André Zamorano Vitorelli

Lentes Gravitacionais Fracas: Técnicas de Medida & Seu Uso Como Sistema de Datação Cósmica

André Zamorano Vitorelli

Lentes Gravitacionais Fracas: Técnicas de Medida & Seu Uso Como Sistema de Datação Cósmica

Tese apresentada ao Departamento de Astronomia do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo como requisito parcial para obtenção do título de Doutor em Ciências.

Área de concentração: Astronomia.
Orientador: Prof. Dr. Eduardo S. Cypriano

Versão Corrigida. O original encontra-se disponível na Unidade.
São Paulo
2019

Resumo

Neste trabalho, discutimos o uso de lentes gravitacionais fracas no contexto de formação da estrutura em larga escala. Usando lentes em correlação cruzada de sistemas empilhados do CFHT Stripe 82 Lensing Survey (CS82), medimos massas e concentrações médias para os sistemas empilhados, ranqueados em tercis por intervalos de magnitude entre a galáxia central e a galáxia satélite mais brilhante, na banda r. Também dividimos as pilhas pelo deslocamento entre a galáxia central e o centróide de luminosidade das galáxias satélites. Encontramos que tanto altos intervalos de magnitude como baixos deslocamentos correspondem a concentrações maiores o que, por sua vez, são indicadoras de histórias de acreção de massa antigas. Adicionalmente, exploramos técnicas de medida de forma de galáxia para o futuro levantamento J-PAS, criando protocolos de medidas de forma que nos permitiram expandir nosso trabalho de astrofísica de aglomerados para novos dados.

Abstract

In this work, we discuss the use of weak gravitational lensing to probe the formation history of the large-scale structure of the Universe. By using cross-correlation lensing of stacked galaxy systems from the CFHT-Stripe 82 Lensing Survey (CS82), we calculate posterior distributions of masses and concentrations of ensembles (stacks) of galaxy systems, ranked into terciles by the magnitude gap in the r-band between the central galaxy and the brightest satellite galaxy. We also divide stacks by offsets between the central galaxy and the luminosity centroid of the satellites. We have found that higher both magnitude gaps and lower offsets correspond to higher concentrations, which in turn are indicators of early mass accretion histories. Additionally, we explore shear measurement techniques for the upcoming J-PAS, from adaptive moments to forward fitting and convolutional neural networks. This, in turn, is a preparation for building repeatable, easily usable, shape measurement protocols that will enable us to expand our cluster astrophysics work into future data.

この道や行く人なしに秋の暮

―松尾芭蕉

ANDRÉ ZAMORANO VITORELLI

WEAK GRAVITATIONAL LENSING: MEASUREMENT TECHNIQUES & THE USE AS A COSMIC DATING SYSTEM

Contents

List of Figures

List of Tables

Acknowledgements

To my mother and brother, who took part in this entire journey to now,

To my father, who first inspired me to look to the stars,

To my beloved partner, Nathalia, for all the love and care and company and inspiration,

To my friends, Caroline, Ana Rita, Luana, mwiimr, Jair, Juliana, Michelle, Adriana, Bárbara, Lídia, Venom, Arthur, Rick, Nexus, Juca, Sher, Priscila, Layla, Andressa, Suzana, Lise, Camilla, Joyce, Catarina, Geeh, Mandi, Gustavo, Vitor, Zé, Adriane and so many others for making my life an endless stream of fruitful discoveries and experiences,

To my coworker-friends Mirian, Hélio, Nathália, Rogério, Raniere, Henrique, Pablo, Stephane, Fábio, Ivan, Johnny, Rafael for their company, counsel, care, and providing for a strong scientific environment.

