the source would have if there were no lens. The position $\boldsymbol{\beta}$ is unobservable, but it can be computed in terms of $\boldsymbol{\theta}$ and the scaled deflection angle $\boldsymbol{\alpha}$:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{4.9}$$

Analyzing the lens equation, we can study how gravitational lensing distorts images. Since lensing does not change the number of photons, the source brightness distribution for a given frequency ν, $I_\nu^{(s)}(\boldsymbol{\beta})$ and the observed one, $I_\nu(\boldsymbol{\theta})$, are related through the mapping $\boldsymbol{\theta} \mapsto \boldsymbol{\beta}$:

$$I_\nu^{(s)}(\boldsymbol{\beta}(\boldsymbol{\theta})) = I_\nu(\boldsymbol{\theta}). \tag{4.10}$$

This work is based on weak lensing, therefore we focus in the following on this regime. To see what is the effect on a background image we linearize the lens mapping. The linear term to the mapping is given by the Jacobian matrix:

$$\mathcal{A}(\boldsymbol{\theta}) = \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \delta_{ij} - \frac{\partial\alpha_i(\boldsymbol{\theta})}{\partial\theta_j} = \begin{pmatrix} 1-\kappa(\boldsymbol{\theta})-\gamma_1(\boldsymbol{\theta}) & -\gamma_2(\boldsymbol{\theta}) \\ -\gamma_2(\boldsymbol{\theta}) & 1-\kappa(\boldsymbol{\theta})+\gamma_1(\boldsymbol{\theta}) \end{pmatrix}. \tag{4.11}$$

This matrix can be decomposed into an identity matrix $\mathbb{1}$, an isotropic magnification $\kappa(\boldsymbol{\theta})\,\mathbb{1}$, and an anisotropic deformation or *shear* $\gamma_1(\boldsymbol{\theta})\sigma_3 + \gamma_2(\boldsymbol{\theta})\sigma_1$, ($\sigma_1$ and σ_3 are Pauli matrices). Note that the isotropic magnification is produced by the convergence κ. Denoting a fixed point on the lens plane with $\boldsymbol{\theta}_0$ and its corresponding point in the source plane with $\boldsymbol{\beta}_0 = \boldsymbol{\beta}(\boldsymbol{\theta}_0)$, we have

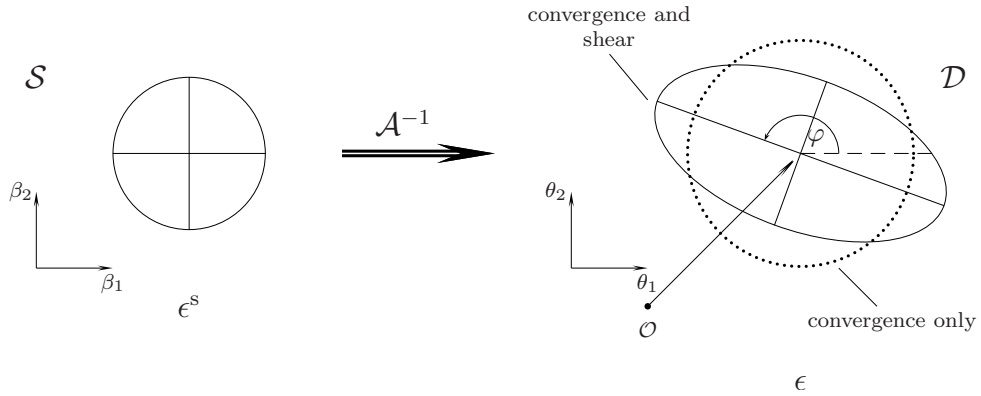

Figure 4.2: Illustration of the weak lensing effect on a circular image. Figure from Schneider (2006b).

$$\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \mathcal{A}(\boldsymbol{\theta}) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0). \tag{4.12}$$

From the last relation we can see that from a circular image with radius R, we obtain an ellipse with semi-axes $R/(1 - \kappa \pm |\gamma|)$, as shown in Fig. 4.2.

The determinant of the Jacobian matrix gives the ratio between the differential solid angle in source plane and lens plane coordinates. The enlargement or decrease of the solid angle due to the lens, changes the observed flux at the image plane. For this reason its inverse is called the *magnification factor*:

$$\mu(\boldsymbol{\theta}) = \frac{1}{\det \mathcal{A}(\boldsymbol{\theta})} = \frac{1}{\big(1 - \kappa(\boldsymbol{\theta})\big)^2 - \gamma_1^2(\boldsymbol{\theta}) - \gamma_2^2(\boldsymbol{\theta})}. \tag{4.13}$$

In principle, the magnification could be infinite, but that would involve unphysical perfect point sources, and in such a case, the geometric optics treatment used here would not be valid.

It is customary to us a complex number representation of the shear to simplify the algebra:

$$\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2 = |\gamma|\, \mathrm{e}^{\mathrm{i}2\varphi} \tag{4.14}$$

4.2 Lensing potential

The solution of many lensing problems can be simplified with a mathematical abstraction, the *lensing potential*. With a little bit of algebra one can see that the scaled deflection angle can be written as the gradient of a potential ψ,

$$\boldsymbol{\alpha} = \nabla\psi. \tag{4.15}$$

This potential allows one to describe both convergence and shear as derivatives of a scalar function. We saw in Eq. (4.11) that

$$\begin{pmatrix} 1 - \alpha_{1,1}(\boldsymbol{\theta}) & -\alpha_{1,2}(\boldsymbol{\theta}) \\ -\alpha_{2,1}(\boldsymbol{\theta}) & 1 - \alpha_{2,2}(\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} 1 - \kappa(\boldsymbol{\theta}) - \gamma_1(\boldsymbol{\theta}) & -\gamma_2(\boldsymbol{\theta}) \\ -\gamma_2(\boldsymbol{\theta}) & 1 - \kappa(\boldsymbol{\theta}) + \gamma_1(\boldsymbol{\theta}) \end{pmatrix}. \tag{4.16}$$

where we have used the notation for the derivatives: $\partial\alpha_i/\partial\theta_j = \alpha_{i,j}$. From last equation and Eq. (4.15), we find the expressions for convergence and shear in terms of the second order derivatives of the potential:

$$\gamma_1 = \frac{1}{2}\left(\psi_{,11} - \psi_{,22}\right), \quad \gamma_2 = \psi_{,12}, \quad \text{and} \quad \kappa = \frac{1}{2}\left(\psi_{,11} + \psi_{,22}\right). \tag{4.17}$$

One can see through the lensing potential that convergence and shear are not independent. We can compute the shear from the convergence using the Kaiser-Squires inversion (Kaiser & Squires 1993):

$$\gamma(\boldsymbol{\theta}) \equiv \frac{1}{\pi}\int_{\mathbb{R}^2} \mathrm{d}^2\theta'\, \mathcal{D}(\boldsymbol{\theta} - \boldsymbol{\theta'})\kappa(\boldsymbol{\theta'}) \quad \text{with} \quad \mathcal{D}(\boldsymbol{\theta}) = \frac{\theta_y^2 - \theta_x^2 - 2\,\mathrm{i}\,\theta_x\theta_y}{|\boldsymbol{\theta}|^4}. \tag{4.18}$$

The relation can be easily found by computing the Fourier transform of γ, κ and the derivatives of the potential ψ. The relations in Eq. (4.17) in Fourier space, allow us to express $\tilde{\gamma}$ in terms of $\tilde{\kappa}$. If we transform this relation back into configuration space, we obtain the Kaiser-Squires inversion. The simulation of the lensing effect therefore can be carried out by: defining the projected mass map, transforming it into convergence and finally computing the shear. This can be done efficiently with Fast Fourier Transform algorithms.

4.3 Observables

We cannot separate the distortion of the image from its intrinsic shape. The estimation of lensing for this reason is difficult in the weak lensing regime. Nonetheless, galaxies are supposed to be randomly oriented, hence the intrinsic ellipticity must average out when using many galaxies. Therefore, the average correlated ellipticity of a set of images to a lensing mass is observable. The correlated magnification could in principle also be studied. For this work however, it is not of importance.

In order to measure such induced ellipticity we proceed as follows. First we define the center of our coordinates according to the brightness distribution of the lensed image:

$$\bar{\boldsymbol{\theta}} = \frac{\int \mathrm{d}^2\theta\, I(\boldsymbol{\theta})\,\boldsymbol{\theta}}{\int \mathrm{d}^2\theta\, I(\boldsymbol{\theta})}, \tag{4.19}$$

and then we compute the second brightness moments:

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta\, I(\boldsymbol{\theta})\,(\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int \mathrm{d}^2\theta\, I(\boldsymbol{\theta})}, \quad i,j \in 1,2. \tag{4.20}$$

From these moments one can construct a complex ellipticity that can be used for our analysis

$$\epsilon \equiv \epsilon_1 + \mathrm{i}\epsilon_2 \equiv \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}. \tag{4.21}$$

One can show that the relation between the observed ellipticity ϵ and the intrinsic one ϵ_i is (Seitz & Schneider 1997):

$$\epsilon_\mathrm{i} = \begin{cases} \dfrac{\epsilon - g}{1 - g^*\epsilon} & \text{if } |g| \leq 1, \\[2ex] \dfrac{1 - g\,\epsilon^*}{\epsilon^* - g^*} & \text{if } |g| > 1, \end{cases} \tag{4.22}$$

where g is the reduced shear, a combination of convergence and shear:

$$g \equiv \frac{\gamma}{1-\kappa} = \frac{\gamma_1 + \mathrm{i}\gamma_2}{1-\kappa}. \tag{4.23}$$

This is of especial importance for us as we work with simulated data. In order to produce realistic shear catalogues we shall introduce an intrinsic ellipticity to our reduced shear catalogue to account for this source of noise. The expected ϵ for an unlensed images is zero, whereas one can show that the expectation value for a lensed image is related to the *reduced shear* g (Seitz & Schneider 1997):

$$E(\epsilon) = \begin{cases} g & \text{if } |g| \leq 1, \\ 1/g^* & \text{if } |g| > 1. \end{cases} \tag{4.24}$$

If we measure the induced reduced shear, we are able to reconstruct the underlying mass that causes the lensing.

Most of the times, there is no precise redshift information of the lensed images, or none at all. In these cases we cannot recover the underlying surface mass density, we can only infer $\kappa(\boldsymbol{\theta})$, and there will be a degeneracy in the mass model. For a given inferred $\kappa(\boldsymbol{\theta})$ model, the new $\kappa_\lambda(\boldsymbol{\theta})$ is

$$\kappa_\lambda(\boldsymbol{\theta}) = 1 + \lambda\big(\kappa(\boldsymbol{\theta}) - 1\big), \quad \text{such as} \quad \lambda\big(1-\kappa(\boldsymbol{\theta})\big) = 1 - \kappa_\lambda(\boldsymbol{\theta}), \tag{4.25}$$

whereas the shear is $\gamma_\lambda = \lambda\,\gamma$. Therefore the reduced shear, the observable, does not change. The only consequence is a change on the source plane, which is physically unobservable. This degeneracy is called *mass sheet degeneracy*, and it is a major issue in many lensing observations.

If one has redshift information, the degeneracy can be broken. The convergence of the same lens for different redshifts has a different proportion to the mass model. If $\Sigma'(\boldsymbol{\theta})$ is the underlying projected mass, for two different redshifts z_1, z_2 we have

$$\kappa_1(z_1, \boldsymbol{\theta}) = \Sigma_{\mathrm{crit}}^{-1}(z_1)\,\Sigma'(\boldsymbol{\theta}), \quad \kappa_2(z_2, \boldsymbol{\theta}) = \Sigma_{\mathrm{crit}}^{-1}(z_2)\,\Sigma'(\boldsymbol{\theta}). \tag{4.26}$$

Since for each redshift the degeneracy is different, we have

$$\begin{aligned} \kappa_{\lambda_1}(z_1, \boldsymbol{\theta}) &= 1 + \lambda_1\big(\Sigma_{\mathrm{crit}}^{-1}(z_1)\,\Sigma'(\boldsymbol{\theta}) - 1\big), \\ \kappa_{\lambda_2}(z_2, \boldsymbol{\theta}) &= 1 + \lambda_2\big(\Sigma_{\mathrm{crit}}^{-1}(z_2)\,\Sigma'(\boldsymbol{\theta}) - 1\big). \end{aligned} \tag{4.27}$$

If we can reconstruct both models for N points, we have a system of $2N$ equation to reconstruct $\Sigma'(\boldsymbol{\theta})$ at N positions plus the two degeneracy parameters λ_1 and λ_2. The degeneracy is in theory broken.

4.4 Galaxy-galaxy lensing

Having presented some very basic concepts about gravitational lensing, we focus on the probe used for this thesis. We measure the average galaxy-galaxy lensing for a galaxy sample. We need a set of galaxies to produce a high enough signal-to-noise ratio, which

in our case are satellite galaxies in clusters. In galaxy-galaxy lensing we measure around each lens the *tangential ellipticity* of the background images

$$\epsilon_{\rm t}(\boldsymbol{\theta}) = -\epsilon_1(\boldsymbol{\theta})\cos(2\phi) - \epsilon_2(\boldsymbol{\theta})\sin(2\phi), \tag{4.28}$$

where ϕ is the polar angle, once we take the lens center as center of coordinates, and $\boldsymbol{\theta}$ the angular position of the image (Fig. 4.3). Then we average over the angle ϕ and over the galaxy sample to improve the quality of the signal.

We want to work in the weak lensing regime. In this case it is common to make the following approximations

$$1 - \kappa \approx 1, \quad g \approx \gamma, \quad \epsilon = \epsilon_{\rm i} + \gamma, \tag{4.29}$$

where we denoted with $\epsilon_{\rm i}$ the intrinsic ellipticity of a lensed image. Nevertheless we shall test the validity of this approximation in Chap. 7. From Eq. (4.24), and under the weak lensing regime we know that the expectation value of the preceeding quantity is the *tangential shear* $\gamma_{\rm t}(\theta)$ as a function of angular separation. An isolated axi-symmetric lens produces only tangential shear.

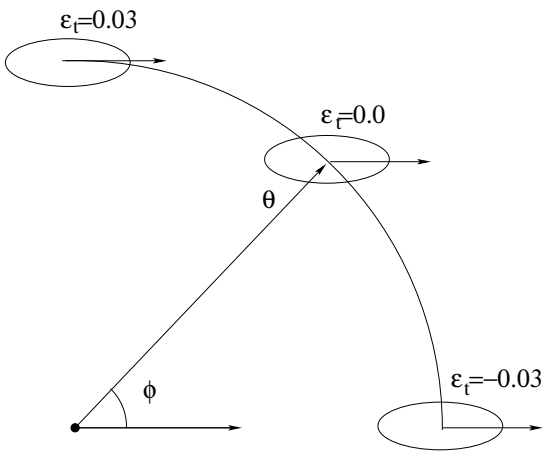

Figure 4.3: Sketch of the measurement of the correlated ellipticity toward a lensing point. The image with an intrinsic ellipticity of $\epsilon_{\rm i} = 0.03$, has a different tangential ellipticity towards the lensing center depending on the relative angle.

We measure $\gamma_{\rm t}(\theta)$ at different angular separations θ from the center of the galaxy. Then we rescale the angular separation θ into physical distance ξ and the tangential $\gamma_{\rm t}(\xi)$ shear into *excess surface mass density* $\Delta\Sigma(\xi)$. One can show (see Schneider 2006b) that $\gamma_{\rm t}(\xi)$ and $\Delta\Sigma(\xi)$ are related by

$$\Delta\Sigma(\xi) = \bar{\Sigma}(\xi) - \Sigma(\xi) = \gamma_{\rm t}(\xi)\,\Sigma_{\rm crit}, \tag{4.30}$$

where $\bar{\Sigma}(\xi)$ denotes the average surface mass density inside a circle of radius ξ and $\Sigma(\xi)$ is the average surface mass density at radius ξ. The estimator we shall use is

$$\widehat{\Delta\Sigma}(\xi) = \Sigma_{\rm crit}(z_{\rm s}, z_{\rm d})\,\epsilon_{\rm t}(\theta\, D_<(z_{\rm d})), \tag{4.31}$$

where $z_{\rm s}$, $z_{\rm d}$ are source and lens redshift, and we compute ξ from its angular distance $D_<(z_{\rm d})$. The expectation value of our estimator, over the ensemble of lenses, assuming that the intrinsic ellipticity noise is not redshift dependent, is the underlying $\Delta\Sigma(\xi)$.

The quantity $\Delta\Sigma(\xi)$ that we measure is insensitive to a constant mass sheet. A constant mass sheet contributes equally to $\bar{\Sigma}(\xi)$ and to $\Sigma(\xi)$, so the final contribution to $\Delta\Sigma(\xi)$ cancels out. This argument reflects the fact that a constant sheet of mass does not produce shear. Therefore, we could always add a constant mass sheet to the mass density profiles derived. In realistic measurements, we can assume that the possible mass sheet is negligible.

Galaxies near in redshift may be dynamically connected, and as a consequence, the intrinsic ellipticity may no longer be randomly oriented. Errors in the redshift determination have also an influence on the final result. However we consider these effects to be higher order corrections. We leave a more thorough study for future works.

The goals we want to achieve as well as a review of previous results are discussed in the next chapter.

4.5 Ray-tracing

Previously we depicted extremely simple situations. On real data many problems have to be taken into account: we must consider lensing contributions from large-scale structure; the signals depend on the redshift distribution of both lenses and images; and we must also consider the impact of observational limits into the final measurement, such as luminosity cuts.

In order to compute realistically what is to be expected from gravitational lensing surveys, numerical simulations must be conducted. With the aid of these simulations, one can predict the final signal and many systematic effects that may occur. In this thesis we use *ray-tracing* simulations. Ray-tracing uses a previously computed cosmological simulation, which in our case was the Millennium Simulation, to obtain simulated lensing surveys.

The ray-tracing we use was produced by Hilbert et al. (2009). Many of the choices for this ray-tracing simulations were due to the characteristics of the Millennium Simulation (see Hartlap 2005). We focus on the choices made to compute our data. In order to obtain galaxy observables, ray-tracing must be combined with the semi-analytical catalogues, which we discussed in Sect. 3.7.

Once the dark matter simulation is completed, the output boxes at different redshifts are combined to form an observed universe. In order to improve the amount of information and to avoid seeing the same structure twice in the same field of view, the output boxes are placed in a skewed angle with respect to line of sight (see Fig. 4.4). Then the matter distribution is discretized into lens planes to allow for an efficient computation.

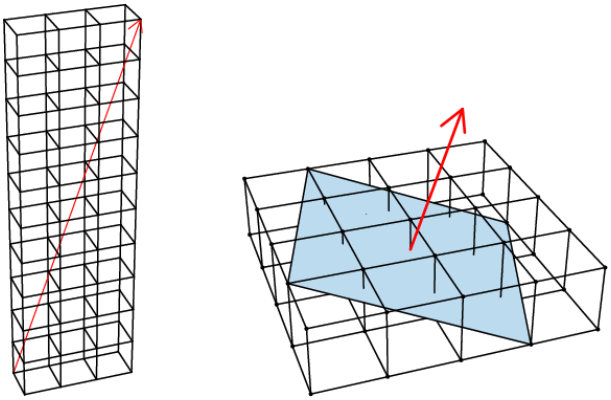

Figure 4.4: Here we show how the cosmological simulation boxes at different redshift are placed with respect the line of sight, and how the planes are build so the observed sky size is maximized. In this way we obtain the matter distribution along the line of sight. From Hilbert et al. (2009).

The separation between these skewed planes is smaller than the box size. With this choice we do not need to project a whole box into a plane, and we avoid a coarse discretization. This particular choice also allows to preserve the large-scale structure between close planes.

The goal is to compute the lensing quantities (shear and convergence) for any light ray coming from any given plane to the observer. We shall define henceforth a map of observed light rays, and compute which angular position $\boldsymbol{\beta}^{(k)}$ in the source plane, for each plane (k), corresponds to an observed angular position. The angular coordinates at the first plane $(\boldsymbol{\beta}^{(0)})$ represent the observed angular position $\boldsymbol{\theta}$. The efficient way to do so, is to define a grid of light rays at the plane $\boldsymbol{\beta}^{(0)}$ and follow them backwards as they cross the different planes in the simulation. The lensing for any position can be computed by interpolation between the array of light rays. The most efficient way to proceed is to find the Jacobian $A_{ij}^{(k)}$ between each of the source planes (k) and the observer.

It is possible to find a recurrence relation between the different Jacobians at different planes. The first step is to compute the *dimensionless surface mass density* $\sigma^{(k)}\left(\boldsymbol{\beta}^{(k)}\right)$

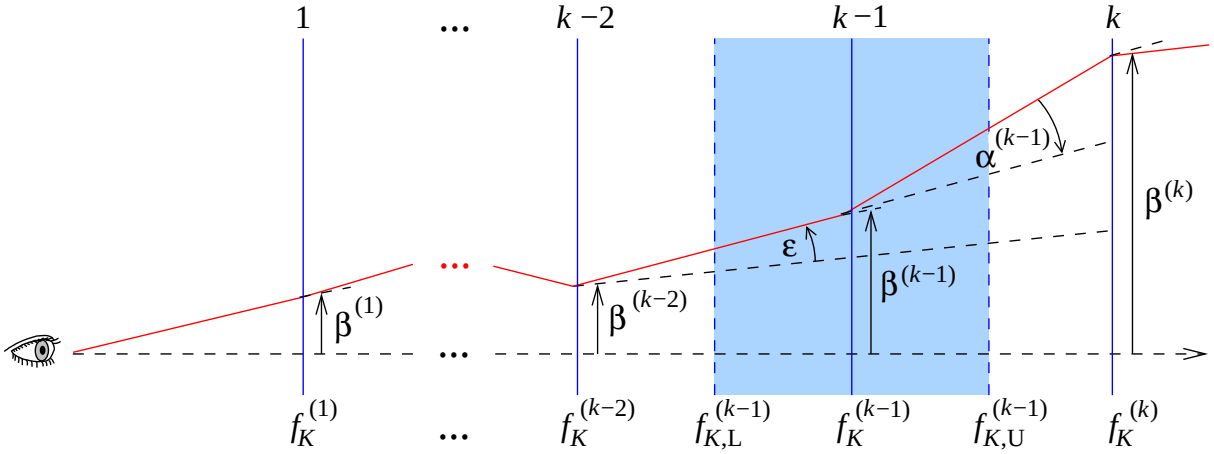

Figure 4.5: Sketch of a ray-tracing situation. The vertical solid blue lines are lens planes where we have projected the mass corresponding to a slab, illustrated in one case with the shaded blue region. In red we display the deflection for one light ray. From Hilbert et al. (2009)

map associated with each plane (k). The plane is at a radial comoving distance $\chi^{(k)}$, defined along the line of sight. For each plane we use the volume slice between the planes at $\chi_U^{(k)} = \chi^{(k)} + \Delta\chi^{(k)}/2$ and $\chi_L^{(k)} = \chi^{(k)} - \Delta\chi^{(k)}/2$. In order to obtain $\sigma^{(k)}\big(\boldsymbol{\beta}^{(k)}\big)$ we project the density contrast $\delta\big(\boldsymbol{\beta}^{(k)}, \chi'\big)$ of this volume slice into the (k) plane:

$$\begin{aligned}\sigma^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) &= \frac{4\pi G}{c^2} D_< \, a^{(k)} \int_{\chi_U^{(k)}}^{\chi_L^{(k)}} \mathrm{d}\chi' \, \rho\big(\boldsymbol{\beta}^{(k)}, \chi'^{(k)}\big) - \bar{\rho}\big(\chi^{(k)}\big) \\ &= \frac{3\, H_0^2 \Omega_\mathrm{m}}{2c^2} \frac{\chi^{(k)}}{a^{(k)}} \int_{\chi_U^{(k)}}^{\chi_L^{(k)}} \mathrm{d}\chi' \, \delta\big(\boldsymbol{\beta}^{(k)}, \chi'\big), \end{aligned} \tag{4.32}$$

We express also $\sigma^{(k)}\big(\boldsymbol{\beta}^{(k)}\big)$ as a function of the matter density to underline its relation with the convergence $\kappa\big(\boldsymbol{\beta}^{(k)}\big)$. The quantity σ equals the convergence κ in the abstract case of sources at infinite distance (light-rays perpendicular to the lens plane). In fact, we can use Eq. (4.18) changing κ for σ and compute the quantity λ, which equals the shear produced by the lensing plane also for sources at infinite distance. With σ and λ we define the *shear matrix* U, at the (k) lens plane

$$U_{ij}^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) = \frac{\partial \alpha_i^{(k)}\big(\boldsymbol{\beta}^k\big)}{\partial \beta_j^{(k)}} = \begin{pmatrix} \sigma^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) + \lambda_1^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) & \lambda_2^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) \\ \lambda_2^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) & \sigma^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) - \lambda_1^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) \end{pmatrix}. \tag{4.33}$$

Now we must define the angular position $\boldsymbol{\beta}^{(k)}$ of a light-ray at the (k) plane, with respect to its position at the previous planes $(k-1)$, and $(k-2)$. From Fig. 4.5 we can extract the necessary relation:

$$\chi^{(k)}\boldsymbol{\beta}^{(k)} = \chi^{(k)}\boldsymbol{\beta}^{(k-2)} + \big(\chi^{(k-2)} - \chi^{(k)}\big)\boldsymbol{\epsilon} - \big(\chi^{(k-1)} - \chi^{(k)}\big)\boldsymbol{\alpha}^{(k-1)}\big(\boldsymbol{\beta}^{(k-1)}\big), \tag{4.34}$$

where

$$\boldsymbol{\epsilon} = \frac{\chi^{(k-1)}}{\chi^{(k-2)} - \chi^{(k-1)}}\big(\boldsymbol{\beta}^{(k-1)} - \boldsymbol{\beta}^{(k-2)}\big). \tag{4.35}$$

This same relation allows us to compute the angular position $\boldsymbol{\beta}^{(k)}$ of the light-ray at the plane (k) as a function of the observed position $\boldsymbol{\theta}$.

In order to compute the lensing effects for each light-ray along its travel, we solve for $\boldsymbol{\beta}^{(k)}$ in Eq. (4.34), and we differentiate the whole expression with respect to $\boldsymbol{\theta}$ to obtain the recursive relation between Jacobians $A_{ij}(\boldsymbol{\theta})$:

$$\begin{aligned} A_{ij}^{(k)}(\boldsymbol{\theta}) = &- \left(\frac{\chi^{(k-1)} - \chi^{(k)}}{\chi^{(k)}} \right) U_{ik}^{(k-1)}\big(\boldsymbol{\beta}^{(k-1)}(\boldsymbol{\theta})\big)\, A_{ik}^{(k-1)}(\boldsymbol{\theta}) \\ &+ \left(\frac{\chi^{(k-1)}}{\chi^{(k)}} \frac{\chi^{(k-2)} - \chi^{(k)}}{\chi^{(k-2)} - \chi^{(k-1)}} \right) A_{ij}^{(k-1)}(\boldsymbol{\theta}) \\ &- \left(\frac{\chi^{(k-2)}}{\chi^{(k)}} \frac{\chi^{(k-1)} - \chi^{(k)}}{\chi^{(k-2)} - \chi^{(k-1)}} \right) A_{ij}^{(k-2)}(\boldsymbol{\theta}). \end{aligned} \tag{4.36}$$

This final relation allows us to compute all Jacobians in a sequence and with them the lensing observables. In order to compute the Jacobian for the first lens plane we assign the identity matrix to the virtual planes $k = 0$ and $k = -1$.

So far we have a total of 512 square degrees up to redshift $z = 4$. The 128 fields of 4 square degrees in which the simulated survey is divided, are almost completely independent. Only in a few occasions the structures observed on different patches correspond to the same object at a different redshift output. The data available for this thesis was composed by the Jacobians $A_{ij}^{(k)}(\boldsymbol{\theta})$ as a function of angular coordinate, between the observed sky plane $(k = 0)$ and the source planes located at the same radial distances as the lens planes. With them we can produce a source distribution at will, and analyze the outcome of different surveys with different specifications.

The final goal of galaxy-galaxy lensing is to infer the underlying projected mass. Therefore, we also analyze the projected mass map of each lens plane. The slices that the planes define are large enough, so almost all the objects we want to study are contained in a single plane. We compare the results from the projected mass maps and from the shear catalogues to estimate the precision of our measurements and the scope of our conclusions.