To professor Thiago Martins, for all the help, without which this work could not have been done, To my advisor Prof. E. S. Cypriano, for the patience and guidance in these years,

To the professors I had the joy to learn from, both at IAG at elsewhere,

To the Brazilian people, who, hereby represented by CAPES and CNPq, entrusted me "to forward the search into the mysteries and marvellous simplicities of this strange and beautiful Universe, our home." -[Misner et al., 1973]

Part I

Theory

1

Introduction

THE STUDY OF SYSTEMS OF GALAXIES, from the small, tens-of-galaxies groups, to the largest structures in the universe that we can identify (galaxy clusters and superclusters), has become one of the main fields of astrophysics. In cosmology, galaxy systems play a fundamental role, because both the content of the universe and gravity determine everything, from their shapes, to statistical properties of their population. On the subject of galaxy formation, they can provide insight of what might create the types of galaxies we observe [Dressler, 1984]. Furthermore, as dynamical structures made mostly of Dark Matter (DM) [Zwicky, 1933], they can be used as "natural particle accelerators"[1] - as cluster mergers can test possible dark matter self-interaction properties [Kahlhoefer et al., 2014]. Finally, as gravitational lenses, they can become natural telescopes to probe high redshift galaxies that would have remained unseen with current technology [eg. Zitrin et al., 2015, Tanvir et al., 2009].

In this work we will explore the physics of formation of galaxy clusters using weak gravitational lensing as a probe of mass distribution. Our objective is to investigate how several observables of these systems correlate with the shape of the mass distribution.

This investigation will focus on two aspects: the techniques by which we can measure the lensing signal, and the application of that measurement in probing galaxy clusters. The technique we use is called weak gravitational lensing, and concerns small shape distortion of background galaxies due to the gravitational effect of galaxy clusters on their foregrounds. The application will be to measure galaxy system masses and the shape of their mass

[1] More generally, natural laboratories as dark matter tests are not confined to the realm of self-interactions.

distribution, projected on the plane of the sky.

The role of this work is twofold: as a doctoral thesis, it summarises the work on the project and its results. Additionally, as a guide for future students, we will try to present the contents in a useful, mostly self-contained manner without turning it into a full scale review of the subjects presented, which would diminish the focus of the work.

In the following sections, we establish a stage in which the evolution of galaxies and galaxy systems take place. The second chapter will deal with the fundamental aspects of the evolution, dynamical properties and observables of galaxy clusters. The third will discuss the main technique used, which is Weak Gravitational Lensing. In chapters chapter 4 & chapter 5 we present our observational studies and their current results. Furthermore, we present in chapter 6 the work to build a new pipeline for future surveys using weak gravitational lensing. Finally, a perspective outlook is offered in the final chapter.

1.1 *The Cosmological Scenario*

PHYSICAL COSMOLOGY is the description of the formation of the large scale structures we observe in the universe through physical principles. It is founded on two cornerstones: a theory of gravity, which is the dominant force in large scales[2], and a model that tells us on what gravity acts to shape the observable universe.

As of this writing, more than a hundred years have passed since the discovery of the theoretical basis of modern cosmology, which is Einstein's Theory of Gravity (ETG) [3]. Tests of ETG, from its effects on time and the bending of light, to parallel transport of angular momentum [Everitt et al., 2011, Will, 2014], and gravitational wave effects [Hulse and Taylor, 1975, LIGO Scientific Collaboration and Virgo Collaboration, 2016], have been extensively carried with continuous success.

Along the establishment of the modern theory of gravity, the field of physical cosmology developed with observation of cosmic expansion, [Slipher, 1917, Hubble, 1929], then with theoretical arguments by Lemaître [1927], Friedmann [1922]. In the following

[2] Gravity and Electromagnetism have infinite range, scaling with r^{-2}, whereas other known forces have very limited range. Electromagnetism, despite being the stronger of the two, is only relevant where charge distribution is not homogeneous. The aspect of homogeneity will be more thoroughly discussed later.

[3] To understand the reluctance of the author in calling the theory as *"General Relativity"*, please refer to the introduction chapter of [Fock, 1955/2015]

decades, CMB observations by Penzias and Wilson [1965] and firm theoretical grounds by Hawking and Ellis [1973] laid a solid foundation for the development of the current standard model. It was only in the past two decades, however, that cosmology has reached its current form - called by some the "golden age of cosmology". After the discovery of the accelerated cosmic expansion by Perlmutter et al. [1999], Riess et al. [1998], new, large surveys (eg. SDSS) to tackle the fundamental questions of the formation of the universe.

ETG is a *metric theory of gravity*, that is, one in which the effects of gravitation are explained as geometric properties of space-time. The main *heuristic* argument (but not fundamental property, again see [Fock, 1955/2015]) of the theory is that a suitable change in geometry can account for gravitational acceleration: free-falling astronauts inside a space capsule, for example, cannot locally distinguish their accelerated movement in a gravitational field or inertial movement in empty space, except for minute tidal forces.

Changes in geometry can be represented by changes in the measurement of distances in small $\mathrm{d}s$ lengths or, more generally, by observing the change in *parallel transport*: in Euclidean space, continuously shifting a vector along a closed trajectory, while maintaining its angle respective to the velocity at each point will bring the vector back to itself, regardless of the trajectory. The same is not true if the space has *curvature*

By requiring that the theory is a *field theory* dependant solely on the metric tensor[4] $g_{\mu\nu}$ and its derivatives, up to second order at most, Lovelock [1971] has shown that we are lead to:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} - \Lambda g_{\mu\nu} \,, \qquad (1.1)$$

where $R_{\mu\nu} = R^{\rho}_{\mu\rho\nu}$[5] and $R = R^{\mu}_{\mu}$ are the Ricci Tensor and scalar, respectively, that measure the distorsion in shape and volume of a small ball moving through space. $T_{\mu\nu}$ is the stress-energy(-momentum) tensor, that is, the collection of 4-momentum fluxes p_{μ} passing through a surface of constant x_{ν}, with $\mu, \nu = 0, 1, 2, 3$ being the time and space dimensions.

These 16 equations [6] are non-linear, since the field equations propagate through the very same space-time metric it governs.[7]

A cosmological model proposes a solution ($g_{\mu\nu}$) of these equations
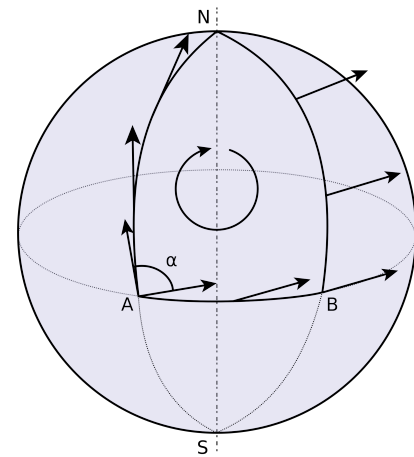
Figure 1.1: Parallel transport of a vector around a closed loop, using sections of grand circles on the sphere. The angle α is proportional to the area inside the loop. *Source: Wikipedia*

[4] The metric tensor contains the recipe for measuring distances along lines, by $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu}$, where we use the conventional Einstein notation: any indices repeated above and below are summed. So

$$g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} := \sum_{\mu=0}^{3}\sum_{\nu=0}^{3}g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} \,.$$

[5] $R^{\alpha}_{\mu\rho\nu}$ is the Riemann Tensor , quantifying the local deviation of the manifold from an isometry of the Euclidean space.

[6] 10 of which will be linearly independent only, as we require that the metric is symmetric $g_{\mu\nu} = g_{\nu\mu}$

[7] The Ricci and Riemann tensors are made of second derivatives and products of first derivatives of the metric $g_{\mu\nu}$, making the equations a system of non-linear partial differential equations in the metric. For a more detailed, step by step introduction to the theory, refer to Foster and Nightingale [2006].

[8] Solutions for which $T_{\mu\nu} = 0$ everywhere, like the Schwarzschild metric, are called *vacuum solutions* - they can still have non-zero energies and be not flat, because $T_{\mu\nu}$ does **not** account for gravitational energy. Gravitational energy is in general ill-defined, because the lack of unique parallel transport complicates generalisations of Stokes' theorem.

[9] in $c = 1$ units, which will help us make the equations more understandable - we can then return to $[c] = \text{km} \cdot \text{s}^{-1}$ with dimensional analysis in the end.