Since only the Jacobians are kept, we had to reconstruct the projected mass maps again. They were obtained inverting the previous process. First, we solved for the shear matrices for each plane (k), out of the Jacobians from the planes (k), $(k+1)$, and $(k-1)$, (Eq. 4.36). Then the dimensionless surface mass density is obtained from their trace:

$$\sigma^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) = \frac{1}{2}\left(U_{11}^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) + U_{22}^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) \right). \tag{4.37}$$

Inverting Eq. (4.32)

$$\Sigma^{(k)}(\boldsymbol{\beta}) = \int_{\chi_U^{(k)}}^{\chi_L^{(k)}} \mathrm{d}\chi'\ \rho\big(\boldsymbol{\beta}^{(k)}, \chi'^{(k)}\big) = \sigma^{(k)}\big(\boldsymbol{\beta}^{(k)}\big) \frac{c^2}{4\pi G\, D_<\, a^{(k)}} + \int_{\chi_U^{(k)}}^{\chi_L^{(k)}} \mathrm{d}\chi'\ \bar{\rho}(\chi^{(k)}),$$

we end up with the reconstruction of the mass maps $\Sigma^{(k)}(\boldsymbol{\beta})$.

5. Substructure & galaxy-galaxy lensing

In previous chapters we reviewed the basic concepts necessary to understand the current standard cosmological model. We also introduced the cosmological probe that we want to use in this thesis, namely gravitational lensing, and we described the data used.

This chapter is devoted to a formal description of the cosmological problem on which this thesis aims to shed light: substructure in galaxy clusters. We propose a particular galaxy-galaxy lensing method to measure mass profiles of satellite galaxies in clusters. With this thesis we want to forecast the lensing signal for our method in future surveys.

The chapter is organized as follows. First we discuss what we can learn with lensing about the mass profiles of sub-halos. We consider previous works and the differences with respect to our proposal. Then we present our galaxy-galaxy lensing method. We inspect its limitations and the impact of the redshift distribution of the galaxies involved. At the end of the previous chapter we have outlined how our lensing simulations are produced, and the data we use. In this chapter we present examples of the projected mass maps $\Sigma(\boldsymbol{\beta})$ that we reconstructed. Finally, we motivate the necessity of a thorough check of the proposed method. We also give details of our simulated cluster sample hosting the sub-halos.

5.1 Substructure in clusters

A paradigm in structure formation is that small mass overdensities, such as galaxy halos, form first. Large ones, corresponding to galaxy clusters, form later through merging processes. This is also what one can observe in simulations. In the current structure formation models, a typical galaxy cluster consists of a massive main halo of dark matter, in the center of which we found a bright central galaxy (BCG). Within this main halo, we can find local overdensities, sub-halos, hosting a satellite galaxy in their centers.

One of the most interesting questions is how sub-halos evolve after they become part of the cluster. This evolution is also essential to understand how hosted galaxies evolve. In the current model, sub-halos are stripped of the outermost layers by gravitational tidal forces. This stripping changes the original radial mass profile. This effect is customarily parameterized by a so-called *truncation radius*. Detecting the truncation radius is one of the main targets of this thesis.

In order to measure the dark matter halo, gravitational lensing is one of the most powerful probes. With lensing we can detect mass at large scales around the galaxy where there is no luminous matter, and it does not depend on its dynamical state.

The mass profiles of sub-halos have already been studied in observations using gravitational lensing. The works published so far focus on analyzing single clusters. In these

works, out of all the constraints coming from both strong and weak lensing analyses, two parameters are estimated for each sub-halo: the already mentioned truncation radius and the central density of the sub-halo. These parameters fix the mass profile. The most relevant papers in this context are: Limousin et al. (2007), based on a method discussed in Limousin et al. (2005) (see also Schneider & Rix 1997); and Natarajan et al. (2007) which coincides in their methods for strong lensing, but differs in the weak lensing treatment (Natarajan & Kneib 1997). The works by Halkola et al. (2007) and Suyu & Halkola (2010) based only on strong lensing analyses, seem to agree with the previous ones.

The method in the papers mentioned involves modeling the global lensing signal coming from the cluster, therefore the results depend on the model for the cluster. Except for Suyu & Halkola (2010), the method also relies on scaling relations between galaxy luminosity and the parameters obtained. These scaling relations are necessary as they are the only way to reduce the number of parameters and produce robust results. However, they introduce strong assumptions. The results are derived from an assumed model for the sub-halo mass profile. There is no direct measurement of the sub-halo profile, but rather an inference.

The method that we develop in this thesis is meant for large cluster samples. Although we are also forced to make assumptions and use priors, we derive the mass profiles in a non-parametric way. Our method is the next step in lensing research on sub-halos. The method we propose uses galaxy-galaxy lensing and is based on previous works e.g. Hoekstra et al. 2004, Mandelbaum et al. 2006, Parker et al. 2007, Mandelbaum et al. 2008, Tian et al. 2009.

5.2 The method

As we introduced in Sect. 4.4, galaxy-galaxy lensing measures the mean tangential shear profile ($\gamma_t(\theta)$) at different radii θ, of a galaxy sample in oder to obtain the average mass profile. If redshift information is available we can convert angular separation θ to physical distances ξ, and transform the tangential shear γ_t into excess surface mass density $\Delta\Sigma$. For the lensing surveys on which we want to apply our method, we consider that we can always measure $\Delta\Sigma(\xi)$.

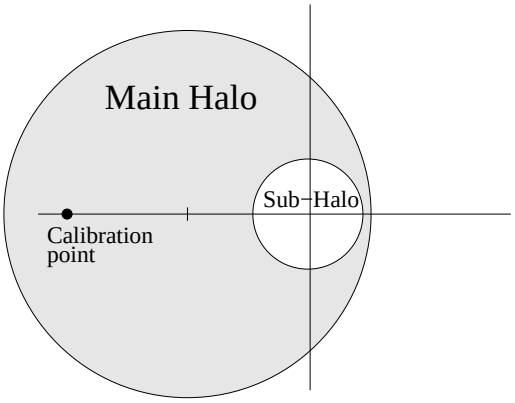

Figure 5.1: Sketch of the situation we want to address with the main halo, a sub-halo and its calibration point. calibration point.

Galaxy-galaxy lensing however cannot be applied directly to sub-halos, at least not without understanding that the signal is going to have a strong contribution coming from the main halo. We are not measuring an isolated object, and therefore our signal is not coming only from our target. To illustrate this, we present Fig. 5.1.

Our method estimates the main halo contamination and subtracts it out from the signal. We measure two signals for each sub-halo: one around the sub-halo center and another around the *calibration point* (Fig. 5.1). We assume that the main halo contribution is the same around both points on average. The calibration point is drawn at the same distance from the main halo center as the sub-halo is, but in the opposite direction as seen from the cluster center. In order to recover the sub-halo $\Delta\Sigma(\xi)$ profile, we subtract

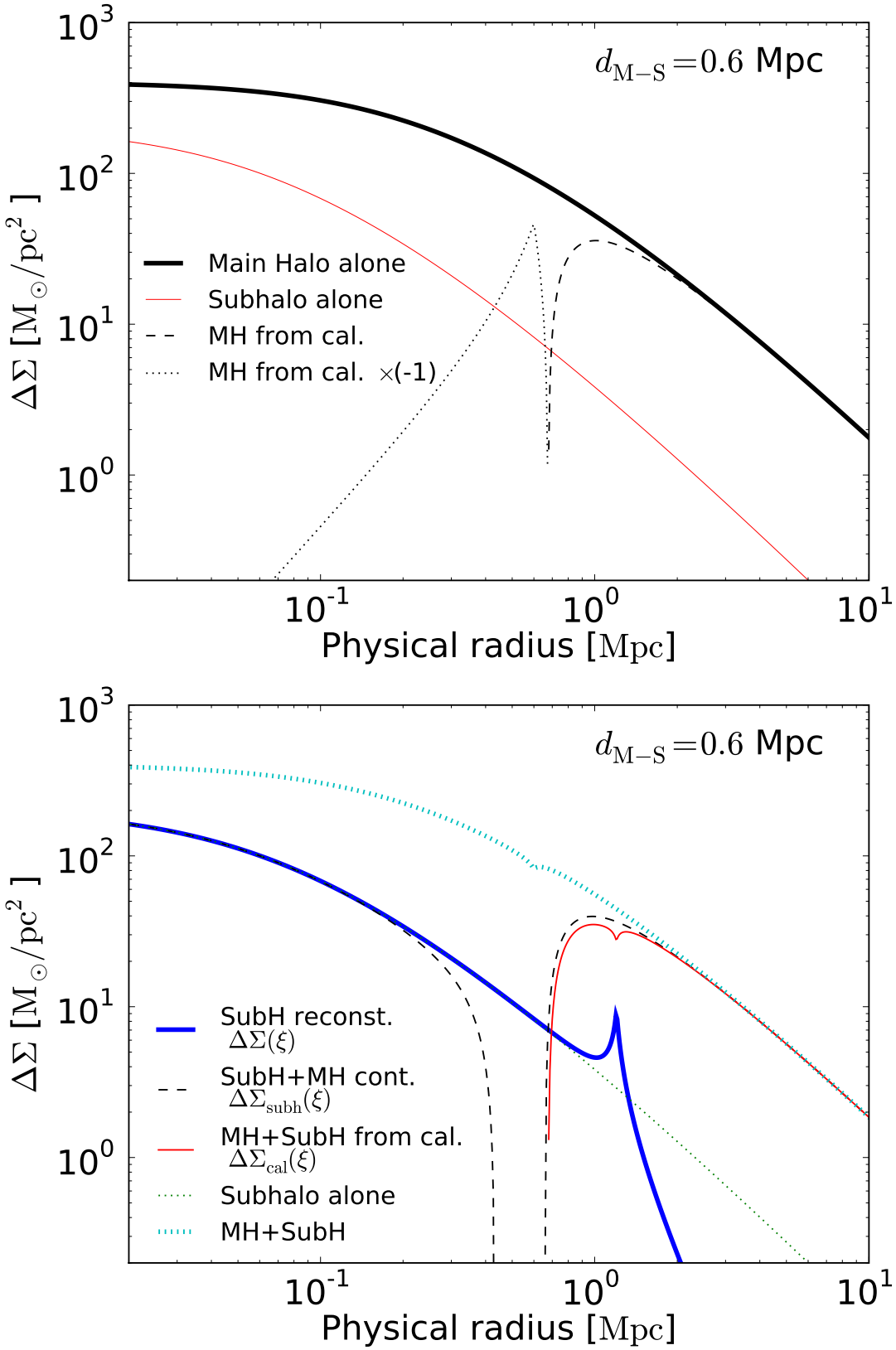

Figure 5.2: Top panel: the signal of the sub-halo alone (thin solid red line) and the main halo alone (thick black line); also the main halo alone seen from the calibration point (dashed line), with the negative part multiplied by −1 (dotted line). Bottom panel: the main halo plus the sub-halo in cyan (dotted line); the original sub-halo in pink (thin dotted line); the observed sub-halo in black (dashed line); the calibration signal in red (solid line) and the reconstructed sub-halo in blue (solid thick line). In these examples, the distance between the main halo center and the sub-halo center is 0.6 Mpc.

the signal from the calibration point from the signal measured around the sub-halos. To use this calibration scheme, we only need to assume that on average the cluster is parity-symmetric with respect to the halo center. Not considering the contamination of the main cluster in the analysis leads to misidentify the sub-halo profile.

We illustrate our method in Fig. 5.2 drawing mass profiles in a situation like the one sketched in Fig. 5.1. We use NFW mass profiles (Navarro et al. 1997) to compute plausible signals. The definition of the profiles, and how we compute the tangential shear from a mass around a point different from its center, can be found in the Appendix. We chose realistic values for the parameters to give an idea of the magnitudes that we can encounter.

In the top panel in Fig. 5.2, we depict the $\Delta\Sigma(\xi)$ profiles that we are going to assign to the main halo and to the sub-halo, using a solid thick black line and a solid red line respectively. Since we are dealing with pure theoretical profiles we can quantify the contamination that the main halo produces in the chosen configuration (dashed and dotted black). The contamination signal becomes negative (dotted black) when $\Sigma(\xi)$ gets bigger than $\bar{\Sigma}(\xi)$ (see Sect. 4.4), which happens if we measure a profile around a point which is not the mass center. We show how to compute it in the Appendix.

In the bottom panel of Fig. 5.2 we show the results after applying of our calibration scheme. One can see how the raw signal measured around the sub-halo (dashed black) is heavily truncated compared to the original one (dotted purple) due to the main halo contamination. We can recover the desired sub-halo signal (thick solid blue) $\widehat{\Delta\Sigma}(\xi)$ by:

$$\widehat{\Delta\Sigma}(\xi) = \widehat{\Delta\Sigma}_{\rm subh}(\xi) - \widehat{\Delta\Sigma}_{\rm cal}(\xi), \tag{5.1}$$

where $\widehat{\Delta\Sigma}_{\rm subh}(\xi)$ is the signal around the sub-halo (dashed black), and $\widehat{\Delta\Sigma}_{\rm cal}(\xi)$ the signal around the calibration point (solid red). We can proceed in this way because any lensing signal is a linear combination of the contribution of every halo in our weak lensing approximation; and the angular averaging needed to obtain a radial profile is also a linear process.

In the signal around the calibration point one can see a *dip* making the line non-smooth, which appears again in the reconstructed sub-halo as a *peak*. This results from the fact that the calibration signal also "sees" the sub-halo. In a realistic measurement, the averaging over signals coming from sub-halos at slightly different distances from the main halo washes out this feature.

For distances beyond the separation between sub-halo center and main halo center $\mathrm{d_{M-S}}$, our calibration is not possible in practice. For twice this separation, the calibration is already not physically possible, as the calibration includes the sub-halo itself. In this case we cannot distinguish between sub-halo and main halo without further assumptions. For distances between $\mathrm{d_{M-S}}$ and $2\,\mathrm{d_{M-S}}$, the calibration is theoretically possible but in practice very difficult. The profiles that we combine are very steep at this physical radius (note in the lower panel, how both measurements are almost vertical lines around 0.7 Mpc). As a consequence, our ability to obtain the reconstructed sub-halo profile above the noise decreases.

We obtain the mass profiles by fitting a model to the measured $\Delta\Sigma(\xi)$ profile. A direct inversion of the relation between mass profile and $\Delta\Sigma(\xi)$ is only effective for a fine sampling, which is hard to achieve. In any case, with galaxy-galaxy lensing we can measure a profile for the sub-halo. The process is therefore more transparent than those

in previous works.

Since we are combining the signals coming from a set of galaxies, the first thing to bear in mind is that we are trading a lower knowledge on individual galaxies for a higher signal-to-noise. One has to classify galaxies and produce sub-samples as homogeneous as possible. Otherwise we may mix different trends and reach wrong conclusions. We shall treat this in more detail in Chap. 7.

A cornerstone of our method is how the different measurements coming from different galaxies are combined into the final result. We address here this issue, as it is common to all our results presented later on. The signals coming from different galaxies do not offer the same amount of information. The background population of galaxies we use to probe the sub-halos has a redshift distribution. Lenses at high redshift therefore are much more sparsely probed than those at low redshift. Moreover, the lens efficiency depends on the observer-lens-source configuration. Altogether, this means we should not consider every measured $\widehat{\Delta\Sigma}_i(\xi)$ equally. In order to obtain the best signal-to-noise we compute the final $\widehat{\Delta\Sigma}(\xi)$ through a weighted mean:

$$\widehat{\Delta\Sigma}(\xi) = \sum_i w_i \, \widehat{\Delta\Sigma}_i(\xi), \tag{5.2}$$

where the sum runs over all pairs formed with a background image and a lensing galaxy. When using shear catalogues the weights for each $\widehat{\Delta\Sigma}_i(\xi)$ are

$$w_i = \frac{\Sigma^{-2}_{\mathrm{crit},i}(z_\mathrm{s}, z_\mathrm{d})}{\sum_j \Sigma^{-2}_{\mathrm{crit},j}(z_\mathrm{s}, z_\mathrm{d})}, \tag{5.3}$$

where $\Sigma_{\mathrm{crit},i}(z_\mathrm{s}, z_\mathrm{d})$ is defined in Sect. 4.1. The weights are a simplified versions of those used by Mandelbaum et al. (2008), since we neglect the error from the ellipticity estimation. We assume a perfect determination of the redshift.

The weighting used makes our estimator more sensitive to lenses in a particular redshift range, which depends on the redshift distribution of both lenses and sources. The weighting also determines the way we obtain the necessary covariance matrices for the model fitting. In this case, the best approach is to compute the covariance matrices with a bootstrapping algorithm.

5.3 Analyzing the method and the data

The data we want to use are simulated shear catalogues resembling those of real surveys. The shear catalogues come from the ray-tracing simulations described in Sect. 4.5. But prior to that we test the method in several ways.

To begin with, we want to know how much the profile that we measure deviates from the underlying "true" one. For that purpose we reconstruct from the lensing Jacobians the projected mass maps $\Sigma(\boldsymbol{\beta})$ for each lens plane as explained at the end of Sect. 4.5. Having $\Sigma(\boldsymbol{\beta})$ allows us to compute $\Delta\Sigma(\xi)$ with a higher precision than with lensing catalogues. We also have the advantage to have the highest possible resolution and no contamination from background or foreground masses, as we consider isolated planes. One can find two examples of these reconstructed maps in Fig. 5.3. All the reconstructed maps are squares

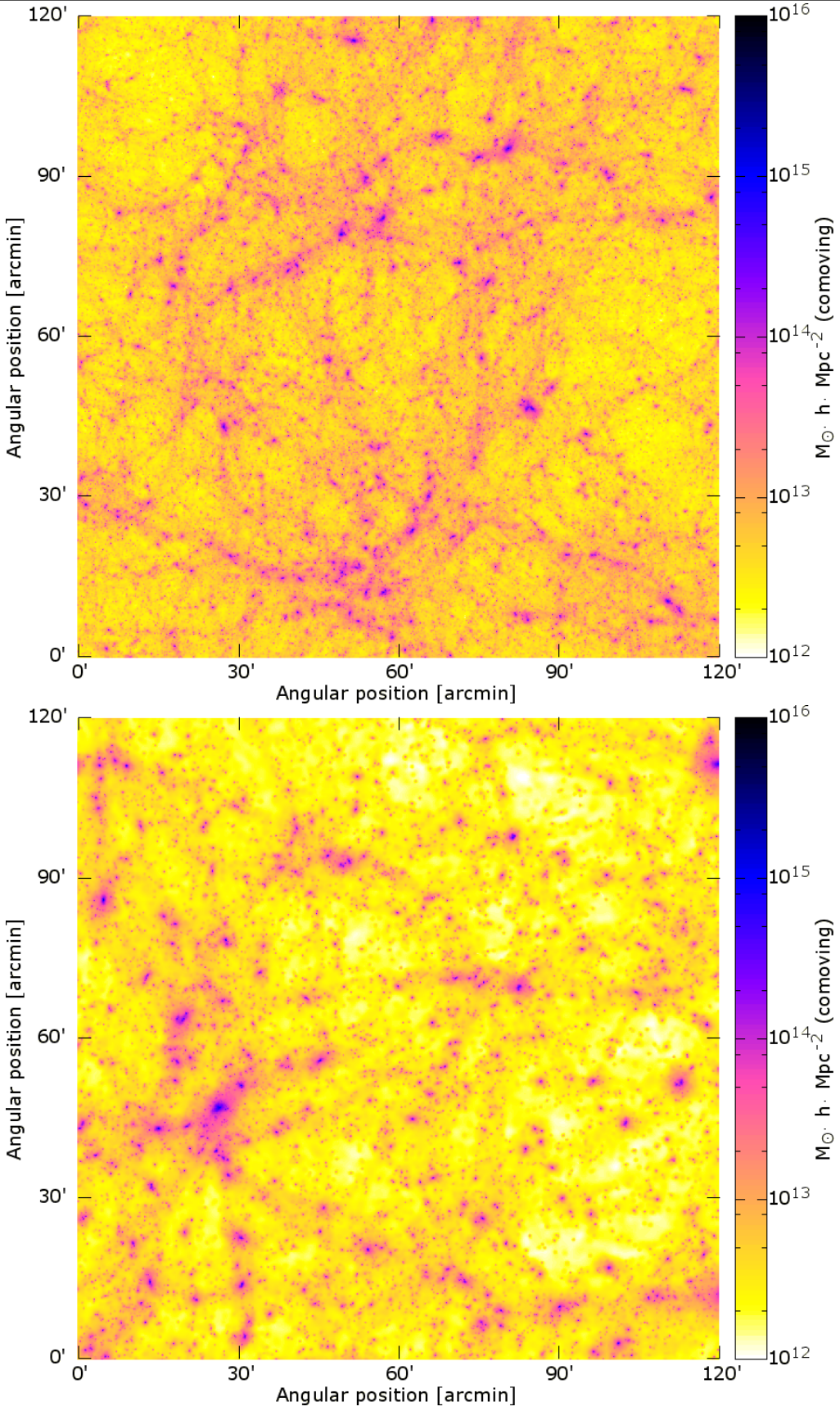

Figure 5.3: Two full lens planes. The side length in each represents an angular size of 2°, with a resolution of 4096 pixels. The top one is at redshift 0.4566, the bottom one at redshift 0.1749. This means the the physical side length is $\sim$ 40.7 Mpc in the top panel and $\sim$ 20.6 Mpc in the bottom one. Note the large overdensity in the bottom panel in the lower left quadrant, as we use it as an example in the following chapters.

with a side length of two degrees. On them one can see overdensities corresponding to isolated galaxies, galaxy groups and clusters.

Measuring $\Delta\Sigma(\xi)$ directly from $\Sigma(\boldsymbol{\beta})$ assures us to get the best signal coming from the simulation. With this measurement we shall characterize the Millennium Simulation sub-halos.

As we mentioned in Sect. 3.6 there is not a unique way of separating the main halo and sub-halos. Nonetheless the validity of our method relies on its ability to distinguish between them. In order to analyze its performance we need data with a clear definition of both contributions. In Sect. 6.1 we propose a test for the calibration, which addresses these problems.

In the tests for our calibration method and in the measurement of $\Delta\Sigma(\xi)$ with projected mass maps, we have pixel maps instead of shear catalogues with discrete galaxies. In these cases we still compute the signal through a weighted mean:

$$\widehat{\Delta\Sigma}(\xi) = \sum_k w'_k \, \widehat{\Delta\Sigma}'_k(\xi), \tag{5.4}$$

but here $\widehat{\Delta\Sigma}'_k(\xi)$ is measured considering the projected mass values within the radius ξ around a lensing sub-halos. The sum runs over all sub-halos in the sample studied. The weights in this case for each $\widehat{\Delta\Sigma}'_k(\xi)$ are

$$w'_k = \frac{n_k}{\sum_j n_j}, \tag{5.5}$$

where n_k is the number of pixels used in each measurement. These weights consider the pixel resolution in each measurement. The covariance matrices that we need in these cases are changed accordingly.

5.3.1 Clusters in the Millennium Simulation

We define as cluster a FOF group whose main halo has a mass according to the SUBFIND algorithm larger than $10^{14}\,\mathrm{M}_\odot/h$. We use this criterion in order to avoid FOF groups where a large fraction of the mass is not gravitationally bound and only close in position. Our simulated survey has 128 patches of 4 square degrees. All clusters we analyze are below $z = 0.9055$. We consider 20 lens planes obtained with the last 20 outputs of the Millennium Simulation. High mass concentrations at high redshifts are likely to have a different evolution and characteristics to those overdensities at low redshift. Including high-redshift objects into the sample is likely to produce less homogeneous samples. Moreover, it is more difficult to probe them with lensing since there are fewer background sources available. Therefore their inclusion offers no improvement in the detectability of the signal. We imposed the redshift limit for this reason.

Clusters near the borders of our survey patches cannot be measured properly, hence we exclude them from the sample. We define a square centered on the main halo using the sub-halo distribution. The dimensions of the rectangle equal the standard deviation of the sub-halo position, along each of the two spatial coordinates. We exclude all clusters whose area, defined in this way, is not fully inside the field of view. The total number of clusters after this filtering is 2991. All of them are used for the tests of our calibration and to characterize sub-halos in the Millennium Simulation.

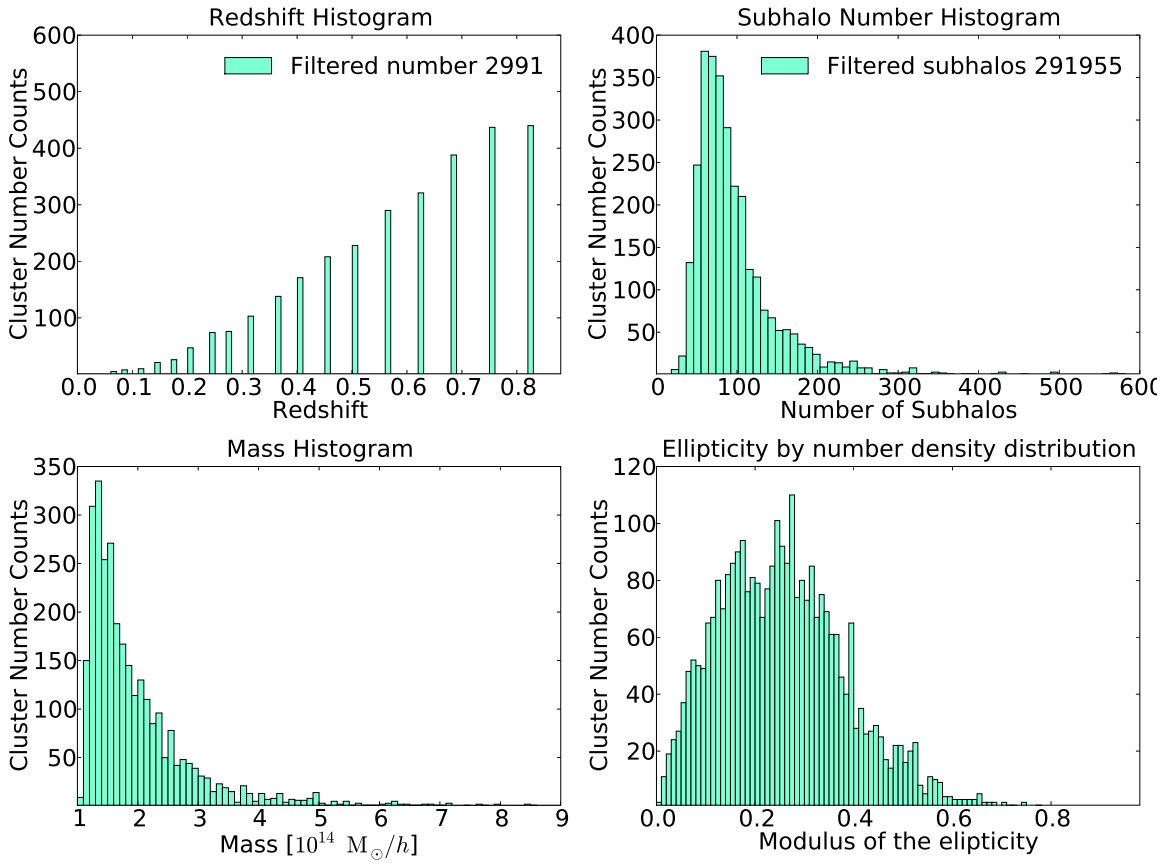

Figure 5.4: Histograms characterizing the cluster sample obtained from the Millennium Simulation. Top left panel: redshift distribution of the clusters once we rejected those too close to the border of a survey patch. Top right panel: cluster histogram according to the number of hosted sub-halos. Bottom left panel: $M_{\rm FOF}$ mass histogram, this mass (the mass of the FOF group, see Sect. 3.6), combines the mass of the sub-halos and main halo mass. Bottom left panel: modulus of the ellipticity of the cluster's shape, according to the galaxy distribution.

In Fig. 5.4 we present some characteristics of the cluster sample. In a ray-tracing simulation, the redshift distribution of the lenses is discrete and one can see on the upper-left panel. The clusters are located in the different planes used for the ray-tracing. We consider 20 planes, but in the first three there are no clusters detected.

The mean sub-halo number (upper-right panel) is around 100. We also count the total number of sub-halos of our sample to be 291955. These are however, very different objects from large sub-halos with a large fraction of the mass of the cluster, to almost unresolved ones in the simulation.

In the lower-left panel we present the histogram of the FOF mass $M_{\rm FOF}$ of the cluster (do not confuse with the main halo $M_{\rm SUBF}$) This mass combines the mass of the sub-halos and main halo mass. As one can see the two lowest mass bins are not heavily populated. This can be explained by our selection criterion. These bins are populated with clusters with small mass but still with a main halo large enough to be inside our sample. These clusters have no or few sub-halos, and as shown in the upper-right panel they are rare.

In the lower-right panel, we present the histogram of the ellipticity modulus of the cluster's shape. The ellipticity was computed using Eq. (4.21), except that we use the galaxy spatial distribution instead of the luminosity function of the cluster. Using the luminosity function does not change the results qualitatively. Note that the clusters have some non-negligible ellipticity.

6. Sub-halos in the Millennium Simulation

In the previous chapter, we investigated how to proceed to apply galaxy-galaxy lensing on sub-halos in order to obtain their average mass profile. Our intention now is to test our method, and to characterize the lensing signal of the sub-halos in the Millennium Simulation. For this purpose we use the projected mass maps $\Sigma(\boldsymbol{\beta})$ defined in Sect. 4.5.

The value of the results that we present in the following lies in two considerations. In the first place, the predictions for galaxies we shall present later on, explicitly depend on the host sub-halos. With the results derived here, we shall be able to avoid most systematics and bias coming from a realistic shear catalogue. Secondly, these results should be representative of realistic lensing signals. The size and resolution of the Millennium Simulation in combination with semi-analytical galaxy catalogues makes our data set state-of-the-art.

Dark matter simulations with a higher resolution than the Millennium simulation have been conducted, i.e. the Aquarius Project (Springel et al. 2008) or Millennium II (Boylan-Kolchin et al. 2009). These are even focused on sub-halo studies. However, the Millennium Simulation is still better suited for our ray-tracing simulations and the masses we want to probe. There are also hydrodynamical simulations available (e.g. Baldi et al. 2010, Kereš et al. 2009) but despite their fast evolution they do not match the size and resolution of the previous ones.

In this chapter we address a number of questions. First we define our calibration tests and we present their results. Afterwards we characterize properly the mass profiles of the sub-halos in our sample. We discuss the detection of a truncation radius and which model describes better the sub-halos. Finally we study how the mass profiles of sub-halos evolve with time.

6.1 Synthetic clusters

In order to test our calibration scheme, we produce a sample of mock clusters. By using them, there is a unique true value for the parameters of the mass profile, and we can quantify how well our method can retrieve them. This mock sample is constructed using truncated NFW theoretical profiles. NFW profiles have been shown to be a good fit for relaxed dark matter halos at the resolution of the Millennium Simulation (Navarro et al. 2004).

Since we want to evaluate the method for situations as realistic as possible, we use the galaxy cluster sample that we have from the Millennium Simulation as a starting point (see Sect. 5.3.1). For this purpose, we use the projected mass maps. For each cluster in the sample we produce a mock one. The synthesis of the mock clusters is done in three

steps:

- characterizing the smooth overall component for each Millennium Simulation cluster by finding the best fitting profile;
- finding approximate relations between sub-halo mass and the parameters of a NFW profile, in order to draw a sub-halo density profile from its mass;
- produce synthetic clusters with the corresponding overall smooth component and the scaling relations derived from the sub-halos.

In order to speed up the computation, for the re-creation of the mock clusters we only consider small fields of view around their center. These fields of view are rectangular and their size is defined by the sub-halo number density distribution.

Synthetic main halo

The main halo is characterized in each case by finding the best fitting profile. We fit two types of profiles, an elliptical truncated NFW, and a truncated NFW profile with no parity symmetry, hereafter called non-elliptical. The elliptical profile is achieved by using elliptical coordinates. Non-elliptical profiles have non-zero higher order moments, presenting banana or triangular-like shapes.

The non-ellipticity is also extracted from the original Millennium cluster sample. In order to achieve it, we proceed as follows. First we fit an elliptical truncated NFW. Then we divide the original projected mass maps by the fitted profile. If the original map is denoted by $\Sigma(\boldsymbol{\beta})$ and the elliptical NFW by $\Sigma_{\rm ellip}(\boldsymbol{\beta})$, we obtain the coefficients

$$\begin{aligned} a_n &= \sum_p^N \frac{\Sigma(\boldsymbol{\beta}_p)}{\Sigma_{\rm ellip}(\boldsymbol{\beta}_p)} \cos\left(n\,\theta(\boldsymbol{\beta}_p)\right), \\ b_n &= \sum_p^N \frac{\Sigma(\boldsymbol{\beta}_p)}{\Sigma_{\rm ellip}(\boldsymbol{\beta}_p)} \sin\left(n\,\theta(\boldsymbol{\beta}_p)\right), \end{aligned} \tag{6.1}$$

where the sum runs over all N pixels $p(\boldsymbol{\beta})$ of the field of view, and $\theta(\boldsymbol{\beta})$ is the polar angle at each position $\boldsymbol{\beta}$. The profiles are always centered on the main halo position given by its potential minimum. We compute a Fourier Series from a discrete function up to order $n = 3$. The higher the coefficients are, the less robust is their computation due to pixelization. For our purposes $n \leq 3$ is sufficient. The final profile in the non-elliptical case $\Sigma_{\rm non-el}(\boldsymbol{\beta})$ is then

$$\Sigma_{\rm non-el}(\boldsymbol{\beta}) = \Sigma_{\rm ellip}(\boldsymbol{\beta}) \left(1 + \sum_{n=1}^{3} \left[a_n \cos\left(n\,\theta(\boldsymbol{\beta})\right) + b_n \sin\left(n\,\theta(\boldsymbol{\beta})\right)\right]\right). \tag{6.2}$$

An example of a non-elliptical truncated NFW profile can be found in Fig. 6.1.

Synthetic sub-halos

The sub-halo characterization is done by a preliminary application of our method. We classify the sub-halos according to $M_{\rm SUBF}$ (see Sect. 3.6). Then we measure the $\Delta\Sigma(\xi)$ as function of the projected physical radius ξ and we obtain the best fitting NFW profile in

each sub-halo class. Finally we find the best fitting power laws for the NFW parameters ($r_\mathrm{s}\,\rho_\mathrm{r}\,\delta_\mathrm{c}$, r_s), as a function of the sub-halo mass (M_SUBF).

We see no advantage in measuring $\Sigma(\xi)$ instead of $\Delta\Sigma(\xi)$, and we can reuse the code if we measure $\Delta\Sigma(\xi)$. In order to avoid biasing from an imprecise calibration we shall only use a limited range in ξ between 0.03 and 0.2 Mpc. These values were a sensible guess out of previous analyses. As we shall see, the final results support our choice. The full measurements are shown in Fig. 6.1, in the middle left panel. The modeling of the two parameters with respect to $\Delta\Sigma(\xi)$ are shown in the two bottom panels. The scaling relations between the parameters and the mass are

$$\begin{aligned}\frac{r_\mathrm{s}}{\mathrm{Mpc}} &= 1.1\times10^{-6}\,(M_\mathrm{SUBF}/\mathrm{M}_\odot)^{0.365},\\ \frac{r_\mathrm{s}\,\rho_\mathrm{r}\,\delta_\mathrm{c}}{(\mathrm{M}_\odot/\mathrm{pc}^2)} &= 1.92\times10^{10}\,(M_\mathrm{SUBF}/\mathrm{M}_\odot)^{0.2402}.\end{aligned} \tag{6.3}$$

Once we have characterized the parameters of the sub-halos, we can convert the mass into a profile. The mass of the NFW profile without truncation diverges if we integrate up to an infinite radius. We expect also a truncation of the profile due to tidal stripping (Hayashi et al. 2003). For these reasons we introduce a truncation radius, obtaining a truncated NFW profile. We apply this modification taking care that we do not alter the scaling relations found.

Table 6.1: Classification according to M_SUBF of all sub-halos in clusters for $z < 0.9055$. Note that the range is in $\mathrm{M}_\odot/\mathrm{h}$ as obtained from the database, and the mean SUBFIND mass $\overline{M_\mathrm{SUBF}}$ in $\mathrm{M}_\odot$. We also present the number of sub-halos in each bin.

Bin	**Range** [$\mathrm{M}_\odot$/h]	**# SH**	$\overline{M_\mathrm{SUBF}}$ [$\mathrm{M}_\odot$]
1	$[10^{11.33} : 10^{11.66})$	31831	4.27×10^{11}
2	$[10^{11.66} : 10^{12}\ )$	16025	9.14×10^{11}
3	$[10^{12}\ : 10^{12.33})$	8273	1.98×10^{12}
4	$[10^{12.33} : 10^{12.66})$	3834	4.23×10^{12}
5	$[10^{12.66} : 10^{13}\ )$	1816	9.15×10^{12}
6	$[10^{13}\ : 10^{13.33})$	810	1.96×10^{13}
7	$[10^{13.33} : 10^{13.66})$	385	4.27×10^{13}
8	$[10^{13.66} : 10^{14}\ )$	162	8.92×10^{13}
9	$[10^{14}\ : 10^{14.33})$	74	1.88×10^{14}

Final synthetic cluster

Each synthetic cluster is constructed by combining the previously defined main halo and sub-halo profiles. The profiles are drawn in pixel maps identical to the original projected mass maps. The center of the main halo and the sub-halos are also conserved. We present an example of how a synthetic cluster is produced in Fig. 6.1. In the top left panel, we show the original Millennium Simulation cluster (note that it can also be seen in the lower right part of the bottom panel in Fig. 5.3). In the top right panel we present the non-elliptical main halo that best fits the original. Finally, in the middle right panel, one can see the final total cluster with the presence of the sub-halos. The reconstruction loses information about the surroundings of the cluster. However, we assume that the

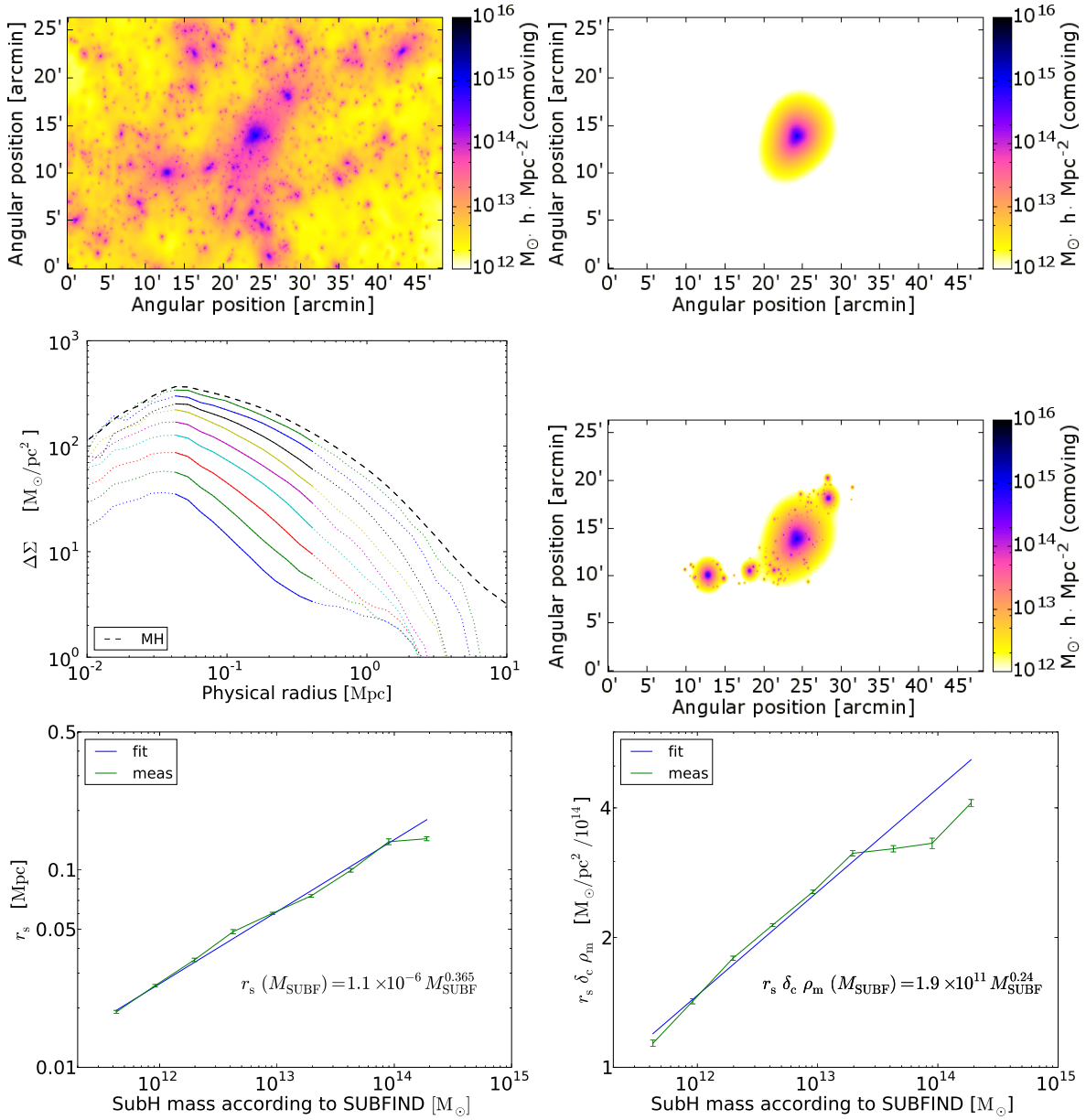

Figure 6.1: Top left: a cluster from the projected mass maps. Top right: best fitting non-elliptical truncated NFW to the main halo of the previous cluster. Middle right: final synthetic cluster produced from the one depicted in the top left panel. Middle left: preliminary measurement of $\Delta\Sigma$ profiles for different sub-halo mass $M_{\rm SUBF}$ bins (see Tab. 6.1) with dotted lines; we highlight the range used for the characterization with solid lines. We show the average signal for the host cluster with a dashed line as a visual reference. Bottom left: NFW scale radius ($r_{\rm s}$) as a function of $M_{\rm SUBF}$, from the measurements depicted in the middle left panel, and the best fitting power law. Bottom right: amplitude of the NFW profile ($r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}$) as a function of $M_{\rm SUBF}$, from the measurements depicted in the middle left panel, and the best fitting power law.

surroundings are uncorrelated with the cluster. Therefore, the average surrounding mass is expected to be a constant mass distribution which cannot produce shear.

The reconstruction fails in the few cases (below 10%), where the main halo is not round enough (bimodality, unrelaxed main halo ...). Nevertheless we are confident that the way we proceed is consistent and robust. In this procedure we take into account:

- the spatial distribution of the sub-halos inside the cluster,
- the sub-halo mass function,
- the dependence between the mass of the sub-halo and its position in the cluster,
- and the correlation between the main halo shape and the position of the sub-halos.

With this procedure we can measure how well our calibration works in the presence of a non-elliptical main halo, or an elliptical one. We can also isolate the impact of the surrounding sub-halos in our measurement by plotting the sub-halos only. In section 6.2 we present the results for all these possibilities.

6.2 Results for the calibration tests

In the previous section, we defined a procedure to produce synthetic clusters with which to test our method. Using them, we present the following three tests. First we derive the radial range where we can measure an unbiased $\Delta\Sigma(\xi)$ profile. Second, we define a criterion to filter the sub-halo sample in order to optimize the results. And third, we discuss the problem of the misidentification of the main halo center. For our calibration tests we divide our sub-halo sample again according to $M_{\rm SUBF}$. In Tab. 6.1 we present the ranges, the mean mass and the number of members for each bin. We have excluded from the sample the lightest bins of sub-halos. Those are the ones with the smallest number of particles and therefore the least resolved.

6.2.1 Range for measuring $\Delta\Sigma$

As mentioned in Sect. 5.2, we can only measure properly $\Delta\Sigma(\xi)$ within a range in radius ξ. Here we derive this range using the synthetic cluster sample.

In Fig. 6.2 we present the results of the test for a selection of the nine sub-halo mass bins. We plot the ratio of the measured $\Delta\Sigma(\xi)$ to the input profile versus physical radius ξ. In the top panel we used a spherical NFW profile for the sub-halos with a truncation radius $r_{\rm tr} = 6.66\, r_{\rm s}$. In the bottom panel we used a different truncation radius, $r_{\rm tr} = 2\, r_{\rm s}$. The first value yields for the least massive sub-halos, with a scale radius of 0.03 Mpc, a truncation radius of 0.2 Mpc. As we shall show, this truncation is above the largest radius for which we measured an unbiased $\Delta\Sigma(\xi)$, for these sub-halos. With it we consider the case where the truncation is difficult to detect. The second value was an sensible choice which yields a much smaller truncation radius and therefore easier to detect.

For each mass bin in each panel, we present the results from different situations. With black dotted-dashed lines, we present the $\Delta\Sigma(\xi)$ signals coming only from the sub-halos, i.e. the main halo was not added. With red dashed lines, we present the results for an elliptical main halo plus the sub-halo population. Finally with solid blue lines, we present the signals from realistic clusters using the best non-elliptical main halos. The shaded region encompasses the one sigma error.

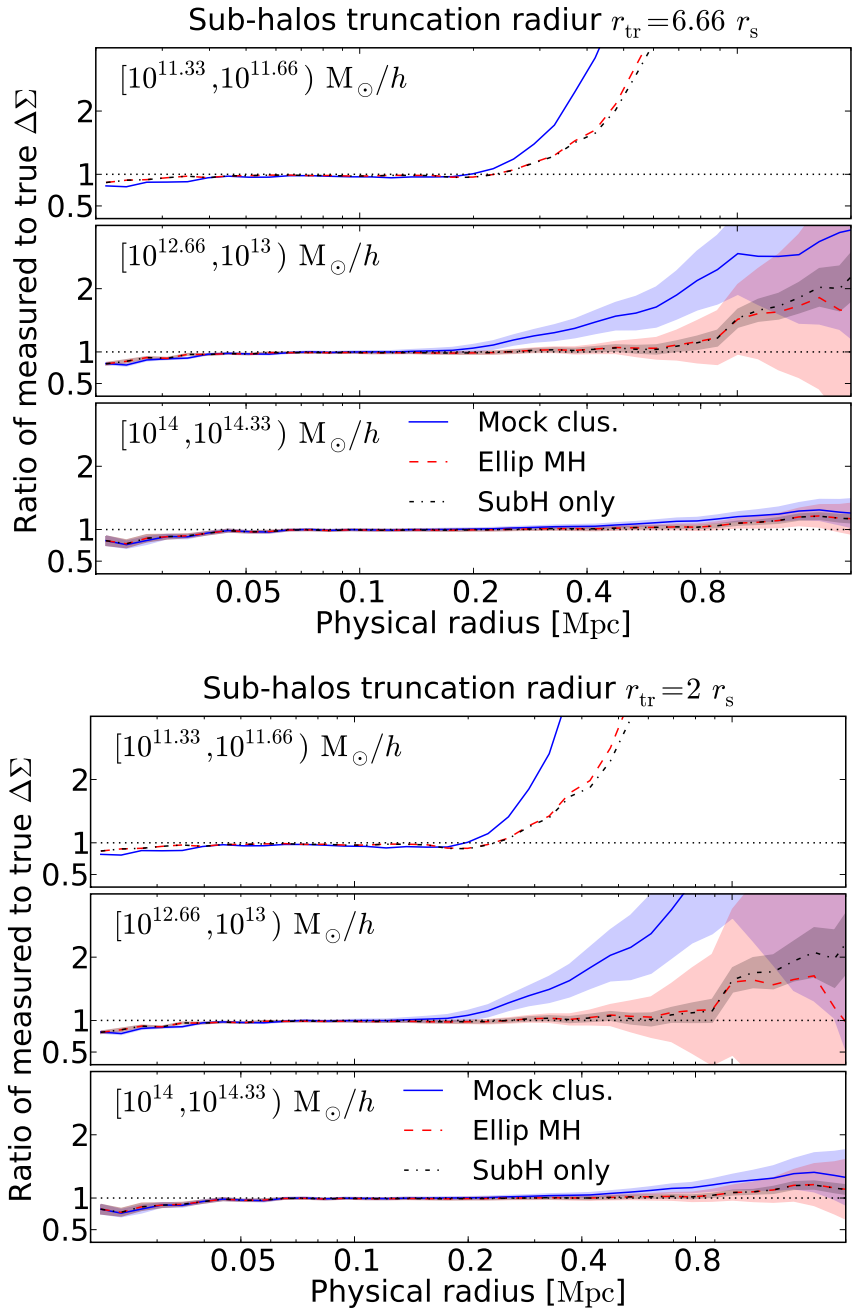

Figure 6.2: Analysis with synthetic clusters, ratio of the measured $\Delta\Sigma(\xi)$ to the one we input, versus the projected physical radius ξ. Top figure: the truncation radius is $r_{\rm tr} = 6.66\, r_{\rm s}$. Bottom figure: $r_{\rm tr} = 2\, r_{\rm s}$. We plot only the results for a selection of mass bins; the mass ranges are written in each plot. Black dotted-dashed lines: measurements done with mock clusters where the main halo was removed. Red dashed lines: clusters with an elliptical main halo. Blue solid lines: clusters with a non-elliptical main halo. The shaded regions correspond to the one sigma error. The dotted black line, where the ratio is one, is a visual reference.

We assume valid for analysis the range of radii for which the measured value deviates less than a 10% from the input model. We will only be able to detect a truncation radius if this scale is well inside the range we can measure. The resolution of our data can also be seen from this analysis. The pixelization of the data limits the resolution of the density profile. The consequence is that the measured signal is smaller than the input one for scales smaller than 0.06 Mpc. We consider that the resolution is good enough to use scales up to 0.03 Mpc. This resolution limit comes from the ray-tracing simulations used, therefore we use it in all our results.

The distribution of the sub-halos is not parity-symmetric with respect to the cluster center. Since the positions of the sub-halos are correlated among themselves, when we measure around a particular sub-halo, we "see" the surrounding ones. If this were not the case, the contribution to our measurement would tend to be zero on average and the black dotted-dashed lines would show no bias. So far we have considered only the contamination of the main halo to the measurement of $\Delta\Sigma(\xi)$. However as it is shown here, there is a non-negligible contamination coming from the population of sub-halos surrounding the one we measure.

A completely elliptical halo can always be calibrated properly. The observed effect is an increased variance that masks the bias introduced by the surrounding sub-halos. The one sigma region clearly encloses the signal computed only with sub-halos (red regions around dashed lines). For a non-elliptical halo, the bias is larger as we are not able to calibrate all of the main halo's contribution (solid blue).

Our method is unable to distinguish between the contamination due to the main halo and due to the sub-halo's spatial distribution. In fact they can be correlated. In any case the sub-halo lensing signal is not only affected by their host main halo but by the other sub-halos.

The difference between the two truncation models does not substantially change the region that we eventually define as suitable to derive the mass profile. Since, as we show later on, our simulations do not present a clear truncation of the mass density profile, we focus in the following in the truncation model where $r_{\rm tr} = 6.66\, r_{\rm s}$.

6.2.2 Discarding sub-halos

The main halo's contribution to the bias in $\Delta\Sigma(\xi)$ is more important for sub-halos near the cluster center. In fact, for certain projected separations, including the sub-halos may decrease the quality of the measured $\Delta\Sigma(\xi)$. Having this in mind, we consider different selection criteria concerning the physical projected separation between the main halo center and the sub-halo center. Note that due to projection effects, on average the true three-dimensional separation between the sub-halo and the main halo is larger than the measured one.

The results are shown in Fig. 6.3. Here we present again the ratio of the measured $\Delta\Sigma(\xi)$ to the input value versus physical radius ξ. In order to be concise we only present some mass bins, and only for the case where sub-halos have a truncation radius $r_{\rm tr} = 6.66\, r_{\rm s}$. This time we compare different selections of sub-halos. In each panel we plot the signal coming from the whole sample in dashed blue; in solid red, the signal where we discard all sub-halos closer than 0.5 Mpc to the main halo center; and finally in black dotted-dashed the signal where we discard all sub-halos closer than 1 Mpc to the main halo center. The shaded regions correspond again to the one sigma error regions.

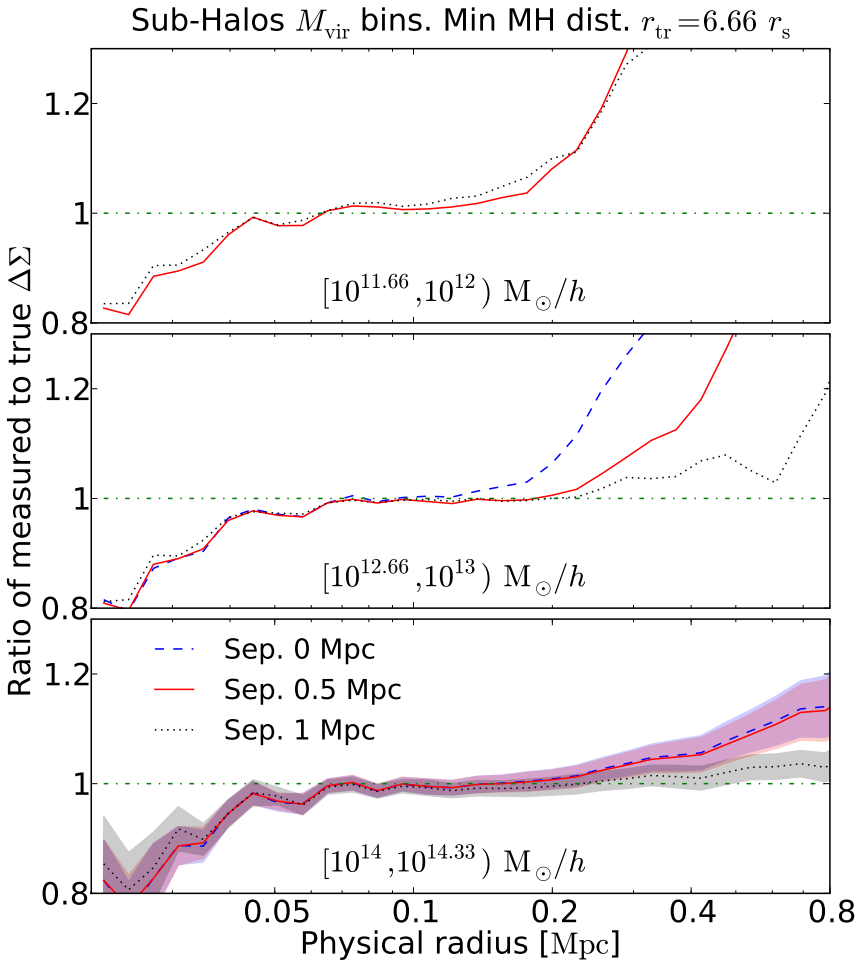

Figure 6.3: Analysis with synthetic clusters, ratio of the measured $\Delta\Sigma(\xi)$ to the one we input, versus the projected physical radius ξ. Only the results for a selection of mass bins, for the case where $r_{\rm tr} = 6.66\, r_{\rm s}$ are presented; the mass ranges are written in each plot. Blue dashed lines: results for the whole sub-halo sample. Solid red lines: results for the sub-halos with a minimal distance of 0.5 Mpc to the cluster center. Black dotted lines: results for the sub-halos with a minimal distance of 1 Mpc to the cluster center. The shaded regions correspond to the one sigma error. Note that, for clarity, we do not plot the dashed blue line in the top panel, its bias and its variance are nevertheless larger than in the other two cases. Note that in the bottom panel the red solid line and the blue dashed one overlap.

The signal improves if we discard sub-halo closer than 0.5 Mpc. Eliminating these sub-halos decreases both bias and variance. The heaviest sub-halos are more likely to be recent mergers and lie far away from the cluster center so the criterion has little effect in this case (lowest panel).

Increasing the minimal separation between cluster center and main halo however, depletes the sample and limits our result to sub-halos in the outskirts of the cluster. Imposing a minimum distance of 1 Mpc increases the range where the bias is below 10% but only at the expense of limiting our sample and increasing the variance (compare the variance for the black dotted line and for the red solid line at the top panel in Fig. 6.3). For the heaviest sub-halos, a larger minimal separation improves the results, and seems to be a better choice. Nevertheless, for simplicity we shall discard sub-halos closer than 0.5 Mpc in projection, regardless of sub-halo type.

6.2.3 Main halo center detection

In the Millennium Simulation, we know where the main halo center is, namely in the potential minimum. Moreover, within the semi-analytical catalogues there is always a galaxy placed at the center of the main halo by construction. The definition of the calibration center is straightforward. However, in real clusters, the correlation between galaxy and main halo center is supposed to be less tight. The definition of the calibration point is henceforth more difficult.

The study that we propose in this thesis is directed to large samples of clusters in lensing catalogues. Nonetheless, it could use additional sources of data, like X-rays surveys, or parallel lensing investigations to determine the center of the clusters. Therefore it is not unlikely that a well determined main halo center can be used. For the rest of the thesis we shall assume the center of the main halo can be found accurately. However, for the sake of completeness we want to consider the possibility of not knowing the main halo center.

The expected spatial correlation between the galaxy hosted by the main halo and its center is difficult to determine. In order to quantify the effect of not knowing the main halo center, we compute the calibration assuming a different cluster center. In Fig. 6.4 we present again, using the synthetic cluster sample, the ratio between the measured $\Delta\Sigma(\xi)$ and the input one, for a selection of sub-halo mass bins. The shaded regions are the one sigma errors.

With blue dashed lines, we plotted the signal previously presented, namely, that obtained as we measure the calibration around a point defined with the help of the main halo center. Then, with dotted black lines, we present the signal where the calibration point is chosen differently. Instead of using the main halo to define the calibration point, we took the average position of all halos in the cluster (main halo and sub-halos). If we assume that the displacements between galaxies and host halos are uncorrelated within the cluster, we only need to detect the hosted galaxies to obtain this position. The center of the cluster defined in this way is independent of the position of the central galaxy with respect to the main halo.

For clusters formed by two large mass overdensities colliding the separation between the two different centers can be up to ~ 2 Mpc. The distribution of separations is highly asymmetric, with a mean of 0.215 ± 0.004 and a median of 0.137.

The new way to define the cluster center results in a much less smooth signal. The

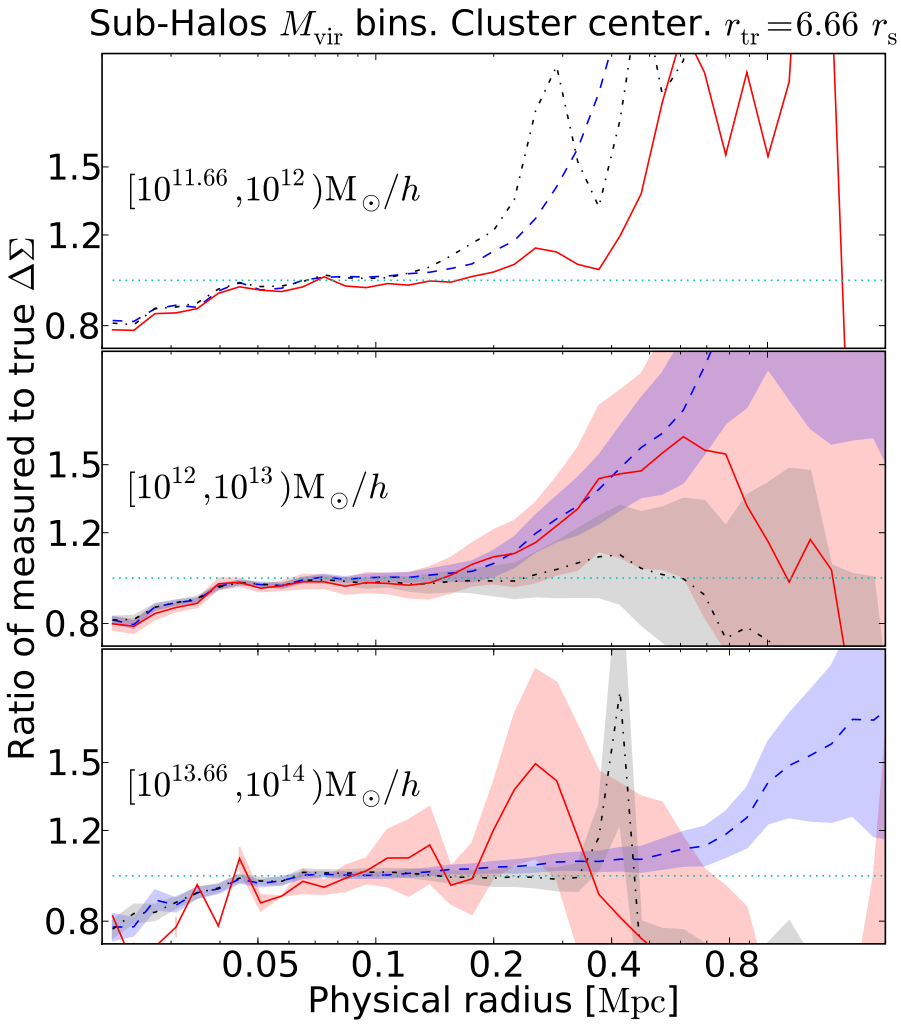

Figure 6.4: Analysis with synthetic clusters, ratio of the measured $\Delta\Sigma(\xi)$ to the one we input, versus the projected physical radius ξ. Only the results for a selection of mass bins, for the case where $r_{\rm tr} = 6.66\, r_{\rm s}$ are presented; the mass ranges are written in each plot. Dashed blue lines: calibration point defined using the main halo center. Black dotted-dashed lines: calibration point defined using the average halo position. Solid red lines: calibration point defined using the average halo position, the clusters were selected such that the separation between BCG and the average halo position was less than 10 kpc.

incorrect calibration of the main halo in some extreme cases appears as a sharp peak in the measurement as seen in the top and in the bottom panel. This identification of the cluster center seems to produce a better calibration for medium mass halos (middle panel). For non-virialized clusters, the new cluster center is closer to the main halo as the sub-halo on average. The main halo contribution measured around the calibration point is larger that the one around the sub-halo, and the final effect is an apparent smaller bias. Nevertheless, the main halo is still not properly calibrated. Choosing the main halo as the cluster center is therefore a more robust method which works consistently on all scales.

Finally, we present a possible improvement to the calibration defined above. We can always discard the clusters which are particularly problematic. In our case we considered the distance between the average halo position and the position of the central galaxy, as filtering criterion. The same criterion can be applied on real observations. We only need to assume that the separations between BCG and the main halo center, and between average galaxy position and main halo center, are uncorrelated. This shall be the case for non-virialized objects, which are responsible for the problem. We plot the result with a red continuous line. We filter out clusters where main halo and the new cluster center are more than 10 kpc away. The signal obtained with the alternative method comes to agree more with the original one once the problematic clusters are discarded. This comes with a price as the variance increases, in some cases like in the bottom panel, compromising the detectability.

We also computed the signal using the luminosity distribution to determine the cluster center. The results reflected the same behavior, and are therefore not shown.

6.3 Sub-halo profiles for different mass bins

Finally, we shall present the analysis of the different sub-halos from the projected mass maps $\Sigma(\boldsymbol{\beta})$ coming from the Millennium Simulation. The mass binning and the number of sub-halos of the samples used are described in Tab. 6.2. In this analysis we fix a minimal distance between the sub-halo and the main halo of 0.5 Mpc as explained in Sect. 6.2.2. The purpose is to find the best possible model and a scaling relation between parameters and mass.

As explained in Sect. 6.2.1, out of all physical radii ξ we can measure, only a limited range can be used. Using the results from Sect. 6.2.1, we present in Fig. 6.5 the data we can use safely. The upper and lower limits are given by the last point which has a bias smaller than 10%. The lower limit is a consequence of the resolution of our simulations. The range grows with the mass bin. First, because a larger sub-halo has a $\Delta\Sigma(\xi)$ signal with a higher amplitude and the contaminations from the main halo and the surrounding sub-halos are relatively less important. Second, because larger sub-halos are on average further away from the center.

In order to characterize the mass profiles we try to find the best fitting model. We fit a NFW profile, a truncated NFW profile and two versions of a PIEMD profile to the signals (see the Appendix for a full description of the profiles). The first version of PIEMD corresponds to the choices made in afore-mentioned works by Limousin et al. (2007) and Natarajan et al. (2007). There, the so-called core radius is fixed a priori to 10^{-4} Mpc and only the cut radius and the central density are fitted. The second version corresponds to fitting all three parameters.

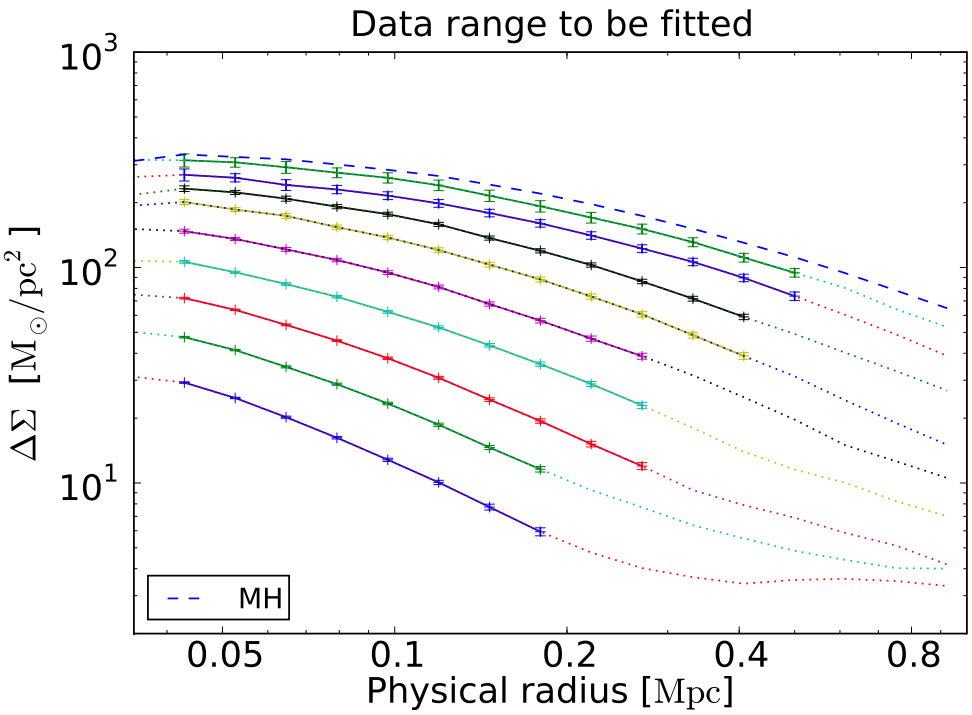

Figure 6.5: The $\Delta\Sigma(\xi)$ profiles measured from the projected mass maps $\Sigma(\boldsymbol{\beta})$, for the different mass bins. We highlight the final range used for the analysis and we plot the corresponding error bars. The range in ξ was defined through the results in Sect. 6.2.1. In dashed blue we show the average signal for the host cluster as a visual reference.

We conducted a Bayesian analysis to get the best fitting parameters for each model in each mass bin. We assumed a Gaussian likelihood in each case. The corresponding covariance matrices were computed from the data using a bootstrap algorithm. As specified in Sect. 5.3 we use weighted means, and so bootstrapping was the most convenient way to compute them. We construct each bootstrap realization by randomly drawing N sub-halos with replacements. We do so for each mass bin, producing 10000 realizations in each case.

In order to rank the different models, we consider the Bayesian evidence. The computation of the evidence was done using a nested sampling algorithm (Skilling 2004). We describe this algorithm and the Bayesian evidence in the Appendix. The results are in Tab. 6.3. The errors presented are the standard error from 10 realizations of the measurement except for the leftmost model where we use a raw estimation coming from the algorithm. The analysis in this case did not require a better accuracy. We write in bold the maximum evidence in each case, which belongs to the best model.

The first thing to underline is that the PIEMD model chosen in the afore-mentioned literature does not fit well sub-halos in the Millennium Simulation. In fact, the model fails to reproduce the simulations. This is important if we assume that the simulation describes well the scales and masses we treat. While our result does not discredit their results, it nevertheless challenges their choices.

The best model, except in two occasions, is a simple NFW profile. Note that the second best model is still the NFW profile. We interpret the two exceptions to the trend as statistical fluctuations. The large significance of the results can be explained by an underestimation of the errors. We compute our covariance matrices with a bootstrapping algorithm. Although bootstrapping is the best method given the characteristics of our

Table 6.2: Classification of all sub-halos in clusters for $z < 0.9055$ according to $M_{\rm SUBF}$. The sub-halos are at least 0.5 Mpc away from the main halo center. This last criterion changes slightly the samples with respect to those in Tab. 6.1. Note that the range is in $M_\odot/h$ as obtained from the database, and the mean SUBFIND mass $\overline{M_{\rm SUBF}}$ in $M_\odot$. We also present the number of sub-halos in each bin.

Bin	**Range** $[M_\odot/h]$	**# SH**	$\overline{M_{\rm SUBF}}$ $[M_\odot]$
1	$[10^{11.33} : 10^{11.66})$	23229	4.27×10^{11}
2	$[10^{11.66} : 10^{12}\ )$	11821	9.14×10^{11}
3	$[10^{12}\ : 10^{12.33})$	6151	1.98×10^{12}
4	$[10^{12.33} : 10^{12.66})$	2837	4.25×10^{12}
5	$[10^{12.66} : 10^{13}\ )$	1476	9.22×10^{12}
6	$[10^{13}\ : 10^{13.33})$	728	1.97×10^{13}
7	$[10^{13.33} : 10^{13.66})$	363	4.30×10^{13}
8	$[10^{13.66} : 10^{14}\ )$	155	8.79×10^{13}
9	$[10^{14}\ : 10^{14.33})$	73	1.89×10^{14}

Table 6.3: Comparison of the logarithm of the evidences for the four models fitted to the $\Delta\Sigma$ signals. In bold we emphasize the best model in each mass bin.

bin	**NFW**	**NFW tr**	**PIEMD 3 par.**	**PIEMD 2 par.**
1	-89.71 ± 0.05	$\mathbf{-89.57\pm0.08}$	-92.89 ± 0.3	-157 ± 1
2	$\mathbf{-88.31\pm0.09}$	-88.54 ± 0.09	-93.38 ± 0.14	-285 ± 1
3	$\mathbf{-88.38\pm0.09}$	-89.38 ± 0.12	-95.99 ± 0.13	-472 ± 1
4	$\mathbf{-85.90\pm0.07}$	-87.72 ± 0.13	-86.33 ± 0.15	-710 ± 3
5	-98.32 ± 0.07	-101.68 ± 0.12	$\mathbf{-96.43\pm0.06}$	-876 ± 3
6	$\mathbf{-93.38\pm0.05}$	-97.04 ± 0.17	-101.02 ± 0.09	-576 ± 1
7	$\mathbf{-99.20\pm0.06}$	-102.63 ± 0.21	-105.43 ± 0.06	-809 ± 2
8	$\mathbf{-88.09\pm0.06}$	-92.80 ± 0.18	-126.59 ± 0.03	-528 ± 1
9	$\mathbf{-85.62\pm0.04}$	-89.42 ± 0.14	-127.03 ± 0.07	-445 ± 1

measurements, we have no control over the precision of the covariance matrices.

With the data used it is not possible to detect a truncation radius. The truncation is beyond the region where our measurement is unbiased. This means that the truncation, if present at all, happens at least at scales larger than 0.2 Mpc, which is roughly four times larger than the results in the literature mentioned. In the case of the PIEMD with no fixed core radius the values for the truncation radius are in some cases consistent with the published values. Nonetheless, this model is strongly disfavored and it is not the one used in the literature.

From the nested sampling algorithm one can also obtain the posterior. We take the mean according to the computed posterior as the best value and the standard deviation as one sigma error. The best values for the NFW profile for each bin are presented in Fig. 6.6. We plot two sets of parameters as function of $M_{\rm SUBF}$. In the top panels we plot the scale radius, and the density parameter which is twice the density at the scale radius. In the bottom panels we plot the concentration and r_{200}. The concentration is the ratio between r_{200} and $r_{\rm s}$. The parameter r_{200} is defined as the radius at which the density of the best fitting NFW profile equals 200 times the mean comoving density of the Universe[1]. Its definition is thought to enclose the region where the virial mass of the halo is found. However r_{200} has no physical meaning in our case, as the sub-halo cannot

[1]The definition of r_{200} is not unique. Other authors use slightly different formulations.

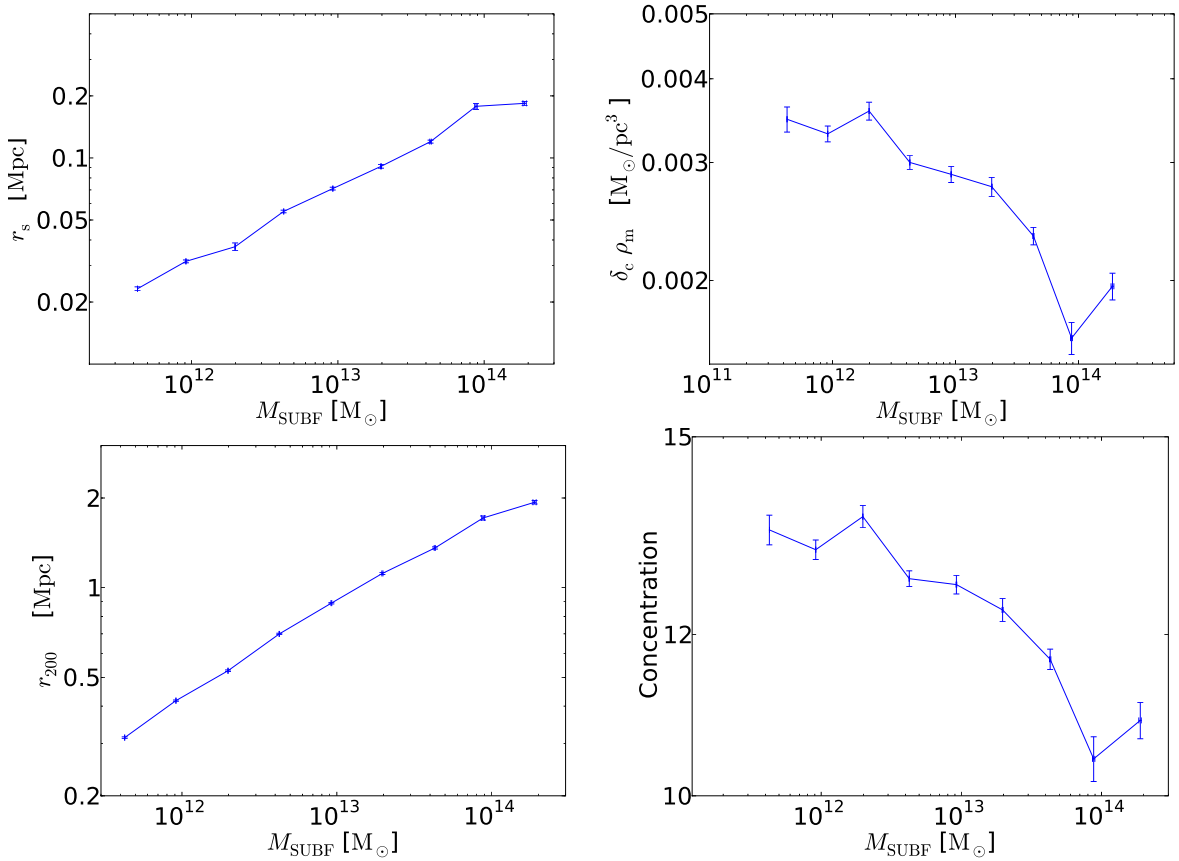

Figure 6.6: NFW parameters fitted to the sub-halo profiles in Fig. 6.5 as a function of $M_{\rm SUBF}$. We show in the top panels the results for $r_{\rm s}$ and $\delta_{\rm c}\,\rho_{\rm m}$, and in the bottom panels for r_{200} and concentration.

extend as far as r_{200}. Nevertheless, Fig. 6.6 shows that this parameterization is a good observable for sub-halo mass.

The correlation of r_{200} and $r_{\rm s}$ with $M_{\rm SUBF}$ allows us to use either one as a proxy for the mass. We shall see that the estimation from r_{200} however is much more robust. We discuss this further in Sect. 7.3 and give supporting arguments.

The best fitting power laws for the left panels in Fig. 6.6 are:

$$M_{\rm SUBF}\,(r_{\rm s})\,/{\rm M}_\odot = 10^B \times (r_{\rm s}/{\rm Mpc})^A \quad \text{with} \quad A = 2.88 \pm 0.07, \quad B = 16.29 \pm 0.13$$
$$M_{\rm SUBF}\,(r_{200})\,/{\rm M}_\odot = 10^B \times (r_{200}/{\rm Mpc})^A \quad \text{with} \quad A = 3.18 \pm 0.05, \quad B = 13.17 \pm 0.06$$

By looking at the left panels Fig. 6.6, one can see that the errors are so small, that the deviation from a simple power law fit should be significant. As we already mentioned, we suspect that the way we computed our covariance matrices makes us to underestimate the errors. Therefore, we see a clear scaling relation between the sub-halo mass $M_{\rm SUBF}$, and r_{200} or $r_{\rm s}$. Nevertheless, in Sect. 7.3 we shall see that the determination of the sub-halo mass has further problems.

Regarding the Millennium Simulation, a truncation radius is not detected. This leads us to conclude that the sub-halos in the Millennium Simulation have a large truncation radius, above the range in radius ξ that we measure.

As yet another check to our results, we compare the mass we can infer from our measurements and $M_{\rm SUBF}$. We compute the mass inside the range in radius ξ where we consider that $\Delta\Sigma$ is unbiased, assuming an NFW profile. In the first column of Tab. 6.4 we present in percentage, the ratio between the measured mass and $M_{\rm SUBF}$. The amount

of mass that we are missing is not large except for the heaviest halos (last row of the table). Therefore we can assume that with our analysis we are taking into account most of the sub-halo spatial extent, except for that case. We interpret this as another indication of a rather sharp truncation for the sub-halos at large radii.

On real lensing surveys we cannot conduct the calibration tests described in Sect. 6.2.1. However, as seen before, the range in ξ where the measured $\Delta\Sigma(\xi)$ is unbiased depends on the mass bin. We propose the following method: one measures $r_{\rm s}$ or r_{200} with a preliminary radius range, i.e. up to ~ 1 Mpc; then from Tab. 6.4, we interpolate the corresponding range and we measure again r_{200}. We iterate the process until we our result converges. The estimation of the range in ξ is in this way coarse but effective. For all the results in the following we proceed as we described here.

Table 6.4: Classification of all sub-halos according to $M_{\rm SUBF}$ in clusters for $z < 0.9055$. The sub-halos are at least 0.5 Mpc away from the main halo center. The mass binning can be found in Tab. 6.2. Here we show the best fitting values for $r_{\rm s}$, r_{200}, the corresponding range in radius where our profiles are not unbiased and the ratio between inferred mass and $M_{\rm SUBF}$.

bin	% $\frac{\text{infered mass}}{M_{\rm SUBF}}$	$r_{\rm s}$ [Mpc]	r_{200} [Mpc]	**Unbiased range** [Mpc]
1	113±16	0.0232±0.0005	0.3147±0.0018	[0.043:0.18]
2	109±10	0.0315±0.0005	0.417±0.002	[0.043:0.18]
3	81±14	0.0371±0.0016	0.526±0.003	[0.043:0.18]
4	119±8	0.0550±0.0008	0.699±0.004	[0.043:0.27]
5	95±7	0.0710±0.0011	0.887±0.005	[0.043:0.27]
6	108±9	0.0910±0.0018	1.116±0.009	[0.043:0.41]
7	78±6	0.120±0.002	1.359±0.011	[0.043:0.41]
8	76±10	0.178±0.006	1.71±0.03	[0.043:0.5]
9	45±4	0.184±0.004	1.94±0.03	[0.043:0.5]

6.4 Evolution of sub-halos

In the previous section we analyzed the average mass profile for sub-halos of different masses from the point of view of a lensing study. We analyzed the detectability of a truncation in the sub-halo's profile as an indication of tidal stripping. Our conclusion is that the values published in the literature are not a good description for the Millennium Simulation. Nevertheless, we have not yet presented any description about how the time spent inside the cluster alters the mass profiles of the sub-halos. We address this matter here.

In order to observe how sub-halos evolve inside the clusters we considered two new parameters: the infall mass $M_{\rm inf}$ and infall snapshot both computed by L. Wang (Wang et al. 2006). The first describes what the halo mass was at the last epoch when it was a central dominant object, i.e. the mass just before it was accreted by a larger structure. The second is the snapshot at which this accretion was detected and allows us to compute the time spent inside the cluster, or the *age* of the sub-halo.

We divide our sample into three classes according to $M_{\rm inf}$, and each of these classes into three sub-classes according to their age. In order to define the classification, we maximized the signal-to-noise defined as the ratio, between the measured $\Delta\Sigma(\xi)$ and its sample variance averaged over ξ. We also avoided too small sub-halo samples. Another

Table 6.5: $M_{\rm inf}$-age classification of all sub-halos in clusters for $z < 0.9055$. The sub-halos are at least 0.5 Mpc away from the main halo center. Here we present the parameters r_{200} and $r_{\rm s}$ as estimators of $\overline{M_{\rm SUBF}}$. 2nd column: number of sub-halos in each sub-sample. 3rd column: infall mass ($[10^{10}\,{\rm M}_\odot/h]$ from the database) ranges $L \to [20:60]$, $M \to [60:150]$, $H \to [150:\infty]$. 4th column: age ranges (Gigayears) I $\to [0:3]$, II $\to [3:4.5]$, III $\to [4.5:\infty]$. The mean separation between the main halo center and the sub-halo center (6th column) is in Mpc. The last two columns (mean SUBFIND mass, and mean infall mass) are in units of $10^{10}\,{\rm M}_\odot$.

bin	# SH	$M_{\rm inf}$	**age**	$\overline{age}$	$\overline{Sep.}$	$r_{\rm s}$ [Mpc]	r_{200} [Mpc]	$\overline{M_{\rm SUBF}}$	$\overline{M_{\rm inf}}$
×–	31408	L	I	1.797	1.149	0.0252±0.0006	0.274±0.018	27.68	46.21
+–	14927	L	II	3.723	1.132	0.0250±0.0011	0.25±0.03	16.75	47.11
•–	10482	L	III	5.608	1.175	0.0240±0.0009	0.21±0.05	22.9*	47.41
×- -	11487	M	I	1.852	1.169	0.0340±0.0007	0.391±0.018	73.8	127.8
+- -	6391	M	II	3.720	1.136	0.0276±0.0007	0.34±0.04	42.9	127.9
•- -	4410	M	III	5.629	1.143	0.0216±0.0010	0.29±0.05	29.0	126.9
× · ·	10792	H	I	1.798	1.260	0.0671±0.0006	0.75±0.04	810	1280
+ · ·	4782	H	II	3.688	1.151	0.0473±0.0010	0.58±0.05	210	682
• · ·	2331	H	III	5.534	1.150	0.0319±0.0010	0.46±0.05	104	488

*The value changes to 12.36 after eliminating two misidentified sub-halos.

issue to consider was the limited time resolution of the snapshots to compute the age of the sub-halos. The ranges for classes and sub-classes are in Tab. 6.5 and the resulting measurements are plotted in Fig. 6.7. The color and the line type in the plot relates to the infall mass range; the symbol of the line to the time spent inside the cluster, as indicated in the first column of Tab. 6.5.

The first thing to consider is that the cuts performed do not always define homogeneous samples. For instance, the mean $M_{\rm inf}$ for the heaviest sub-halos (last three rows, tenth column in Tab. 6.5), are not similar. The three sub-halo classes did not have a similar mass right before they merged into the cluster. The evolution of the different sub-samples appears to have been different. Therefore, for these cases our classification is not optimal. This is an intrinsic problem that will become worse when using observables instead of direct quantities. Nevertheless we detect that the mass decreases as time evolves in all cases by comparing the ratio $M_{\rm inf}/M_{\rm SUBF}$ for each age sub-classes within this infall mass range.

The oldest sub-class for the least massive sub-halos (third row, ninth column in Tab. 6.5), does not follow the trend where older sub-halos have smaller $M_{\rm SUBF}$. Nevertheless, we find that the last value decreases from 22.9×10^{10} to 12.36×10^{10} if we eliminate two detected outliers. In these cases, a visual inspection of the clusters revealed that the identity of a large sub-halo and a close-by merging one must have been swapped. The spatial proximity and a likely matter exchange can produce the misidentifications.

The profiles do not get truncated. Instead they show a change at all radii. This change is consistent with a mass loss at all scales. Therefore, the sub-halo is stripped in the outermost layers while the innermost mass is reheated and redistributed. This is in agreement with the results by Hayashi et al. (2003). The truncation is always at a larger radius ξ than the range that we can safely probe. The measured ranges for the sub-halo radius were taken using Tab. 6.4 as explained at the end of Sect. 6.3.

Although the plots for sub-halos of the same $M_{\rm inf}$ but with different ages show a clear variation, it is not clear how the difference can be quantified. In Tab. 6.5 we present the best fitting values for $r_{\rm s}$ and r_{200}. The values decrease as the age increases for each $M_{\rm inf}$

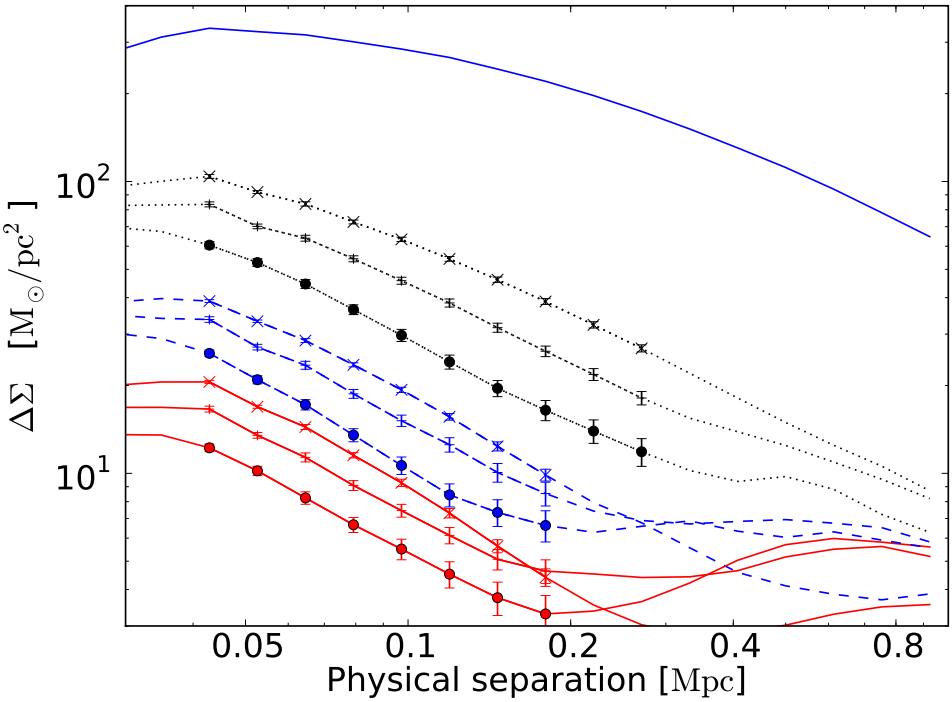

Figure 6.7: Profiles for different infall mass and time spent inside the cluster: with error bars the regions defined as unbiased. The color and the line type denote infall mass range. The symbols denote age range. See Tab. 6.5 for details. In solid blue we plot the average signal for the host cluster as a visual reference.

bin. Both parameters show mass loss but the errors for the medium and light sub-halos are quite large (first six rows, seventh and eighth column).

The quantification of the mass loss with a single parameter like r_s or r_{200} has large errors. Therefore it may not be the best approach. A detailed matter evolution model for sub-halos could improve the significance of the measurements.

7. Galaxies in clusters in the Millennium Simulation

In previous chapters we defined our method to compute galaxy-galaxy lensing on sub-halos. We also characterized the simulated data which we use to study its application. In this chapter, we forecast the signals that one could measure with our method in real lensing surveys.

The method that we proposed requires a large sample of galaxy clusters. Therefore large optical imaging surveys are needed. The goal is to use data coming from current surveys like SDSS (Abazajian et al. 2009), ongoing like DES (The Dark Energy Survey Collaboration 2005), KIDS[1] or Pan-STARRS (Kaiser et al. 2010), and future ones like LSST (Ivezic et al. 2008) or Euclid (Refregier et al. 2010). We should also benefit from Sunyaev-Zel'dovich effect surveys (e.g. SPT, Carlstrom et al. 2009) or X-rays surveys (e.g. eRosita, Predehl et al. 2010) to detect clusters. These surveys should also proportionate cluster characteristics such as shape, or help to locate the density maxima, which would improve the results of our lensing analysis.

For our method it is important to properly classify the sub-halos in the sample. On real data, the classification of the sub-halos must be done through the characteristics of the hosted galaxies. The properties of the galaxies in our case are given by the semi-analytic galaxy models by De Lucia & Blaizot (2007). It is therefore important to discuss which are the most useful galaxy observables and what is to be expected from them.

Our previous results were computed with the projected mass maps $\Sigma(\boldsymbol{\beta})$, but never with a true shear catalogue. We shall therefore describe the shear catalogues from which we compute the measurements. Once we have presented our lensing data set, we address the detectability of the signals in different surveys. We also discuss the best way to estimate the sub-halo mass. Then we check the impact of using the weak lensing approximation on sub-halo profiles. This approximation is based on the condition $\kappa \ll 1$ (see Sect. 4.4), which is often not fulfilled inside clusters.

Finally, we investigate what can be learned by using our method considering surveys like LSST and DES. First, we consider how well we can estimate a mass-luminosity relation from our measurements. Then we study how to detect and quantify the time evolution of the mass profiles.

Throughout the analysis, we eliminate the galaxies without sub-halos that exist in our catalogues (see next section). We conclude by analyzing what would be the impact of including these galaxies in our measurements.

[1] http://www.astro-wise.org/projects/KIDS/

7.1 Observables in the semi-analytical models

In the previous chapter, the sub-halos were divided into mass bins to better characterize the mass profiles. We also classified them according to infall mass and time spent inside the cluster to optimize the detection of their evolution. We recall that infall mass was the mass of the sub-halo before merging into a cluster, when it was an isolated halo. Unfortunately these quantities are not available in real galaxies. The only possibility, in order to perform similar studies on real data, is to rely on correlated observables.

The use of proxies for the desired physical quantities decreases our ability to produce homogeneous samples of sub-halos. Understanding how the different observables depend on the underlying sub-halo characteristics is essential to obtain optimal results with our method. Investigating these relations is beyond the scope of this thesis, we only make use of the galaxy catalogues at our reach. Our intention is to describe them properly in order to interpret our results.

We must make a remark: there are galaxies inside the galaxy catalogues without sub-halos. Sub-halos, due to tidal forces eventually dissolve into the main halo. As galaxy catalogues are produced, whenever the sub-halo is no longer detected, the hosted galaxy is still maintained and placed at the position of the most bound particle of the formerly detected sub-halo. These galaxies eventually disappear as they merge into the main halo galaxy, but only after a period of time. We do not use these so-called *type-2* galaxies. Including them in the analysis would only produce noise, and there are several reasons not to consider them:

- they are hosted by sub-halos below or around the mass resolution limit;
- they exist under the assumption that the stripping of a large part of the sub-halo does not destroy the hosted galaxy;
- the survival time is heavily influenced by the model.

At the end of this Chapter, we study nevertheless how our results would change by including them.

For our investigation we have studied proxies for the mass, the infall mass and the age of the sub-halos (time spent inside the cluster). The results are shown in Fig. 7.1. There, we have marked those proxies that we use for our analysis with a black frame. In the rest of the plots we present proxies which show a weak correlation with the desired quantity. We only used galaxies hosted by sub-halos and imposed an apparent magnitude limit of 22 in the SDSS r band in order not to include too faint galaxies.

Among the quantities we want to infer from the observables, the age of the sub-halos is the most difficult to obtain. We could not find any observable in the semi-analytical catalogues with a strong correlation with this quantity. In Fig. 7.1, on the top left panel we have plotted the projected radial separation versus the age of the sub-halo. Sub-halos fall gradually into the main halo, and so the distance should be an indicator of the age. However, we cannot measure the true separation, only the projected one, and this is too weakly correlated.

In the top right and the middle left panels, we plot the redshift-corrected $r - i$ and $u - r$, color indices as a function of sub-halo age. Color indices change with redshift. However, it is known that red cluster galaxies at a given redshift gather in a thin color range: the red cluster sequence. We identify for each redshift the red cluster sequence, and we rescale its mean color to zero. By doing this, we can compare galaxies at different

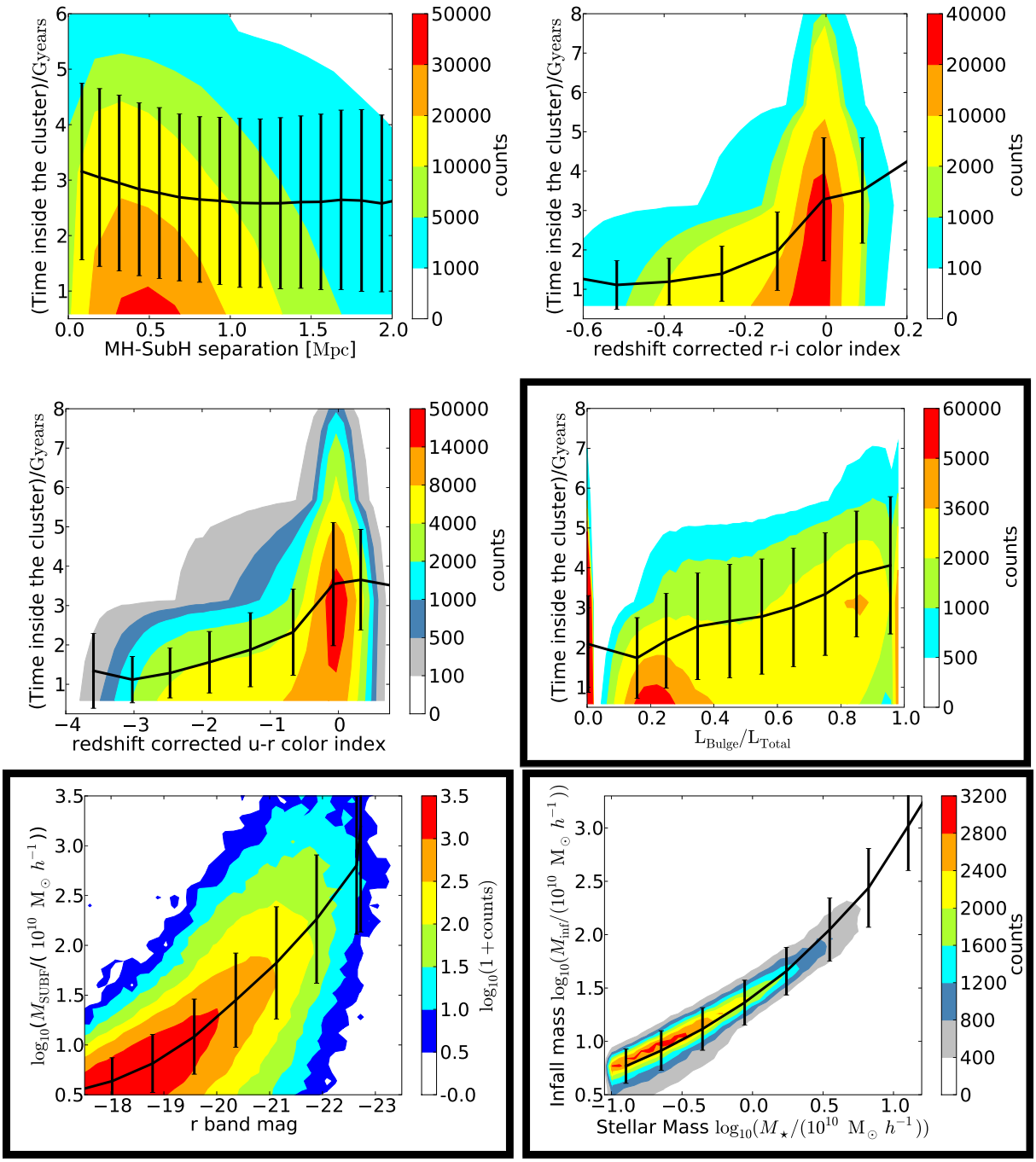

Figure 7.1: Distribution contours of different observables studied to classify the galaxies in our samples. With a black frames, we outline those finally used. We overplot the mean value of the studied sub-halo quantity for each bin in the galaxy observable. We plot the standard deviation as error. Top left: age of the sub-halo versus the projected radial separation of the sub-halo from the main halo center. Top right: age of the sub-halo versus redshift-corrected $r-i$ color. Middle left: age of the sub-halo versus redshift-corrected $u-r$ color. See text for a description of the redshift correction. Middle right: age of the sub-halo versus our morphology estimator. Bottom left: $M_{\rm SUBF}$ versus absolute observer frame SDSS r band magnitude. Bottom right: $M_{\rm inf}$ versus stellar mass (this plot reproduces the results from Wang et al. 2006). Note that the contour levels are not equally spaced except for the bottom left panel. The contour levels were chosen to show the overall shape of the distribution.

redshifts. In both cases one can see that this observable is weakly correlated with the sub-halo age. Note that the contour levels are not equally spaced.

The most sensible proxy for the age of the sub-halos that we found is their morphology. There is no proper morphology parameter in the catalogues, so we used the ratio between the luminosities of the bulge and the disk (middle right panel). We must remark that there is a large amount of galaxies which only have a disk component, or only have a bulge. In both cases the age of these sub-halos is highly degenerate. These two cases are filtered as they do not help our analysis and should be easy to avoid on real surveys. Morphology can be obtained easily in real data even for poorly resolved galaxies for which we can use color gradients (Park & Choi 2005).

The commonly used proxy for mass is luminosity. We experiment with the r band absolute observer frame magnitude as a proxy for $M_{\rm SUBF}$ (see Sect. 3.6 for a definition of this mass). One can observe the relation in the bottom left panel. Other luminosity bands give qualitatively similar results.

Finally we use the stellar mass as a proxy for infall mass $M_{\rm inf}$. This last result is shown in the bottom right panel. The results from Wang et al. (2006) show that there is a strong correlation between the two quantities.

The accuracy of the semi-analytical catalogues is beyond the scope of our work, and yet the results we present in the following depend strongly on them. Our aim is to show the feasibility of extracting information with the method we propose. Corrections to the relation between sub-halo mass and galaxy properties may change the predictions we offer. It is unlikely though that these corrections invalidate the studies presented in this thesis.

7.2 Simulated Surveys

Different surveys result in different galaxy catalogues. In any lensing analysis we split the galaxy catalogue into a foreground sample which contains the lenses, and a background sample where the galaxy images are lensed. The first determines which type of galaxy we can study, whereas the second defines how good our measurements can be. Both depend on the magnitude limits and the solid angle of the survey. The foreground sample has been previously defined: in Sect. 5.3.1 we considered the host clusters; in Chap. 6 we described the host sub-halos; and in the last section we described the observables which we use to characterize the galaxies. Now we discuss the background sample.

For the background galaxy sample, we place points uniformly and randomly across the survey area. Using true galaxy positions would introduce systematics coming from the clustering of galaxies. We leave these problems for future analysis. For the redshift distribution of the galaxy sample we use the well-known formula

$$p(z) = z^{\alpha}\, \exp\left(-(z/z_0)^{\beta}\right)\, \frac{\beta}{(z_0)^{1+\alpha}\, \Gamma\left(\frac{1+\alpha}{\beta}\right)}, \tag{7.1}$$

where $\alpha = 2$, $\beta = 1.5$ and z_0 is a function of the median redshift of the survey. Once the random redshift distribution is produced, we project the points into the planes of our ray-tracing simulations, in order to compute the shear and magnification. This process discretizes our background redshift sample.

We add an intrinsic ellipticity to the shear in order to mimic real galaxy images (Eq. 4.22). The modulus of the intrinsic ellipticity is drawn from a Gaussian distribution

with zero mean and standard deviation $\sigma_{|\epsilon_i|}$. This intrinsic ellipticity is the dominant source of noise in our measurement. The only way to reduce it is by increasing the size of the sample where galaxy-galaxy lensing is computed. We can safely assume that the final error in the measurement due to the intrinsic ellipticity ϵ_i is $\sigma_{|\epsilon_i|}/\sqrt{n}$, where n is the total number of foreground-background galaxy pairs used to measure the final $\Delta\Sigma(\xi)$. This error can be reduced by observing fainter galaxies and by considering larger portions of the sky.

We can change the magnitude limits of our sample, but unfortunately, the size of the simulated survey that we obtain by ray-tracing is limited. Although the simulated area is large, it is still below the size for which we want to make predictions.

We consider that reducing $\sigma_{|\epsilon_i|}$ is equivalent to an increase in the galaxy sample. Assuming this relation allows us to simulate larger survey areas. From galaxy studies, the standard deviation of intrinsic ellipticity noise is typically $\sigma_{|\epsilon_i|} = 0.3$. Under the previous arguments, we achieve the results from a 36 times larger survey than the simulated one, by using a value of $\sigma_{|\epsilon_i|} = 0.05$. This size is approximately the size of Pan-STARRS or LSST survey ($36 \times 512 = 18432$ square degrees).

In order to examine the expected signal quality, we study surveys with different background populations (parameterized with the galaxy number density and the median redshift), and with different survey areas (parameterized with the effective $\sigma_{|\epsilon_i|}$). We focus on DES and LSST, exploring also variations of the main parameters that define them. The different surveys that we investigate are in Tab. 7.1. For each survey, we compute a mean signal-to-noise ratio (SNR) for a set of luminosity bins. To do this, for each bin we average over eight logarithmic bins in projected radius ξ in the range 0.043 to 0.12 Mpc. Our signal-to-noise ratio can be expressed as:

$$SNR = \left\langle \frac{\epsilon_t(\xi)}{\sigma\big(\epsilon_t(\xi)\big)} \right\rangle_\xi \tag{7.2}$$

Note that doubling the number of bins roughly halves the signal-to-noise. The values for each luminosity neither take into account correlated noise, nor express how well a particular feature can be detected. Nonetheless, they are compact estimators and offer a rule of thumb for the expected SNR.

The LSST survey and the DES survey, are expected to be around 10% larger than the ones that we consider. We take with it a realistic approach and we consider the possibility of nuissances that may reduce the usable area. With our mock survey DES-WIDE we analyze the impact of having a survey like DES but with the solid angle of LSST. With

Table 7.1: Parameters for the different surveys considered in our detectability study. We present the median redshift, the background galaxy number density, the standard deviation of the intrinsic ellipticity and its equivalent survey solid angle.

survey	**Median** z	$\frac{\text{gal}}{\text{arcmin}^2}$	$\sigma_{\lvert\epsilon_i\rvert}$	degrees^2
LSST	1.2	40	0.05	18432
INT	0.9	25	0.1	4608
DES-WIDE	0.68	12	0.05	18432
DES	0.68	12	0.1	4608

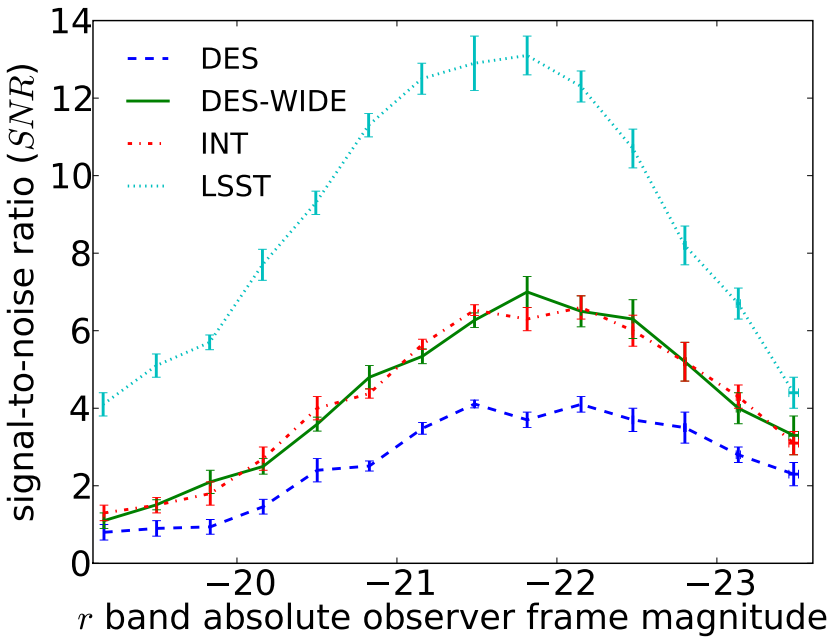

Figure 7.2: Our compact SNR estimator,for the different luminosity bins for each mock survey considered. For each luminosity bin we plot the mean magnitude. The errors plotted are standard errors obtained from the sample.

INT we inspect the effect of achieving more depth on the same angular size as DES. These two cases do not correspond to any real or planned survey.

In Fig. 7.2, we present our SNR estimation for different luminosity bins in r band absolute magnitude, for the four mock surveys considered. We define 14 bins in luminosity, with a bin width of 0.33 magnitudes, between the values -19 and -23.66. For the LSST-like survey, we only considered lens galaxies with an apparent magnitude smaller than 26 in r band. For the other three surveys the magnitude limit was 22 in r band. We impose in all cases a maximum redshift of $z = 0.9055$ for the foreground sample as explained in Sect. 5.3.1. In Fig. 7.2, one can see the average detectability of a profile for a given luminosity. Bright galaxies have a larger halo with a stronger signal. Faint galaxies are more numerous and their samples have a smaller statistical error. The signal-to-noise depends on these two competing effects. Due to this, the maximal signal-to-noise ratios are around the ninth bin. The least luminous bins are below the two sigma level except for a LSST-like survey. For those cases we must use broader luminosity bins.

Another point is the fact that, although INT has nominally half the number of galaxies than DES-WIDE, the SNR values are very similar. Note that DES-WIDE has a smaller galaxy number density, but more than four times the area of INT. This suggests that the sensitivity of the survey is more important than its area.

7.3 Optimal sub-halo mass proxy

The study that we propose is designed to measure the mass profile of the cluster sub-halos hosting satellite galaxies. The accuracy of the measurement depends on the survey characteristics. In the worst case scenario we still want to be able to compress the information coming from the whole profile into a meaningful quantity. This section is dedicated to

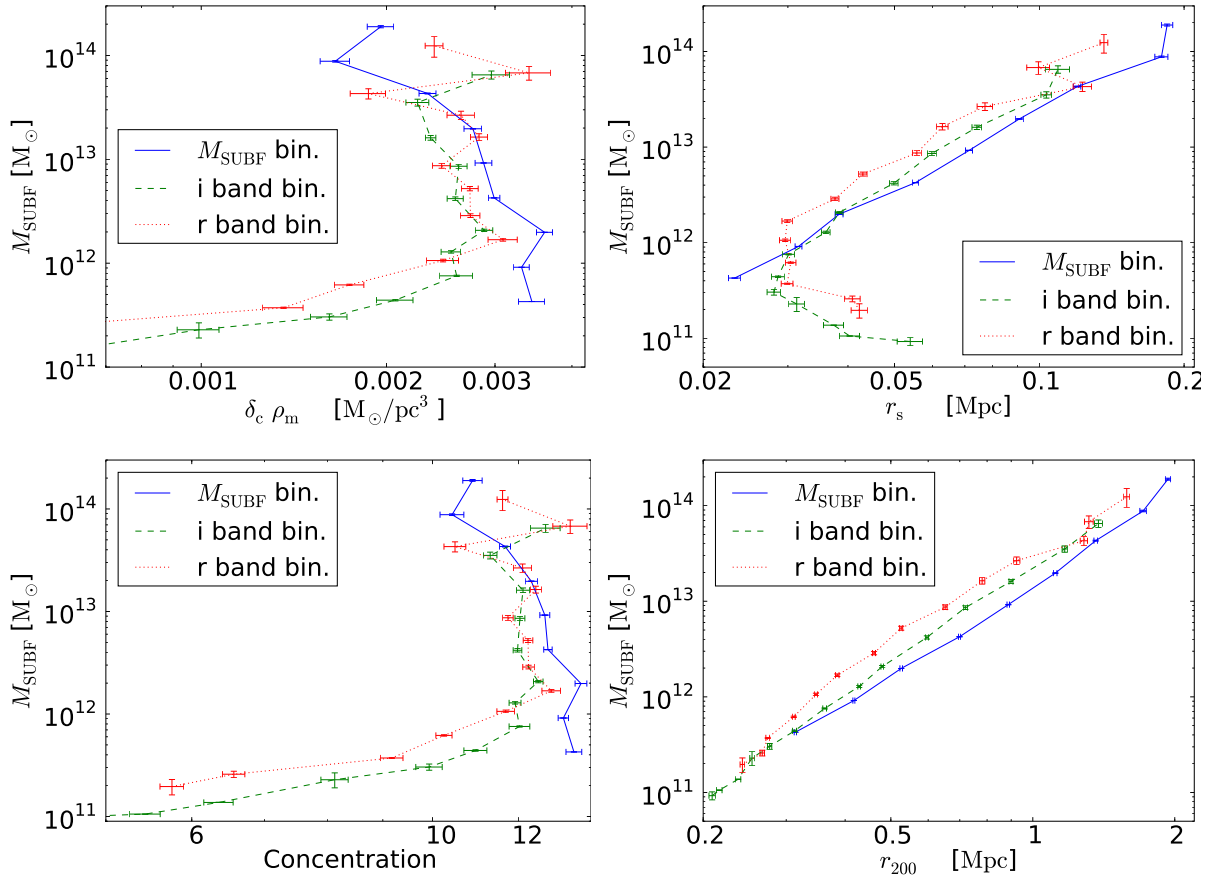

Figure 7.3: Mass proxies for different binnings. We used projected mass maps $\Sigma(\boldsymbol{\beta})$. In the top panels we show the NFW parameter set $r_{\rm s}$ and $\delta_{\rm c}\,\rho_{\rm m}$, whereas in the bottom ones we present r_{200} and concentration. Note the similarity between concentration and $\delta_{\rm c}\rho_{\rm m}$ is only coincidental.

explore which quantity is more robust to be useful as mass proxy. We assume $M_{\rm SUBF}$ as the best possible definition of sub-halo mass. Again, in order to get the best estimation of the parameters we want to study, we use the projected mass maps $\Sigma(\boldsymbol{\beta})$ defined in Sect. 4.5.

The following results depend on the underlying profile which in our case we have shown is well represented by an NFW. If halos of real galaxies differ from the simulated ones, the conclusions derived have to be revised.

In Sect. 6.3, we identified two possible parameters as proxies for $M_{\rm SUBF}$: $r_{\rm s}$ and r_{200}. The first is a parameter of the NFW profile, the second parameterizes the mass in isolated halos but a priori does not have any meaning for sub-halos. In Fig. 7.3, we present $M_{\rm SUBF}$ versus $r_{\rm s}$, $\delta_{\rm c}\rho_{\rm m}$, concentration, and r_{200}. We have binned the sub-halos in three different ways: according to $M_{\rm SUBF}$ as in Fig. 6.6; according to the observer frame absolute magnitude in the r band; and according to the observer frame absolute magnitude in the i band. The binning according to mass is different from the luminosity binning. The consequence is that the mean profile for the same given mean mass changes. The luminosity ranges are the same as for Fig. 7.2 for both r and i bands, and the mass ranges are described in Tab. 6.1. The best quantity to infer mass is the measured r_{200}. The correlation with $M_{\rm SUBF}$ is stronger than for the rest of parameters that we analyzed. In Tab. 7.2 we list the values for the best fitting power laws for $M_{\rm SUBF}$ as a function of r_{200}. We are not able to estimate $M_{\rm SUBF}$ with high precision through r_{200} due to the dependence

on the galaxy classification, but we may be able to derive useful constraints. We can also apply the scaling relation here to define a proper mass-luminosity relation later on.

Table 7.2: Best fitting parameters for the relation $M_{\rm SUBF}(r_{200})/{\rm M}_\odot = A \times (r_{200}/{\rm Mpc})^n$ coming from the data displayed in Fig. 7.3 in the bottom right panel. We present the mean and the standard error of the posterior of the parameters.

Data	**n**	$\log_{10}(\mathbf{A})$
Mass bin.	3.18±0.05	13.172±0.013
i band bin.	3.62±0.15	13.64±0.07
r band bin.	3.57±0.12	13.42±0.05

7.4 Weak lensing approximation

In Sect. 4.4, we assumed that the weak lensing approximation is valid in our measurements. This means that $\kappa \ll 1$, and that the reduced shear effectively equals the shear. This assumption is motivated by the fact that we only aim to use weakly lensed images where this should be the case. However, sub-halos are always in high density regions where the convergence is large, therefore we need to check the assumption. For this purpose, we compare the signal measured from a shear catalogue and the signal measured from a reduced shear catalogue. In Fig. 7.4, we plot a selection of luminosity bins for a LSST-like survey. We used red solid lines for the profiles measured using reduced shear, and blue dashed lines for the measurements using only shear. In the range in radius where we consider that our measurement is unbiased, we plotted error bars.

When we compare the measurements done using reduced shear and shear, we see a systematic difference for the most luminous sub-halos. One can see from the plots that those coming from a reduced shear catalogue are steeper. The effect is stronger the more luminous the galaxies are, and occurs at the innermost part of the profile. The position of the sub-halo inside the cluster plays no role. Smaller halos are closer to the cluster center, where the density is higher, but for these the effect is not visible.

The difference between reduced shear and shear is small and only important for the most massive sub-halos (note the two uppermost profiles with error bars in Fig. 7.4). The shape of the profile is modified. This is a source of uncertainty in the parameter estimation. If we can no longer assume that we measure shear, the expectation value of our estimation does not follow Eq. (4.31). Instead we have:

$$\begin{aligned}\langle\widehat{\Delta\Sigma}(\xi)\rangle &= \langle\Sigma_{\rm crit}(z_{\rm s}, z_{\rm d})\,\epsilon_{\rm t}(\xi)\rangle \\ &= \left\langle \frac{\gamma_{\rm t}(\xi)\,\Sigma_{\rm crit}(z_{\rm s}, z_{\rm d})}{1-\Sigma(\xi)/\Sigma_{\rm crit}(z_{\rm s}, z_{\rm d})} \right\rangle \neq \frac{\Delta\Sigma}{1-\Sigma(\xi)/\langle\Sigma_{\rm crit}(z_{\rm s}, z_{\rm d})\rangle}. \end{aligned} \tag{7.3}$$

Note that in the second step we averaged over the distribution of intrinsic ellipticities and obtained $\gamma_{\rm t}(\xi)$ from $\epsilon_{\rm t}(\xi)$. We assumed that the intrinsic ellipticity noise is redshift independent, and we denoted with $z_{\rm s}$ and $z_{\rm d}$ source and lens redshift respectively. With the last step of the previous equation we want to stress that the averaging procedure does not commute with non-linear operations. Therefore, the $1-\kappa$ correction cannot be easily accounted for. The exact modeling of the measured signal includes many non-trivial contributions.

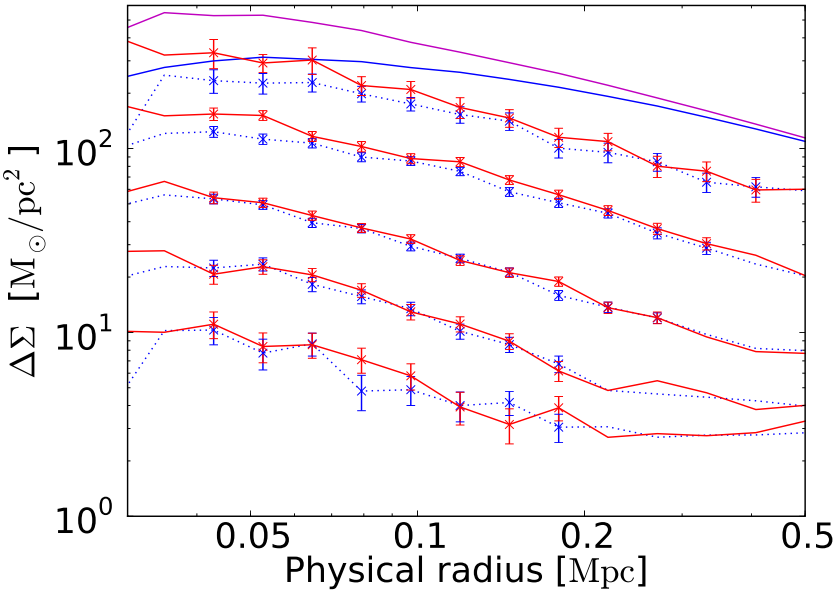

Figure 7.4: Comparison between the profiles measured with a shear catalogue (blue dashed-dotted) and a reduced shear catalogue (solid red line). The errors are the standard error for each radial bin. We present a selection of luminosity bins for a LSST-like survey. The magnitude limit is 26 in apparent magnitude in r band. The uppermost profiles, plotted without errors are the profiles for the host cluster, shown as a visual reference.

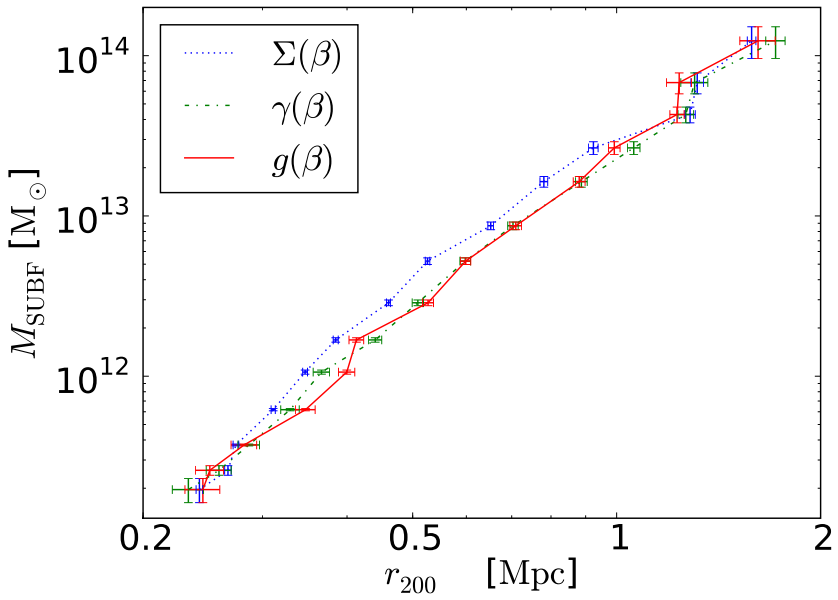

Figure 7.5: Determination of r_{200} as a function of $M_{\rm SUBF}$. We plot the results coming from different ways of computing $\Delta\Sigma(\xi)$: from projected mass maps $\Sigma(\boldsymbol{\beta})$; from a shear catalogue $\gamma(\boldsymbol{\beta})$; and from a reduced shear catalogue $g(\boldsymbol{\beta})$. The binning of the galaxies is done according to its r band absolute magnitude in the observer frame. The magnitude limit is 26 in apparent magnitude in r band.

However, the use of the weak lensing approximation is justified. We estimate r_{200} using the projected mass maps $\Sigma(\boldsymbol{\beta})$, the shear catalogue $\gamma(\boldsymbol{\beta})$ and the reduced shear catalogue $g(\boldsymbol{\beta})$, to quantify its accuracy. The results are presented in Fig. 7.5. The values for r_{200} are almost insensitive to the initial data set.

The difference between the values from $\Sigma(\boldsymbol{\beta})$, and the other two is not unexpected. When we use the projected mass maps $\Sigma(\boldsymbol{\beta})$, there is no contamination from matter structures in the foreground our background. Moreover, the measurements made with $g(\boldsymbol{\beta})$ or $\gamma(\boldsymbol{\beta})$ due to our weighting scheme (see Sect. 5.2) depend on the redshift distribution of the background sources, which is not the case for the measurements done with $\Sigma(\boldsymbol{\beta})$.

The weak lensing approximation introduces a systematic error in the estimation of the $\Delta\Sigma$ profile, for which it is difficult to account. On the other hand, it affects only the innermost part of the profile of massive sub-halos. The analysis of a possible truncation radius or other time evolution features is possible. As shown, it does not affect the estimation of the mass through r_{200}. Therefore, we shall use the weak lensing approximation in the following.

7.5 Luminosity

One of the essential analyses in galaxy evolution is the relation between mass and luminosity in different bands. After defining the lensing and the shear catalogues we address this point.

We want to treat in detail the estimation of a mass-luminosity relation, focusing on the results coming from a LSST-like and a DES-like survey. We defined the luminosity bins regarding the detectability of measurements, which we presented in Fig. 7.2. The SDSS r band is the one used most frequently for weak-lensing analysis. Therefore, it is the one that we use in our work. A mass-luminosity relation in different bands is left for future analyses. The magnitude limit in the case of the DES-like survey is 22 in apparent magnitude, whereas in the case of the LSST-like survey it is 26.

In the right panel in Fig. 7.6, we present our prediction for a LSST-like survey, and in the left panel for a DES-like survey. The signal-to-noise is lower in a DES-like survey and we must use fewer luminosity bins in order to achieve a sensible result. The binning characterization is shown in Tab. 7.3.

In order to get the sub-halo mass M_{SUBF}, we need to measure r_{200}, and derive it using a scaling relation between M_{SUBF} and r_{200}. The mass that we estimate therefore depends first on the mass definition and second in the scaling between the mass and the measured r_{200}. In any case, we observe that it is possible to infer constraints for a mass-luminosity relation. The measurement of such a relation can also help to distinguish between two competing model in sub-structure formation. With our method, we probe larger scales than stellar dynamics and we consider large samples of galaxies compared to strong lensing analysis. Therefore, the data that this analysis provides is difficult to obtain by other means. The final mass-luminosity relations, for both a LSST and a DES-like survey, are shown in Fig. 7.7.

In principle, from our analysis we conclude that we can constrain the mass-luminosity relation over two decades in mass. The different luminosity binning and samples used in each survey produce differences. The brightest bins are populated with the same objects and therefore they are similar. At the fainter end, the different magnitude cuts in combination with the broader bins changes the mass-luminosity relation (Fig. 7.7, top

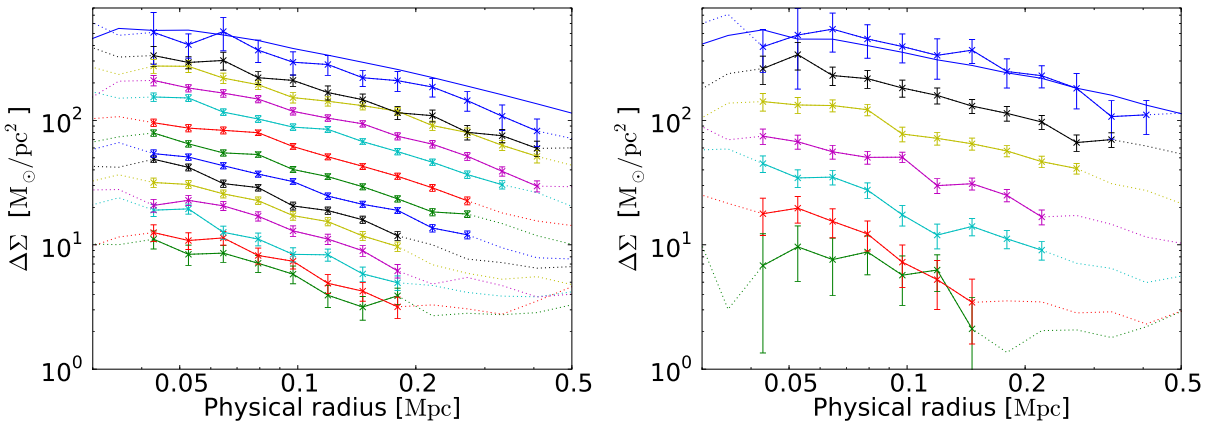

Figure 7.6: Measured $\Delta\Sigma(\xi)$ profiles of galaxies binned by absolute magnitude in r band. Left panel: LSST-like survey, with a limiting magnitude of 26 in r band apparent magnitude. Right panel: DES-like survey, with a limiting magnitude of 22 in r band apparent magnitude. With solid lines the data used for the analysis later on, the range was derive from our results in Tab. 6.4. The solid blue line without error bars is the measurement for the host cluster shown as a visual reference.

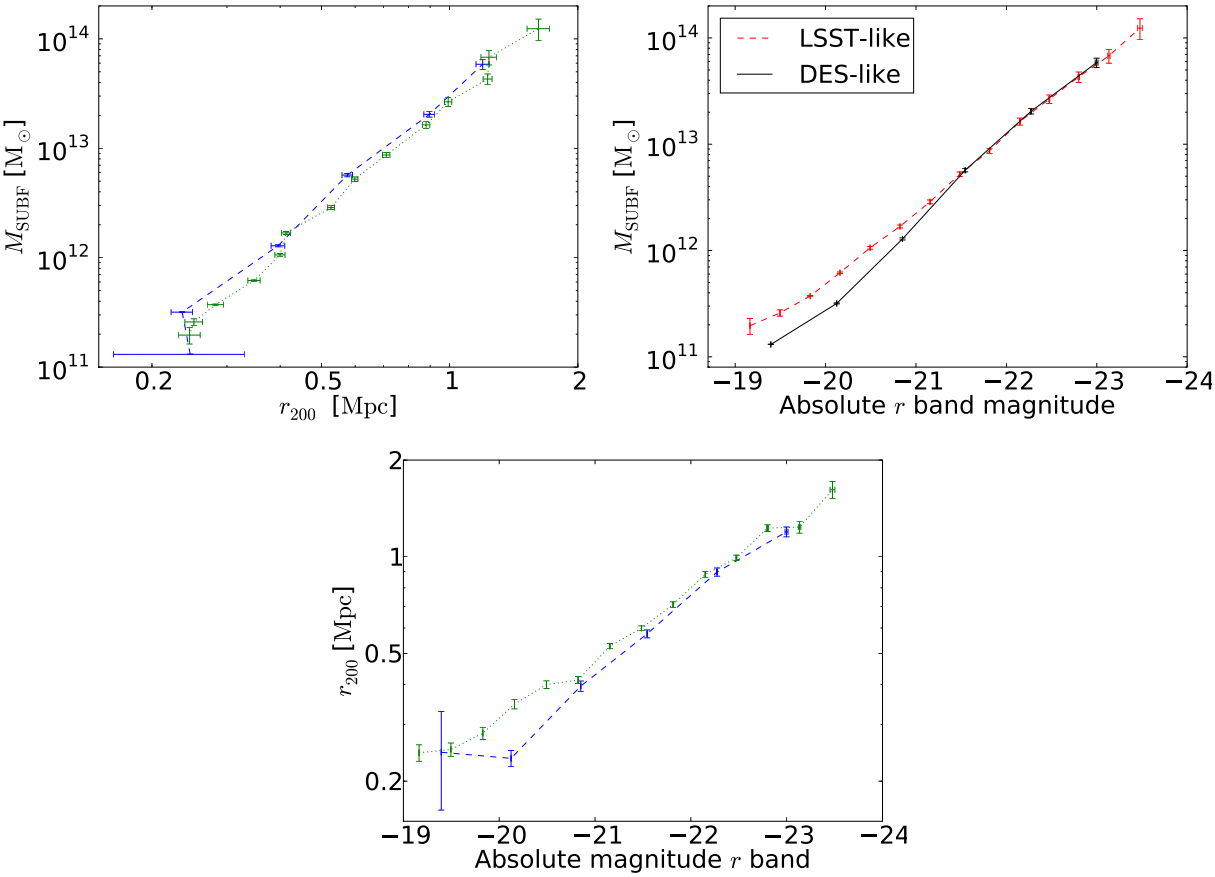

Figure 7.7: Top left: relation between the fitted parameter r_{200} and $M_{\rm SUBF}$. Top right: the mass-luminosity relation obtained from the catalogues. Bottom: the relation between the fitted parameter r_{200} and the absolute observer frame luminosity in r band magnitude. We plot in blue the results for a DES-like survey, and in green for a LSST-like survey.

Table 7.3: Left: binning used for the LSST-like survey, with r band absolute observer frame magnitude. Right: the same for the DES-like survey. We present also the number of sub-halos.

Bin	**# SH**	**r band range**
1	22698	$[-19.00{:}-19.33]$
2	19684	$[-19.33{:}-19.66]$
3	16625	$[-19.66{:}-20.00]$
4	13145	$[-20.00{:}-20.33]$
5	10472	$[-20.33{:}-20.66]$
6	7694	$[-20.66{:}-21.00]$
7	5166	$[-21.00{:}-21.33]$
8	3127	$[-21.33{:}-21.66]$
9	1598	$[-21.66{:}-22.00]$
10	752	$[-22.00{:}-22.33]$
11	354	$[-22.33{:}-22.66]$
12	128	$[-22.66{:}-23.00]$
13	56	$[-23.00{:}-23.33]$
14	15	$[-23.33{:}-23.66]$

Bin	**# SH**	**r band range**
1	11590	$[-19.00{:}-19.75]$
2	13740	$[-19.75{:}-20.50]$
3	12824	$[-20.50{:}-21.25]$
4	5639	$[-21.25{:}-22.00]$
5	1154	$[-22.00{:}-22.75]$
6	145	$[-22.75{:}-23.50]$
7	10	$[-23.50{:}-24.25]$

right panel). Nonetheless the relation between r_{200} and the mass is the same (within the errors) in both cases. The smaller number density of faint galaxies in DES compared to the LSST survey reduces in this case the range and the resolution that we can measure for the mass-luminosity relation. Note that the faintest luminosity bin for the DES-like survey may not be usable as the measurement error becomes too large.

7.6 Stellar mass & morphology

At the end of the previous chapter, we saw how the sub-halo mass profiles change with time. For this purpose, we classified them according to the time spent inside the cluster and the infall mass. The conclusion was that it is not possible to detect a truncation in the original sense, and that the mass density decreased at all scales consistent with a reheating process.

Our goal now is to see whether this behavior can be detected using exclusively galaxy observables. As stated previously, we can only derive conclusions assuming that the semi-analytical catalogues are valid. We can estimate infall mass with some accuracy using stellar mass, but the galaxies in these catalogues have no observable which is strongly correlated with the age of the sub-halo. We take morphology as the best option.

The predicted signals for both LSST and DES are shown in Fig. 7.8 (top and bottom panel respectively). We plot with different symbols the different stellar mass bins. We have also divided each of the former sets according to morphology. The classification ranges for morphology and stellar mass are described in Tab. 7.4 for the top panel and in Tab. 7.5 for the bottom panel. The dashed red lines belong to galaxies with a relatively large bulge, which should belong to old sub-halos. With blue dotted lines we plotted the signals from galaxies with a large disk, which should belong to younger sub-halos. The number of bins we can analyze according to mass is small, since we need to split them into young and old sub-halos. The stellar mass ranges were obtained by experimentation.

Table 7.4: LSST-like survey data. Morphology-stellar mass classification of all sub-halos in clusters for $z < 0.9055$ with a magnitude limit of 26 in r band apparent magnitude. The sub-halos are at least 0.5 Mpc away from the main halo center. Each bin is denoted with the symbol and line type used in the plots. Mass is in units of $10^{10}\,M_\odot$. 2nd column: number of sub-halos. 3rd column: morphology ranges, Disk $\rightarrow 0 < L_{\rm Bulge}/L_{\rm Total} \leq 0.6$, Bulge $\rightarrow 0.6 < L_{\rm Bulge}/L_{\rm Total} < 0.98$. 4th column: mean value of the previous quantity. 5th column: mean age in Gigayears. 6th column: stellar mass bin, I $\rightarrow$ [0.14 : 1.4], II $\rightarrow$ [1.4 : 4.11], III $\rightarrow$ [4.11 : 10.96], IV $\rightarrow$ [10.96 : ∞]. 7th column: mean stellar mass. 8th column: mean infall mass. 9th column: mean SUBFIND mass. 10th column: estimated r_{200} in Mpc.

bin	**# SH**	**Mor.**	$\left\langle \frac{L_{\rm Bulge}}{L_{\rm Total}} \right\rangle$	$\overline{\rm age}$	$M_\star$	$\overline{M_\star}$	$\overline{M_{\rm inf}}$	$\overline{M_{\rm SUBF}}$	$\overline{r_{200}}$
$\times\,\cdot\,\cdot$	66703	Disk	0.3212	2.362	I	0.5839	19.94	11.23	0.216±0.009
$\times$- -	31587	Bulge	0.7526	3.329	I	0.7203	19.91	8.2	0.20±0.02
$+\,\cdot\,\cdot$	20733	Disk	0.3562	2.229	II	2.326	761.3	43.2	0.339±0.010
$+$- -	21952	Bulge	0.7956	3.651	II	2.390	657	23.1	0.284±0.010
$\star\,\cdot\,\cdot$	9101	Disk	0.4157	2.160	III	6.299	2900	162	0.476±0.009
$\star$- -	13003	Bulge	0.8008	3.522	III	6.491	2840	119	0.456±0.009
$\bullet\,\cdot\,\cdot$	762	Disk	0.468	1.91	IV	13.45	860	485	0.62±0.02
$\bullet$- -	2063	Bulge	0.799	2.78	IV	15.60	2070	1280	0.850±0.014

Table 7.5: DES-like survey. Morphology-stellar mass classification of all sub-halos in clusters for $z < 0.9055$ with a magnitude limit of 22 in r band apparent magnitude. Note that we denote each bin with the symbol an line type that we used in the plots. The sub-halos are at least 0.5 Mpc away from the main halo center. We use the same notation as in the previous table, except for the stellar mass ranges: I $\rightarrow$ [1.1 : 4.1], II $\rightarrow$ [4.1 : 13.7].

bin	**# SH**	**Mor.**	$\left\langle \frac{L_{\rm Bulge}}{L_{\rm Total}} \right\rangle$	$\overline{\rm age}$	$M_\star$	$\overline{M_\star}$	$\overline{M_{\rm inf}}$	$\overline{M_{\rm SUBF}}$	$\overline{r_{200}}$
$+\,\cdot\,\cdot$	8378	Disk	0.3494	2.386	I	2.233	69	42.7	0.257±0.013
$+$- -	8809	Bulge	0.8029	4.30	I	2.308	60.9	20.8	0.250±0.015
$\star\,\cdot\,\cdot$	4577	Disk	0.4186	2.32	II	6.92	302	178	0.446±0.010
$\star$- -	7531	Bulge	0.8066	3.90	II	7.23	244	153	0.467±0.010

In the case of the DES-like survey we focused on a particular stellar mass range where the effect was strongest. Note that we only plot error bars for the range where we consider that our measure is unbiased.

The sub-division according to morphology does not always produce the desired results. One can see from the last two rows in Tab. 7.4, that for the most massive satellite galaxies, the infall mass (eighth column) of the two morphology bins are very different. Older sub-halos have a significantly larger mean infall mass, which indicates that they evolved in a different manner. This difference masks any possible detection of mass loss due to tidal stripping by the cluster. Since the number of sub-halos used for this bin is much smaller than for other bins, we cannot exclude that the sample was too small. A deeper understanding of galaxy evolution is needed in order to fully assert the causes.

Except for the previously referred bin, it is easy to detect from the top plot in Fig. 7.8 that the profile amplitude of bulge-dominated galaxies is smaller respect to those with a larger disk. This allows us to infer a change of the mass profile as predicted in Sect. 6.4. For the DES-like survey the detection is unfortunately not so strong.

Again, the main problem is how to quantify the change on the profile. When we used only sub-halos without involving galaxies we already saw that r_{200} (the best estimator according to Sect. 7.3) cannot quantify the mass loss with precision. Here this is also the

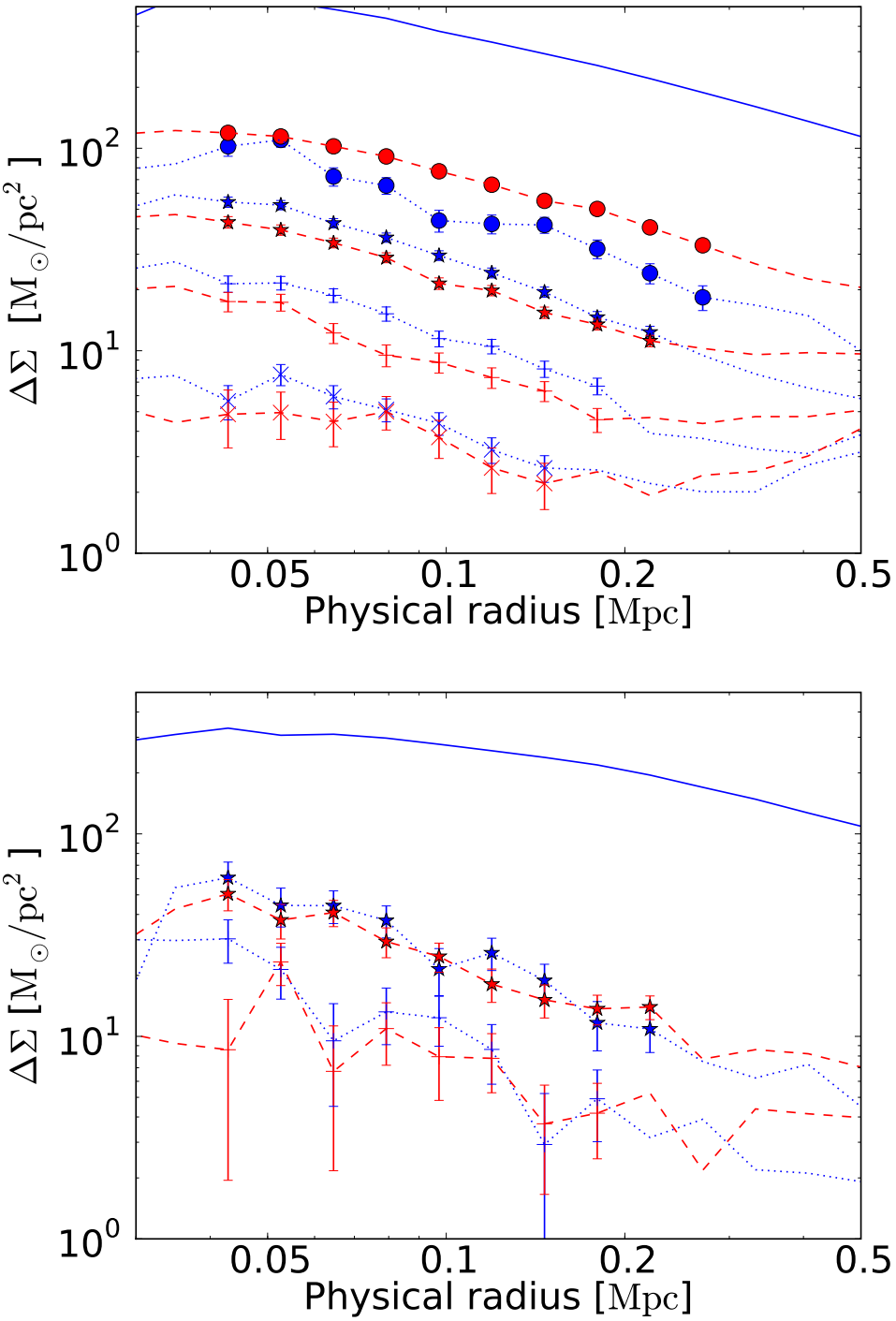

Figure 7.8: Top panel: morphology-stellar mass classification for a LSST-like survey. Bottom panel: morphology-stellar mass classification for a DES-like survey. Blue dotted lines correspond to galaxies with a relatively small bulge. Red dashed lines correspond to galaxies with a large bulge which spent more time inside a cluster. The different symbols distinguish different stellar mass bins. The classification ranges for morphology and stellar mass are described in Tab. 7.4 for the top panel and in Tab. 7.5 for the bottom panel. The blue solid line is the measurement for the host cluster shown for visual reference. We only plot error bars where we consider that the measure is unbiased.

case. Compare the r_{200} values and their errors for galaxies of the same stellar mass in Tab. 7.4 and Tab. 7.5. The difference between them for old and young sub-halos is within the errors. In the case of a DES-like survey the quantification of the change from r_{200} is not possible. Nevertheless, this measurements could help to constrain the mass loss in combination with other methods like e.g. the analysis of the stellar dynamics of the galaxy.

7.7 Type-2 galaxies

As previously stated, the semi-analytical catalogues are populated with satellite galaxies without a host sub-halo (type-2 galaxies). The sub-halo could not be detected in those cases. The sub-halos are then below the resolution limit of the simulation, or dissolved due to tidal stripping by the host main halo, or they are sub-halos of sub-halos (as explained in Sect. 3.6, SUBFIND only considers one level of substructure). The existence in reality of satellite galaxies with very small dark matter sub-halos, such as type-2 galaxies represent, is difficult to assert. Moreover, even if we assume that they are plausible, it is impossible to evaluate how realistic our set of type-2 galaxies is, at least with the data we have available. Nevertheless, to understand what can be the impact of such galaxies on our measurements, we analyze how our results are altered once we include them.

For a broad luminosity range, the number of type-2 galaxies is comparable to galaxies hosted by sub-halos (type-1 galaxies). In Fig. 7.9 we present the luminosity histograms for both. If we want to study sub-halo mass profiles we must identify type-2 galaxies in order to avoid uncertainty in our modeling.

In Fig. 7.10, we present the $\Delta\Sigma(\xi)$ measurements for type-2 galaxies only. We plot a selection of luminosity bins from Tab. 7.3 with a solid red line. For comparison, we plot with dashed-dotted blue lines the previous measurements for type-1 galaxies. We only plot the region assumed to be unbiased. Note that there are no type-2 galaxies in the bin with the highest luminosity. As one can see, the $\Delta\Sigma(\xi)$ profiles for the type-2 are not well described by an NFW profile. Two of the type-2 profiles show an increase, whereas the NFW profile decreases monotonically. For the third profile, the best fitting scale radius for an NFW profile would be larger than the one of the host main halo, and the characteristic density would be extremely small.

One possible explanation is that, on average, type-2 galaxies are correlated with the position of sub-halos. The presence of a nearby sub-halo would make it more difficult to detect a sub-halo for these galaxies (if there were any at all) and help to explain why none could be found.

To verify whether type-2 galaxies are correlated with sub-halo positions is a complex task since we are studying galaxies in clusters. Our galaxies are not uniformly distributed, but localized in a host cluster, and with a higher density towards the cluster center. If we try to quantify the afore-mentioned correlation without considering this fact, our results will only reflect that the galaxies clump towards the center of the cluster.

In order to quantify the correlation between type-a galaxies and type-b galaxies within a cluster, we proceed as follows. We compute the number density of type-b galaxies as a function of radial separation from a type-a galaxy, and from the calibration points of type-a galaxies. If we assume that clusters are parity-symmetric, the inclusion of measurements around the calibration points allows us to quantify the expectation value of the background number density. The ratio between the two number densities (from

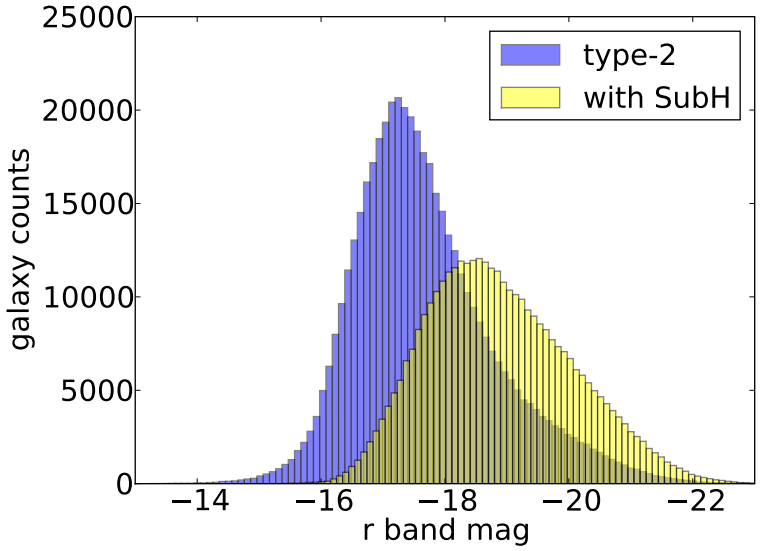

Figure 7.9: Luminosity histograms of galaxies with (type-1) and without a host sub-halo (type-2). The luminosity is in SDSS r band absolute magnitudes. All galaxies have a smaller apparent magnitude than 22 in r band.

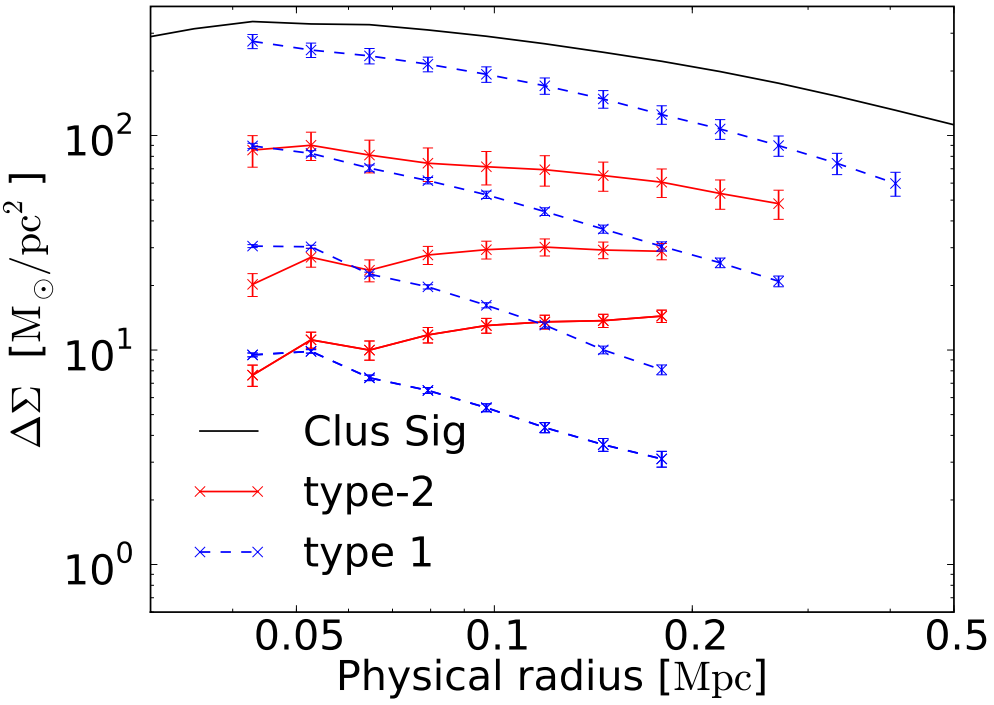

Figure 7.10: Excess surface mass density $\Delta\Sigma(\xi)$ for a few r band luminosity bins for a LSST-like survey (see Tab. 7.3). Here we used the projected mass maps $\Sigma(\boldsymbol{\beta})$ to obtain the measurements. With solid red lines we plot the results from type-2 galaxies only. With blue dashed lines we plot the previous results, for galaxies with a sub-halo. The range where we assume the measurement is unbiased was derived from our results in Tab. 6.4. These are our results for a LSST-like survey. The solid black line is the measurement for the average cluster, shown as a visual reference.

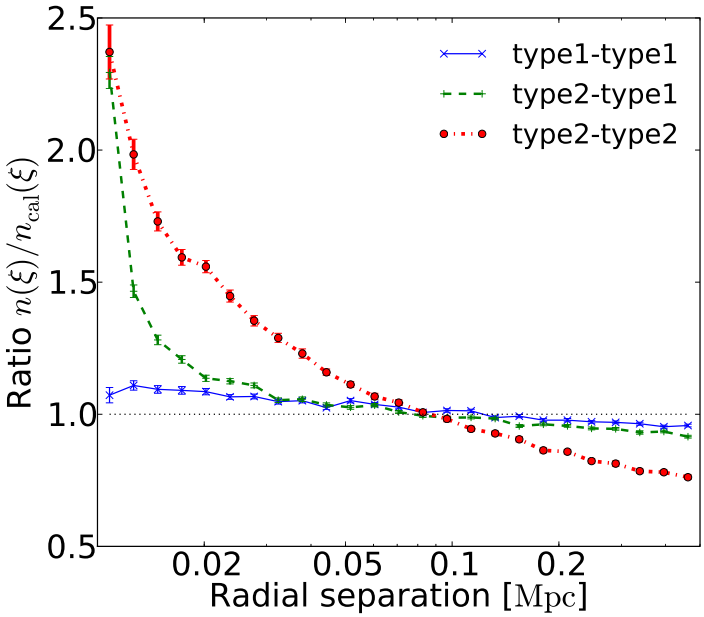

Figure 7.11: Ratio between the galaxy number density around a galaxy and around its corresponding calibration point, as a function of distance. We also plot as visual reference the line where the ratio is 1. The dotted-dashed red line corresponds to the excess type-2 number density around type-2. The solid blue line corresponds to the excess type-1 number density around type-1. The green dashed line corresponds to the excess type-1 number density around type-2. In all cases we only considered galaxies 0.5 Mpc away from the cluster center.

the type-a position and from the corresponding calibration point), is then an estimate of the desired correlation. In Fig. 7.11, we show this estimate for the three different possible cases. The values are given by the mean over all the survey patches (128 patches), and the errors are the standard errors of the samples. As one can see, the excess number density for type-1 pairs is small (solid blue line), especially if we compare it to the excess number density of type-2 pairs (dotted-dashed red line) or even compared to the pairs formed by a type-2 and a type-1 galaxy. For scales above 0.1 Mpc, the excess number density tends to be smaller than one, which suggests that the existence of an average clustering scale for satellite galaxies within themselves inside the cluster.

The signals in Fig. 7.11 show a correlation in the position of type-2 and type-1 as we previously suggested, that we can actually quantify. However, the source for the signals measured around type-2 galaxies in Fig. 7.10 is not solely due to nearby sub-halos. A visual inspection of the projected mass maps reveals a more complex situation. In Fig. 7.12, we present a zoom of the same cluster that we showed in Fig. 5.3 and in Fig. 6.1. It is the upper-right fraction of the central part of the mentioned cluster. The center of the cluster is marked with a large white circle. We overplot to the mass density color code the position of the type-1 (pluses), and type-2 galaxies (diamonds).

At first sight, type-2 galaxies traces the cluster mass density and do not seem to have any related sub-halo. Nevertheless, the correlation between sub-halos and type-2 galaxies can be observed in a few spots. We highlight with tailed arrows the cases which correspond to the correlation between type-1 and type-2 galaxies. In some cases, which we marked with simple arrows, the type-2 galaxy seems to be located on top of another type-1 galaxy. These could be a projection effect, but also two satellite galaxies merging among themselves. Finally, with double-tipped arrows we show cases where we can detect a mass overdensity which appears to host a type-2 galaxy but no type-1 galaxy. These, which

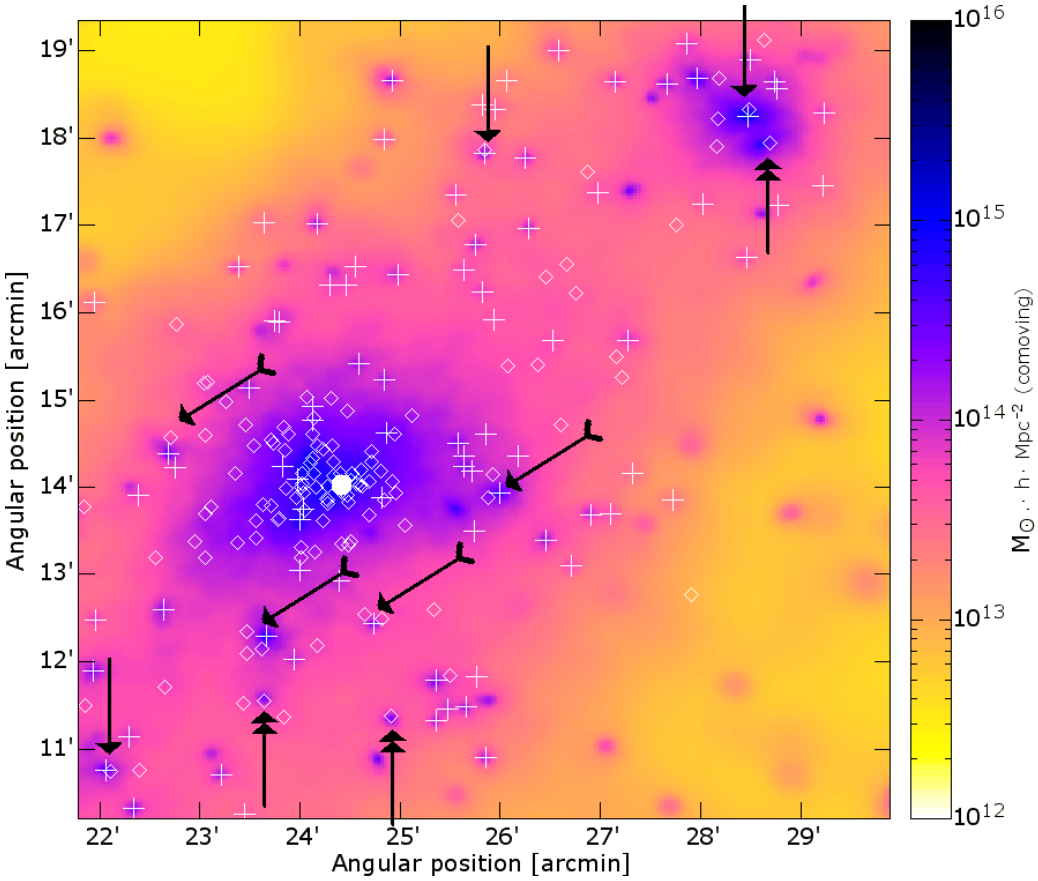

Figure 7.12: A detailed view of central part of the projected mass map of the cluster shown in Fig. 6.1. The proportion between physical separation and angular separation is approximately 0.170 Mpc per arcminute. We show the center of the cluster with a white circle. We overplot the position of type-1 galaxies with pluses and the position of type-2 galaxies with diamonds. We highlight with tailed arrows the position of type-2 galaxies which are near to a type-1 galaxy; with simple arrows, cases where a type-2 seems to be on top of a type-1 galaxy; and with doubled-tipped arrows, type-2 galaxies which appear on top of an overdensity without type-1 galaxy.

can be seen in many clusters, show some kind of unexplained behavior of the SUBFIND algorithm (see Sect. 3.6).

We check our conclusion about the source for the measured $\Delta\Sigma(\xi)$ in type-2 galaxies. In Fig. 7.13, we present $\Delta\Sigma(\xi)$ for type-2 galaxies, for a selection of luminosity bins. In both panels we plot the signals for all type-2 galaxies in red with pluses, (the same signals as in Fig. 7.10). In the left panel, we present with black crosses the signals for type-2 galaxies, but rejecting from the sample those which were nearer than 0.05 Mpc to a type-1 galaxy. In the right panel, we present with green squares the signals for type-2 galaxies which have no other galaxy nearer than 0.05 Mpc (both type-1 and type-2). The line type allows to compare the lines for the same luminosity. Note that some values get negative which corresponds to a mass profile that increases at larger radii (see Sect. 5.2).

As one can observe, filtering out type-2 galaxies which are too close to a type-1 galaxy reduces the amplitude of the measured $\Delta\Sigma(\xi)$. The amplitude of $\Delta\Sigma(\xi)$ decreases further by filtering out also type-2 galaxies which are too close among themselves. The measured $\Delta\Sigma(\xi)$ is therefore due to mass overdensities associated with type-2 galaxies, and as shown in Fig. 7.12, there are several effects responsible for this.

Finally, we present how our measurements would be affected assuming that type-2 are realistic and that we cannot distinguish between them and type-1 galaxies. In Fig. 7.14,

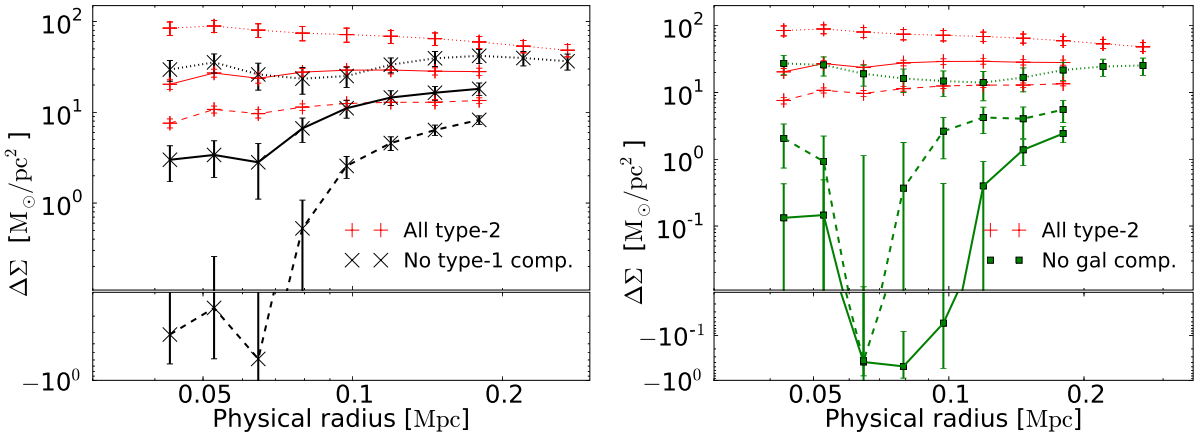

Figure 7.13: Excess surface mass density $\Delta\Sigma(\xi)$ for a few r band luminosity for a LSST-like survey (see Tab. 7.3). Here we used the projected mass maps $\Sigma(\boldsymbol{\theta})$ to obtain the measurements. With red crosses we plot the results from type-2 galaxies only. In the left panel, with black pluses, we plot the results from type-2 galaxies which are at least 0.05 Mpc away from a type-1 galaxy. In the right panel, with green squares, we plot the results from type-2 galaxies which are at least 0.05 Mpc away from a galaxy of any type. These are our results for an LSST-like. The line type denotes the different luminosity bins, in order to compare the profiles for all type-2 and for the filtered samples. Note that we plot the negative values in the lower panels.

we present $\Delta\Sigma(\xi)$ for a selection of luminosity bins for type-1 galaxies (with dashed blue lines) and the same luminosity bins but for the combination of type-2 and type-1 galaxies (with solid black lines). As one can see, the inclusion of type-2 galaxies change the measurement, especially at large radii. The change is larger for fainter galaxies. The new profiles are flatter, and there is a systematic error in the determination of the mass profile. This effect can also interfere with our ability to assign any kind of mass estimate.

In Fig. 7.15, we present the same analysis as in the top panel in Fig. 7.8, but with both type-2 and type-1 galaxies. Recall that the line and color type encodes the morphology (as a proxy for the sub-halos' age), and the symbol encodes the stellar mass (as a proxy for the sub-halo infall mass). The inclusion of type-2 galaxies is more important if we aim to estimate the mass loss on sub-halos induced by tidal stripping. Our former ability to measure this effect was small, and the change in the profiles that type-2 galaxies produce, makes it impossible.

The presence of galaxies with no halos makes our measurements less precise. However, if such galaxies exists, we could use our analysis to determine the amount of galaxies of such type, and to distinguish them from galaxies with a host halo. The methods proposed in this thesis are therefore of scientific interest in any case.

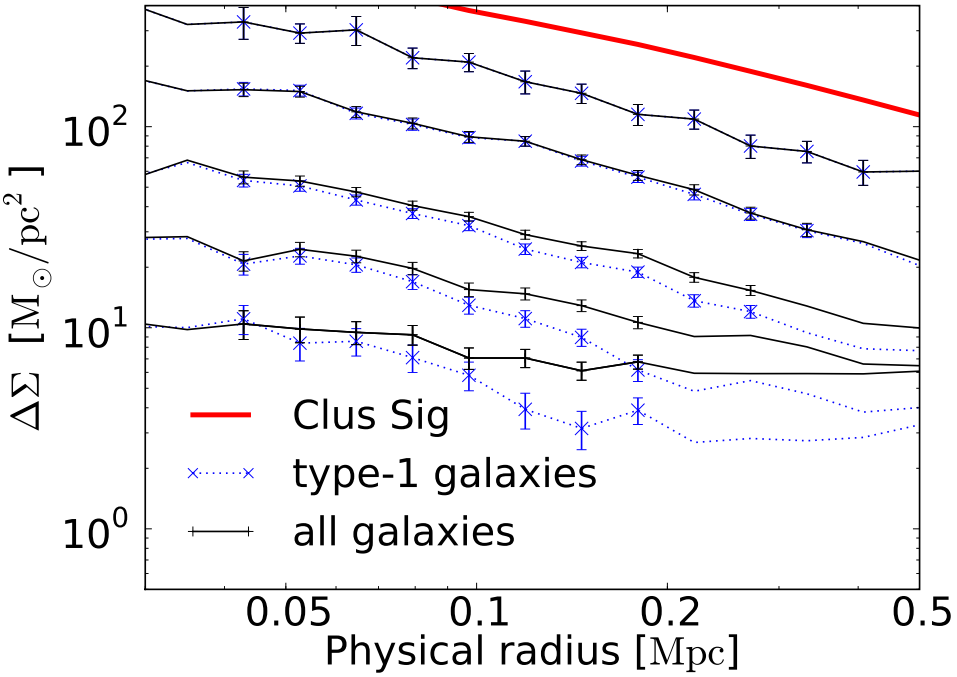

Figure 7.14: Excess surface mass density $\Delta\Sigma(\xi)$ for a few r band luminosity bins. Here we used the reduced shear catalogues $g(\boldsymbol{\theta})$ simulating a LSST-like survey, to obtain the measurements. The blue dashed lines correspond to the measurements only for type-1 galaxies. The solid black lines correspond to the combination of both type-1 and type-2 galaxies. The range where we assume the measurement is unbiased is highlighted, and it was derived using our results in Tab. 6.4 on the type-1 only profiles. The solid thick red line is the signal for the average cluster, shown for visual reference.

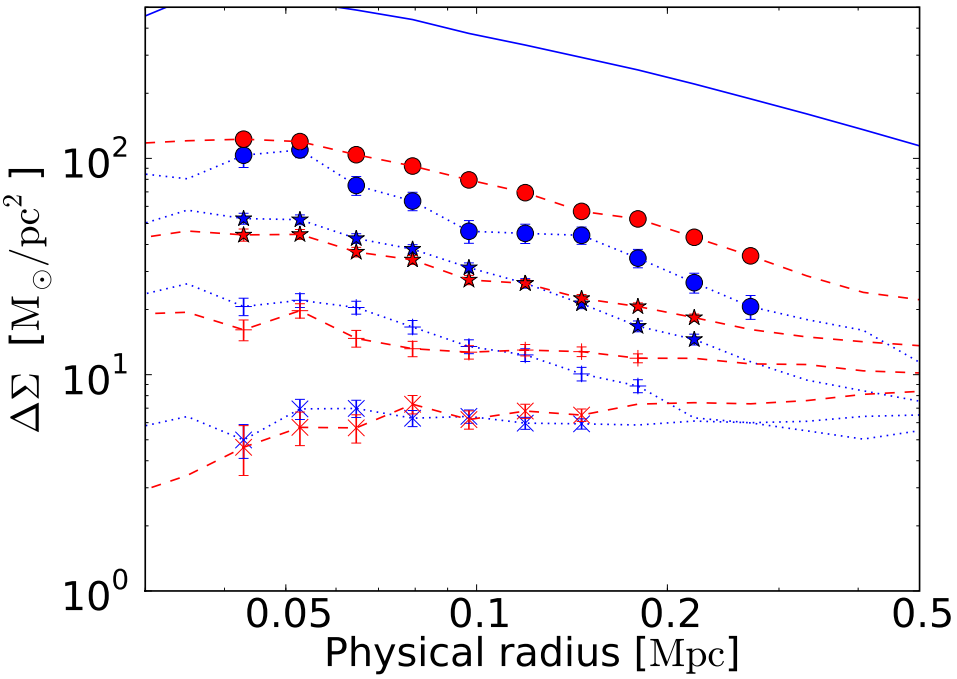

Figure 7.15: Morphology-stellar mass classification for a LSST-like survey including both type-1 and type-2 galaxies. Blue dotted lines correspond to galaxies with a relatively small bulge, red dashed to those with a large bulge. Red lines correspond to galaxies which spent more time inside a cluster. The different symbols distinguish different stellar mass bins. The range where we assume the measurement is unbiased is highlighted, and it was derived using our results in Tab. 6.4 from the type-1 only profiles. The classification ranges for morphology and stellar mass are described in Tab. 7.4 The blue solid line with the highest amplitude is the measurement for the host cluster shown for visual reference.

8. Summary & conclusions

In order to understand the evolution of galaxy clusters, it is essential to study the matter profiles of satellite galaxies. Analyzing how these profiles evolve with time can improve our knowledge on galaxy evolution and on matter structure formation.

In this work we analyzed the use of weak gravitational lensing on satellite galaxies inside clusters, in particular the use of galaxy-galaxy lensing. With this probe, we can measure in a non-parametric way the average projected mass profiles of the host matter halos of galaxies. This cosmological probe correlates the image distortion (shear) of a background galaxy with the mass of a foreground galaxy.

Galaxy-galaxy lensing needs large galaxy samples, for this reason it has not been up to now fully exploited on cluster galaxies. In our work, we forecast results for future surveys using the Millennium Simulation (Springel et al. 2005), ray-tracing simulations (Hilbert et al. 2009) and the galaxy semi-analytical catalogues by De Lucia & Blaizot (2007).

We presented the details of galaxy-galaxy lensing measurements on satellite galaxies and we exposed the contamination from the host cluster. In order to overcome the contamination from the host cluster we proposed a calibration method, which can solve the problem up to a certain range in radius. For each sub-halo, we define a point at the same distance from the main halo as the sub-halo is, but in the opposite direction as seen from the cluster center. We can estimate the contribution of the main halo measuring around this new point, under the assumption that the cluster is parity-symmetric.

We replicated the mass maps of the clusters in our simulations using theoretical profiles with known parameters. With this mock cluster sample we were able to test our calibration method taking into account realistic characteristics for our cluster samples such as halo spatial distribution or mass function. We estimated the performance of our measurements at different radii and we defined a minimal separation between the sub-halo and the main halo center to optimize the signals. Finally, we also discussed how to proceed if the central galaxy position and the main halo center are weakly correlated.

With the previous tests, we could characterize the sub-halos in the Millennium Simulation using projected mass maps. The weak lensing signal of sub-halos is well described by a simple NFW profile, and it was not possible to estimate the spatial extent of the sub-halos. Our results are consistent with an abrupt truncation of the mass profile at radii larger than 0.2 Mpc. There are certain discrepancies between our work and the previously published works from Limousin et al. (2007), Natarajan et al. (2007), Halkola et al. (2007) and Suyu & Halkola (2010). These authors measured a much smaller extent of the sub-halos using gravitational lensing on a few observed clusters. The results were derived using parametric models for the mass profiles of the sub-halos and the main halo. The models used by these authors are unable to fit sub-halos in the Millennium Simulation. Since the Millennium Simulation only contains dark matter, there is not absolute certainty that NFW profiles should be the best description for halos of real galaxies.

Nevertheless, our results challenge the choices made in the afore-mentioned works. The truncation radii that they measured can be also interpreted as a consequence of the parametric method used. Therefore, further analysis and better data are needed in order to solve the discrepancies.

We also were able to characterize the evolution of the mass profiles. We show that the lensing profiles decrease in amplitude with time. This is consistent with a mass loss at all scales. Due to the tidal forces exerted by the host halo, sub-halos are stripped of the mass at the outermost radii, and at the same time the mass at the inner regions is redistributed. This is consistent with the work by Hayashi et al. (2003) and also supports the idea of large truncation radii.

After describing the sub-halo profiles we used simulated galaxy catalogues to forecast signals for future surveys, focusing on DES (The Dark Energy Survey Collaboration 2005) and LSST (Ivezic et al. 2008). We analyzed the semi-analytical catalogues in order to classify the galaxy samples to optimize the measurements. With the result from this analysis we derived the following results.

We predict the detectability of the signals using a compact estimator of the signal-to-noise ratio. The cluster sample required is very large, but we can already expect signals from a DES-like survey roughly above the three sigma level. The data from a LSST-like survey is optimal for the studies that we proposed.

There is not a unique way of separating between sub-halo and host halo mass. In this thesis, we considered that the mass of the sub-halo is well estimated in our simulations by the SUBFIND algorithm (Springel et al. 2001) ($M_{\rm SUBF}$). We modeled the dependence between the measured NFW profiles and the gravitationally bound mass of the sub-halo mass $M_{\rm SUBF}$. We also checked that the weak lensing approximation is valid.

According to the semi-analytical catalogues luminosity in the SDSS r band is a good proxy for mass. We predict that it is possible even for a DES-like survey to constrain the mass-luminosity relations of sub-halos over two decades in mass, from around $5 \times 10^{11}\,{\rm M}_{\odot}$ to $10^{14}\,{\rm M}_{\odot}$.

In order to study in real data the time evolution of the profiles, we binned the galaxies according the mass of the sub-halo prior to falling into the cluster and the time spent inside the cluster (sub-halo age). Within the semi-analytical catalogues, it is possible to infer their initial mass, with its stellar mass. On the other hand we could not find any galaxy observable strongly correlated with the sub-halo age. With these observables, we could only put weak constraints to the evolution of sub-halo matter profiles. Nevertheless, our results are subjected to the models of galaxy evolution, which are still in constant development.

In our analyses we neglected the galaxies without a host sub-halo which one finds in the semi-analytical catalogues. A priori, these galaxies are not fully reliable as they populate mass overdensities below the resolution limit of the Millennium Simulation. Nevertheless, for completeness we considered their influence. These galaxies show a lensing signal with a high amplitude, which can be explained as being produced by correlated halos. We also were able to quantify the correlation between the position of type-2 galaxies and other sub-halos. Finally, we presented how our previous analysis is affected by assuming that these galaxies are realistic and that we are not able to distinguish them from galaxies with a host sub-halo. However, we consider that they can be safely neglected. Further simulations are needed in order to investigate galaxies of such characteristics.

A. Shear Profiles

A.1 NFW profiles

The NFW profile (Navarro et al. 1997) is a model which fits well halos in dark matter simulations regardless of the mass. It has a radial density of

$$\rho(r) = \frac{\delta_{\rm c}\,\rho_{\rm m}}{(r/r_{\rm s})(1+r/r_{\rm s})^2}, \tag{A.1}$$

where $\delta_{\rm c}$ is a characteristic number, $\rho_{\rm m}$ is the mean matter density of the universe at the halo redshift, and $r_{\rm s}$ is the *scale radius*.

In order to define the extension of the halos we define a characteristic density, which is a factor $\Delta_{\rm vir}$ times the comoving mean density of the Universe $\rho_{\rm m}$. However there are other definitions based on the critical density or the physical mean density of the Universe. The factor $\Delta_{\rm vir}$ is derived from the spherical collapse model and it is an approximation to the density of a virialized object. The radius inside which the average density of an object equals $\Delta_{\rm vir}\rho_{\rm m}$ is called the *virial radius* $r_{\rm vir}$. Since the extension so defined is only an estimation of the virial radius, it is also customary to use the notation $r_{\Delta_{\rm vir}}$ instead of $r_{\rm vir}$. In this thesis since we assume $\Delta_{\rm vir} = 200$, we shall use r_{200}.

The mass within the radius r is given by:

$$M(<r) = 4\pi\delta_{\rm c}\rho_{\rm m}r_{\rm s}^3\left[\ln\left(1+\frac{r}{r_{\rm s}}\right) - \frac{1}{1+r_{\rm s}/r}\right]. \tag{A.2}$$

A commonly used parameter for NFW profiles is the ratio between the virial radius and the scale radius, the concentration:

$$c = \frac{r_{200}}{r_{\rm s}}, \tag{A.3}$$

which is related to the characteristic density $\delta_{\rm c}$

$$\delta_{\rm c} = \frac{\Delta_{\rm vir}}{3}\frac{c^3}{\ln(1+c) - c/(1+c)}. \tag{A.4}$$

The projected mass density $\Sigma_{\rm NFW}$ at a projected radius ξ of an NFW profile has been derived by Bartelmann (1996). If we define the parameter $x = \xi/r_{\rm s}$, the expression for $\Sigma_{\rm NFW}$ is:

$$\Sigma_{\rm NFW}(x) = \begin{cases} \dfrac{2\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}}{x^2-1}\left(1 - \dfrac{2}{\sqrt{1-x^2}}\operatorname{arctanh}\sqrt{\dfrac{1-x}{1+x}}\right) & \text{for} \quad x<1, \\ \dfrac{2\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}}{3} & \text{for} \quad x=1, \\ \dfrac{2\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}}{x^2-1}\left(1 - \dfrac{2}{\sqrt{x^2-1}}\arctan\sqrt{\dfrac{x-1}{1+x}}\right) & \text{for} \quad x>1. \end{cases} \tag{A.5}$$

This expression can be integrated to obtain the mean projected mass density $\Delta\Sigma$ inside the radius ξ, which yields:

$$\bar{\Sigma}_{\rm NFW}(x) = \begin{cases} \dfrac{4\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}}{x^2}\left(\dfrac{2}{\sqrt{1-x^2}}\,\mathrm{arctanh}\sqrt{\dfrac{1-x}{1+x}} + \ln(x/2)\right) & \text{for} \quad x<1, \\ 4\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}\,[1+\ln(1/2)] & \text{for} \quad x=1, \\ \dfrac{4\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}}{x^2}\left(\dfrac{2}{\sqrt{x^2-1}}\,\arctan\sqrt{\dfrac{x-1}{1+x}} + \ln(x/2)\right) & \text{for} \quad x>1. \end{cases} \tag{A.6}$$

The excess surface mass density comes straightforward as

$$\Delta\Sigma_{\rm NFW}(x) = \bar{\Sigma}_{\rm NFW}(x) - \Sigma_{\rm NFW}(x). \tag{A.7}$$

A.2 Truncated NFW profiles

The truncated NFW profile we use in this thesis is described by Baltz et al. (2009). It is a modification of the original and reads

$$\rho(r) = \frac{\delta_{\rm c}\,\rho_{\rm m}}{(r/r_{\rm s})(1+r/r_{\rm s})^2\,[1+(r/r_{\rm tr})^2]}. \tag{A.8}$$

In order to define the lensing profiles we shall define again $x = \xi/r_{\rm s}$, and the function

$$F(x) = \begin{cases} \dfrac{\ln\left(1/x - \sqrt{1/x^2-1}\right)}{\sqrt{1-x^2}} & \text{for} \quad x<1, \\ 1 & \text{for} \quad x=1, \\ \dfrac{\arccos(1/x)}{\sqrt{x^2-1}} & \text{for} \quad x>1. \end{cases}$$

With this function and defining $\tau = r_{\rm tr}/r_{\rm s}$ for convenience, we have

$$\Sigma(x) = 2\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}\frac{\tau^2}{(\tau^2+1)^2}\left(\frac{\tau^2+1}{x^2-1}[1-F(x)] + 2\,F(x) \right. \\ \left. -\frac{\pi}{\sqrt{\tau^2+x^2}} + \frac{\tau^2-1}{\tau\sqrt{\tau^2+x^2}}\log\left(\frac{x}{\sqrt{\tau^2+x^2}+\tau}\right)\right). \tag{A.9}$$

The mean density within a radius x is

$$\bar{\Sigma}(x) = \frac{4\,r_{\rm s}\,\delta_{\rm c}\,\rho_{\rm m}}{x^2}\frac{\tau^2}{(\tau^2+1)^2}\left(\left[\tau^2+1+2(x^2-1)\right]F(x) + \tau\,\pi + (\tau^2-1)\log\tau \right. \\ \left. +\sqrt{\tau^2+x^2}\left[-\pi + \frac{\tau^2-1}{\tau}\log\left(\frac{x}{\sqrt{\tau^2+x^2}+\tau}\right)\right]\right). \tag{A.10}$$

A.3 PIEMD profiles

The pseudo-isothermal elliptical mass distribution (PIEMD), was derived by Kassiola & Kovner (1993). The profile we use is however a version of the former, the truncated PIEMD or the dual pseudo isothermal elliptical mass distribution (dPIE) (Elíasdóttir et al. 2007). nevertheless for this work we consider only the circular case. Its three dimensional distribution reads :

$$\rho(r) = \frac{\rho_0}{(1 + r^2/r_{\rm core}^2)(1 + r^2/r_{\rm cut}^2)}. \tag{A.11}$$

The correspondent surface mass density distribution is

$$\Sigma(\xi) = \frac{\rho_0\, r_{\rm core}^2\, r_{\rm cut}^2\, \pi}{r_{\rm cut}^2 - r_{\rm core}^2} \left(\frac{1}{\sqrt{r_{\rm core}^2 + \xi^2}} - \frac{1}{\sqrt{r_{\rm cut}^2 + \xi^2}} \right). \tag{A.12}$$

The mean surface mass density distribution is

$$\bar{\Sigma}(\xi) = \frac{2\,\rho_0\, r_{\rm core}^2\, r_{\rm cut}^2\, \pi}{\xi^2 (r_{\rm cut} + r_{\rm core})} \left(1 - \frac{\sqrt{r_{\rm cut}^2 + \xi^2} - \sqrt{r_{\rm core}^2 + \xi^2}}{(r_{\rm cut} - r_{\rm core})} \right). \tag{A.13}$$

The parameter $r_{\rm cut}$ changes the profile such as the total mass does not diverges, in the same way as the truncation radius in the NFW profile. For this reason we shall consider it also as a truncation scale.

B. Halo shear signal around any point

We want to describe the shear signal for a circular symmetric halo, seen from a point different from the halo center. We present a sketch in Fig. B.1.

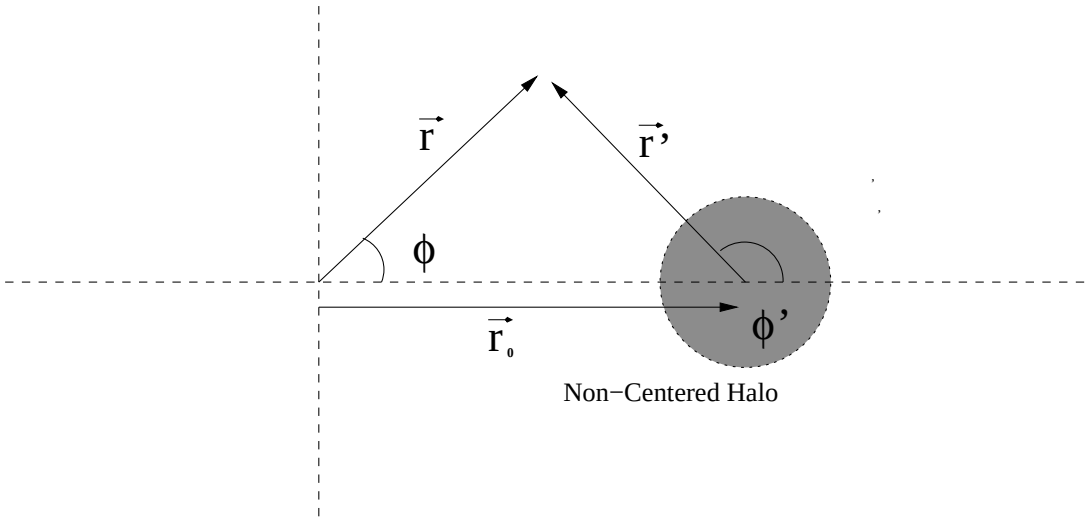

Figure B.1: Sketch of the situation we want to treat. The origin of coordinates, (also the point around which we want to measure shear), is where the dashed lines cross.

Because of the symmetry, the modulus of the shear only depends on the distance r' to the center. If the modulus is described by the function $f(r')$, the shear produced at a position $\mathbf{r}' = (r', \phi')$ is:

$$\gamma(\mathbf{r}') = f(r') \exp\left(2\mathrm{i}\left[\phi' + \frac{\pi}{2}\right]\right) = -f(r') \exp(2\mathrm{i}\phi'). \tag{B.1}$$

The additional $\pi/2$ phase on the exponential accounts for the fact that the shear aligns perpendicular to the position vector.

Now we define the tangential shear with respect any origin, taking an offset frame of reference ($\mathbf{r}$) with respect the one centered on the halo ($\mathbf{r}'$). The tangential shear around the new origin is

$$\gamma_{\mathrm{t}}(\mathbf{r}) = -\gamma_1\big(\mathbf{r}\big)\cos(2\phi) - \gamma_2\big(\mathbf{r}\big)\sin(2\phi) \tag{B.2}$$

In the assumed case where the shear field is only produced by the halo, we can compute from Eq. (B.1) the Cartesian components of the shear field $\gamma_1(\mathbf{r}') + \mathrm{i}\gamma_2(\mathbf{r}')$, in the old frame of reference ($\mathbf{r}'$) and express them into the new one ($\mathbf{r}$):

$$\gamma_{\rm t}(\mathbf{r}) = -\gamma_1\big(\mathbf{r}'(\mathbf{r})\big)\cos(2\phi) - \gamma_2\big(\mathbf{r}'(\mathbf{r})\big)\sin(2\phi)$$

$$= f\big(\mathbf{r}'(\mathbf{r})\big)\Big[\cos\big(2\phi'(\mathbf{r})\big)\cos(2\phi) + \sin\big(2\phi'(\mathbf{r})\big)\sin(2\phi)\Big]. \tag{B.3}$$

From Fig. B.1 we can derive the necessary trigonometric relations between both coordinate frames:

$$\left.\begin{array}{l} r'\cos\phi' = r\cos\phi - r_0 \\ \\ r'\sin\phi' = r\sin\phi \end{array}\right\} \Longrightarrow \begin{cases} \cos\phi' = \dfrac{r\cos\phi - r_o}{r'}, \\ \\ \sin\phi' = \dfrac{r\sin\phi}{r'}, \\ \\ r'^2 = r^2 + r_0^2 - 2r_0 r\cos\phi. \end{cases} \tag{B.4}$$

Now we apply these trigonometric equalities and express $r' \to r,\ \theta' \to \theta$:

$$\cos(2\phi') = \cos^2(\phi') - \sin^2(\phi') = \frac{r^2\cos(2\phi) + r_0^2 - 2r_0 r\cos(\phi)}{r^2 + r_0^2 - 2r_0 r\cos(\phi)}, \tag{B.5}$$

$$\sin(2\phi') = 2\sin(\phi')\cos(\phi') = \frac{r^2\sin(2\phi) - 2r_0 r\sin(\phi)}{r^2 + r_0^2 - 2r_0 r\cos(\phi)}. \tag{B.6}$$

Putting altogether we arrive to the final expression of the tangential shear around any point:

$$\gamma_{\rm t}(\mathbf{r}) = \frac{f\left(\sqrt{r^2 + r_0^2 - 2r_0 r\cos(\phi)}\right)}{r^2 + r_0^2 - 2r_0 r\cos(\phi)} \times \Bigg(\cos(2\phi)\Big[r^2\cos(2\phi) + r_0^2 - 2r_0 r\cos(\phi)\Big] + \sin(2\phi)\Big[r^2\sin(2\phi) - 2r_0 r\sin(\phi)\Big]\Bigg),$$

$$\gamma_{\rm t}(\mathbf{r}) = f\left(\sqrt{r^2 + r_0^2 - 2r_0 r\cos(\phi)}\right)\frac{r^2 + r_0^2\cos(2\phi) - 2r_0 r\cos(\phi)}{r^2 + r_0^2 - 2r_0 r\cos(\phi)},$$

$$\gamma_{\rm t}(\mathbf{r}) = f\left(\sqrt{r^2 + r_0^2 - 2r_0 r\cos(\phi)}\right)\left[1 + \frac{r_0^2[\cos(2\phi) - 1]}{r^2 + r_0^2 - 2r_0 r\cos(\phi)}\right]. \tag{B.7}$$

C. Bayes' theorem and model comparison

The basis of Bayesian analysis is Bayes' theorem. The theorem relates the conditional probability of an hypothesis H or model, given a data set D, $p(H|D)$, with the conditional probability of the data given the hypothesis $p(D|H)$. This allows to discriminate between hypothesis through a much easier analysis of likelihood of the data. The theorem reads:

$$p(H|D)\,p(D) = p(D|H)\,p(H). \tag{C.1}$$

where $p(D)$ is the likelihood of the data or *evidence* and $p(H)$ the likelihood of the hypothesis or *prior*. The prior encodes our previous knowledge.

From Bayes' theorem we can derive what model is favored by the data. The ratio between $p(H|D)$ of two different hypothesis (H_1 and H_2) yields

$$\frac{p(H_1|D)}{p(H_2|D)} = \frac{p(D|H_1)}{p(D|H_2)}\frac{p(H_1)}{p(H_2)}. \tag{C.2}$$

The l.h.s. of the equation is the posterior ratio of the probabilities of both hypothesis. The second term in the r.h.s. of the equation is the prior ratio, and the remaining term is the Bayes' factor. The Bayes' factor relates the prior to the posterior, therefore it carries the information given by the data, concerning the validity of both hypothesis.

The Bayes' factor can be compute from the likelihood of the data:

$$B_{12} = \frac{p(D|H_1)}{p(D|H_2)} = \frac{\int p(D|\theta_1, H_1)\,\pi(\theta_1|H_1)\,\mathrm{d}\theta_1}{\int p(D|\theta_2, H_2)\,\pi(\theta_2|H_2)\,\mathrm{d}\theta_2}, \tag{C.3}$$

where θ_1, θ_2 are the set of parameters for the model given by the hypothesis H_1, H_2 and $\pi(\theta_1|H_1)$, $\pi(\theta|H_2)$ are the respective prior densities of the parameters. One can develop empirically a scale for its interpretation. In Tab. C.1 we present the version by Kass & Raftery (1995).

Table C.1: Scale for rating a the likelihood of hypothesis H_1 against H_2 according to its evidence Kass & Raftery (1995).

Bayes Factor B_{12}	Evidence against H_2
<1	Disfavored against H_1
1:3	Barely significant
3:20	Positive
20:150	Strong
>150	Very strong

D. Nested Sampling

Nested sampling is an algorithm designed to compute the Bayesian evidence. With it, it is possible to rank the performance of different models. The Bayesian posterior is also computed as a secondary product. This is only an overview, for a detailed description see Skilling (2004). The essence of the algorithm is to use Monte-Carlo methods to compute the integral expression of the evidence. The sampling is design to converge geometrically towards the point with maximum likelihood, and solve the multidimensional curse.

The evidence is defined as

$$Z = \int L(D|\theta, H)\, \pi(\theta|H)\, \mathrm{d}\theta, \tag{D.1}$$

where $L(D|\theta, H)$ is the likelihood of the data D, assuming the hypothesis H. We also have assumed a particular parameterization of the problem θ. With $\pi(\theta|H)$ we denote the prior over the parameters' space. The first step is to define the prior mass

$$X(\lambda) = \int_{L>\lambda} \pi(\theta|H)\, \mathrm{d}\theta, \tag{D.2}$$

which accounts for the parameters' space with a likelihood greater than λ. Expressing the likelihood L as a function of the prior mass $L(X)$, and normalizing the total prior mass such as $X(0) = 1$ and $X(L_{\max}) = 0$, the evidence becomes a one-dimensional integral over the unit range:

$$Z = \int_0^1 L(X)\mathrm{d}X. \tag{D.3}$$

The numerical computation of this integral involves dividing the total prior mass and sorting the resulting elements according to its likelihood. One complication that may occur is ties. For practical purposes, we break the degeneracy by adding a very small random number to each point with the same likelihood.

In order to understand the algorithm, we discuss the first iteration in the central loop. We start by sampling uniformly the weighted parameters' space with N points. We denote the points like t_i, and we order them according to their likelihood. We can statistically estimate the prior mass for the smallest likelihood sampled $L(t_N)$ from the N points. In a random sampling of the unit interval, the probability distribution function (PDF) of the maximum value X follows

$$\mathrm{PDF}(X) = N\, X^{N-1} \quad \text{in} \quad (0, 1), \tag{D.4}$$

where N is the number of random points in the sample. The prior mass defined by the point with the smallest likelihood $X(t_N)$ follows this PDF. For an elegant derivation of the final result, we compute the mean and the standard deviation of $\ln(X)$ of this probability distribution function:

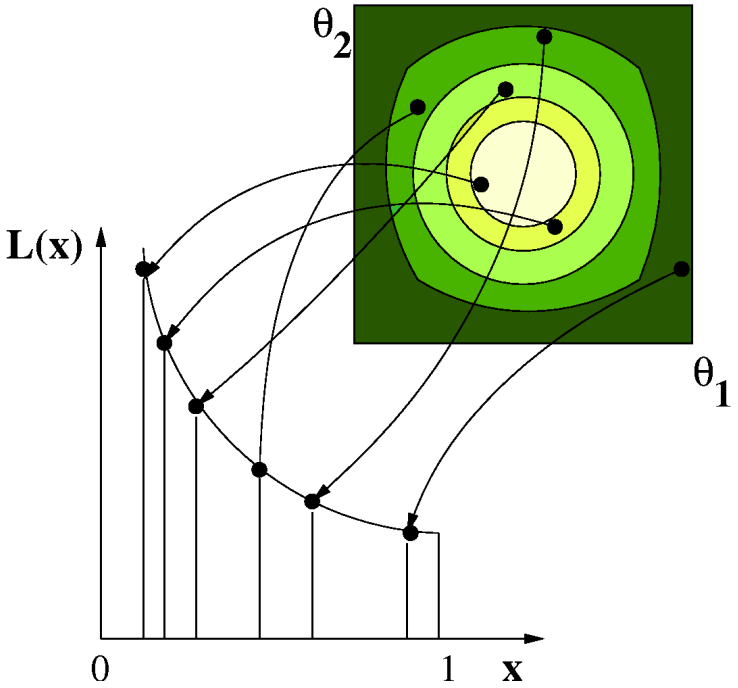

Figure D.1: Illustration of the algorithm. The square with different contour lines represents our parameter space with the likelihood coded into the color. The different points form a sample with which to compute the evidence following the algorithm. Original picture extracted from (Mukherjee et al. 2006).

$$\big\langle \ln(X)\big\rangle = -1/N, \quad \sigma\big(\ln(X)\big) = 1/N. \tag{D.5}$$

The average contribution to the total evidence for the prior space between $L = 0$ and $L(t_N)$ is $L(t_N) \times \big[1 - \exp(-1/N)\big]$. This is a coarse estimate of this particular contribution of the evidence. On the other hand, the evidence fraction contained in this parameter volume fraction is small for a sufficient number of sampling points, and the error introduced small.

In the next iteration of the central loop of the algorithm, we repeat the process re-sampling the weighted parameters' space between $L_{\max}$ and $L(t_{\rm N})$ with N more points. Now, we estimate the average contribution to the total evidence between $L(t_{\rm N})$ and the next smallest likelihood value.

The sampling has a higher density as we probe regions with a larger likelihood. After j iterations, which we construct such that they are independent, the prior mass is expected to shrink to:

$$\ln(X_j) \approx -\Big(j \pm \sqrt{j}\Big)/N \quad \text{such as} \quad X_j = \exp(-j/N). \tag{D.6}$$

We can therefore assign in each iteration of our loop, the prior mass fraction w_j, for the value with the smallest likelihood $L(t_j)$

$$w_j = X_{j-1} - X_j. \tag{D.7}$$

Our final estimation of the Bayesian evidence after m iterations is therefore:

$$Z(t) = \sum_{j=1}^{m} L(t_j) w_j = \sum_{j=1}^{m} L(t_j)\Big(\exp(-j/N) - \exp(-[j-1]/N)\Big). \tag{D.8}$$

We denote with $L(t_j)$ the lowest likelihood sampled in each iteration. The algorithm can be terminated when the last iteration does not increases the evidence by more than a small fraction.

The sample of stored points at each iteration can be used to reconstruct the posterior. Through Bayes' theorem, the posterior is:

$$P(\theta, H|D) = \frac{L(D|\theta, H)\pi(\theta|H)}{Z(D)}. \tag{D.9}$$

In order to computed the expectation value of the quantity $Q(\theta, H|D)$ from the posterior we only need to compute:

$$E(Q) = \frac{\sum_{j=1}^{m} Q(\theta, H|D)\, L\big(t_j(\theta, H|D)\big) w_j}{\sum_{j=1}^{m} L(t_j) w_j}. \tag{D.10}$$

The idea can be solved implementing the following loop:

1. Sample uniformly the parameters' space with N points t_x, accepting only those whose likelihood is larger than L_{lim} which we initialize to 0.
2. Store the point t_j with the lowest likelihood $L(t_j)$.
3. Estimate the prior mass fraction $\Delta X = w_j$ between L_{lim} and $L(t_j)$, which the former sampling defines.
4. Compute the contribution to the total evidence of the current iteration, approximated by $L(t_j)w_j$.
5. Update the likelihood acceptance limit to the stored value $L_{\text{lim}} = L(t_j)$.
6. Iterate until the integral converges.

There is actually no need to produce N points in each iteration. We can reuse the $N-1$ points from the previous step with a larger likelihood than L_j, improving the performance of the algorithm.

Acknowledgments

The first time I desired to do scientific investigation I was a 12 years old boy. I must thank Professor Peter Schneider for giving me the opportunity to do a Ph.D. and for taking care of me.

This thesis cost me a piece of my life: it cost me time, energy and pain. I benefited from the help of many people. But the main reason for which I could bring it to an end was the support of two people: Dr. Stephan Hilbert and Dr. Jan Hartlap. My work could not have been done without them. I have a debt with them and here I leave written proof that I will make it up to them.

One of the hardest thing I had to do to finish this thesis was to believe I could do it. I would not have done it without my friends. I believe I would have got lost without them. I would like to mention, for their special support, Nina Roth, Tim Eifler, and Amru Hussein. They were next to me on very special occasions (and they even read chapters of my thesis). Other people who were present along these year were Mauricio and Cristina, Emilio Biancardi and Guillem Sagué. My parents to whom I dedicate this thesis also belong here.

I also would like to mention the people that invested his own time to read and correct the thesis. Apart from P. Schneider, J. Harlap and S. Hilbert (who read almost all of it) I want to thank Phil Bett, and Cristiano for reading my thesis. I also want to thank Sherry Suyu, Mischa Schirmer, Holger Israel and Jackeline Chen, whom I constantly consulted and were always ready to help.

I am going to miss the members of my AIfA group, those who already went away some time ago, and those who I leave behind. I shall carry them in my heart forever.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Alcock, C., Allsman, R. A., Alves, D. R., et al. 2000, ApJ, 542, 281

Baldi, M., Pettorino, V., Robbers, G., & Springel, V. 2010, MNRAS, 403, 1684

Baltz, E. A., Marshall, P., & Oguri, M. 2009, Journal of Cosmology and Astro-Particle Physics, 1, 15

Bartelmann, M. 1996, A&A, 313, 697

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Bennett, C. L., Bay, M., Halpern, M., et al. 2003, ApJ, 583, 1

Bertschinger, E. 2001, astro-ph/0101.009

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, MNRAS, 398, 1150

Carlstrom, J. E., Ade, P. A. R., Aird, K. A., et al. 2009, astro-ph/0907.4445

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

De Lucia, G. & Blaizot, J. 2007, MNRAS, 375, 2

De Lucia, G., Kauffmann, G., & White, S. D. M. 2004, MNRAS, 349, 1101

Dodelson, S. 2003, Modern cosmology (Amsterdam (Netherlands): Academic Press. 2003)

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

Elíasdóttir, Á., Limousin, M., Richard, J., et al. 2007, astro-ph/0710.5636

Etherington, I. M. H. 1933, Philosophical Magazine, 15, 761

Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, ApJ, 553, 47

Halkola, A., Seitz, S., & Pannella, M. 2007, ApJ, 656, 739

Hartlap, J. 2005, Studying Galaxy-Galaxy-Lensing Using Ray-Tracing Simulations (Rheinischen Friedrich-Wilhelms-Universität Bonn. Institut für Astrophysik und Extraterrestrische Forschung, (IAEF))

Hayashi, E., Navarro, J. F., Taylor, J. E., Stadel, J., & Quinn, T. 2003, ApJ, 584, 541

Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, A&A, 499, 31

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2004, ApJ, 606, 67

Ivezic, Z., Tyson, J. A., Allsman, R., et al. 2008, astro-ph/0805.2366

Kaiser, N., Burgett, W., Chambers, K., et al. 2010, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 7733, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Kaiser, N. & Squires, G. 1993, ApJ, 404, 441

Kass, R. E. & Raftery, A. E. 1995, Journal of the American Statistical Association, 90, 773

Kassiola, A. & Kovner, I. 1993, ApJ, 417, 450

Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2010, astro-ph/1001.4538

Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, ApJ, 609, 482

Limousin, M., Kneib, J., & Natarajan, P. 2005, MNRAS, 356, 309

Limousin, M., Kneib, J. P., Bardeau, S., et al. 2007, A&A, 461, 881

Linde, A. 2005, Journal of Physics Conference Series, 24, 151

Mandelbaum, R., Seljak, U., & Hirata, C. M. 2008, J. Cosmology Astropart. Phys., 8, 6

Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, MNRAS, 368, 715

Meylan, G., Jetzer, P., North, P., et al., eds. 2006, Gravitational Lensing: Strong, Weak and Micro (Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro)

Misner, C. W., Thorne, K. S., & Wheeler, J. A. 1973, Gravitation (San Francisco: W.H. Freeman and Co., 1973)

Mukherjee, P., Parkinson, D., & Liddle, A. R. 2006, astro-ph/0508461v2

Muldrew, S. I., Pearce, F. R., & Power, C. 2010, astro-ph/1008.2903

Natarajan, P., De Lucia, G., & Springel, V. 2007, MNRAS, 376, 180

Natarajan, P. & Kneib, J. 1997, MNRAS, 287, 833

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Navarro, J. F., Hayashi, E., Power, C., et al. 2004, MNRAS, 349, 1039

Park, C. & Choi, Y. 2005, ApJ, 635, L29

Parker, L. C., Hoekstra, H., Hudson, M. J., van Waerbeke, L., & Mellier, Y. 2007, ApJ, 669, 21

Particle Data Group, Amsler, C., Doser, M., et al. 2008, Physics Letters B, 667, 1

Peacock, J. A. 1999, Cosmological Physics (Cambridge University Press, January 1999)

Percival, W. J., Reid, B. A., Eisenstein, D. J., et al. 2010, MNRAS, 401, 2148

Predehl, P., Boehringer, H., Brunner, H., et al. 2010, astro-ph/1001.2502

Refregier, A., Amara, A., Kitching, T. D., et al. 2010, astro-ph/1001.0061

Renault, C., Afonso, C., Aubourg, E., et al. 1997, A&A, 324, L69

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Riess, A. G., Macri, L., Casertano, S., et al. 2009, ApJ, 699, 539

Schmidt, B. P., Suntzeff, N. B., Phillips, M. M., et al. 1998, ApJ, 507, 46

Schneider, P. 2006a, Extragalactic Astronomy and Cosmology (Berlin: Springer, 2006.)

Schneider, P. 2006b, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 269–451

Schneider, P. & Rix, H. 1997, ApJ, 474, 25

Seitz, C. & Schneider, P. 1997, A&A, 318, 687

Seljak, U. & Zaldarriaga, M. 1996, ApJ, 469, 437

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, American Institute of Physics Conference Series, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Suyu, S. H. & Halkola, A. 2010, astro-ph/1007.4815

The Dark Energy Survey Collaboration. 2005, astro-ph/0510346

Tian, L., Hoekstra, H., & Zhao, H. 2009, MNRAS, 393, 885

Tisserand, P., Le Guillou, L., Afonso, C., et al. 2007, A&A, 469, 387

Tytler, D., O'Meara, J. M., Suzuki, N., & Lubin, D. 2000, Phys. Rep., 333, 409

Wang, L., Li, C., Kauffmann, G., & De Lucia, G. 2006, MNRAS, 371, 537

Weinberg, S. 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity (Wiley-VCH,1972)

White, S. D. M., Davis, M., & Frenk, C. S. 1984, MNRAS, 209, 27P

White, S. D. M., Frenk, C. S., & Davis, M. 1983, ApJ, 274, L1