[10] This argument follows [Ellis, 2007]. In a nutshell, we discuss the dynamics of fluids in ETG, then impose the *cosmological principle* as symmetries: isotropy and homogeneity.

for a particular $T_{\mu\nu}$ that attempts to describe the observed universe in the largest scales. As solving systems of non-linear PDEs is **hard**, a strategy is to: 1) choose desirable symmetries from general physical arguments and 2) find suitable metric $g_{\mu\nu}$ and $T_{\mu\nu}$ that have these symmetries and are solutions of equation 1.1. $T_{\mu\nu}$ also must, of course, be related to some fundamental physics of the matter/energy content being considered.[8]

We will begin looking for our model by considering a mixture of fluids with physically motivated equations of state for constituents of the universe. The reason for doing so is that in both fluid motion and cosmology, we can work with a field of bulk velocities [Ehlers, 1993]. So, let us consider a congruence of fluid particle world lines, having timelike 4-velocities[9] $u^\alpha u_\alpha = -1$, where $u^\alpha = \mathrm{d}x^\alpha/\mathrm{d}\tau$, under the effect of their own gravity. The acceleration vector is given by $\dot{u}^\alpha = u^\alpha_{;\beta}u^\beta$ and vanishes if the world lines are geodesics. Then, motion of these fluid particles can then be characterised by distinct effects as [10]:

- expansion/contractions of volume, which are given by divergences of u^α, as $\theta := u^\alpha_{;\alpha}$,

- distortions in shape without change of volume which are trace free and orthogonal to u^α and are defined by the symmetric tensor,

$$\sigma_{\mu\nu} := u_{(\mu;\nu)} - \frac{1}{3}\theta h_{\mu\nu} - \dot{u}_{(\mu}u_{\nu)},$$

where $u_{(\mu;\nu)} = (u_{\mu;\nu} + u_{\nu;\mu})/2$ and $h_{\mu\nu} = g_{\mu\nu} - u_\mu u_\nu$.

- rotation and vorticity, without change in shape which are anti-symmetric and given by $\omega_{\mu\nu} := u_{[\mu;\nu]} - \dot{u}_{[\mu}u_{\nu]}$ with $u_{[\mu;\nu]} = (u_{\mu;\nu} - u_{\nu;\mu})/2$, and

- acceleration due to non-gravitational forces, like pressure gradients, given by the time derivative of the 4-velocity $\dot{u}_\mu = u_{\mu;\nu}u^\nu$.

The covariant derivative of the 4-velocity can now be written as

$$u_{\mu;\nu} = \sigma_{\mu\nu} + \omega_{\mu\nu} + \frac{1}{3}h_{\mu\nu} + \dot{u}_\mu u_\nu. \qquad (1.2)$$

Using this, the Riemann curvature tensor can be expressed as[Ellis, 2007]:

$$u^\alpha_{;\nu\beta}u_\mu - u^\alpha_{;\beta\nu}u_\mu =: R^\alpha_{\mu\beta\nu}. \qquad (1.3)$$

Contracting the expression above with $\alpha = \beta$ and multiplying by $u^\mu u^\nu$, we have the scalar equation

$$\theta_{;\nu}u^\nu - u^\alpha_{;\nu\alpha}u^\nu = R_{\mu\nu}u^\mu u^\nu\,,$$

which gives us the *Ricci identity*:

$$\dot{\theta} - \dot{u}^\alpha_{;\alpha} + 2(\sigma^2 - \omega^2) + \frac{1}{3}\theta^2 = R_{\mu\nu}u^\mu u^\nu\,, \qquad (1.4)$$

where σ and ω are the traces of their respective tensors.

Now, the stress-energy tensor of a perfect fluid in thermodynamic equilibrium is given by:

$$T_{\mu\nu} = \left(\rho + \frac{p}{c^2}\right)u_\mu u_\nu + pg_{\mu\nu}\,, \qquad (1.5)$$

where ρ is the matter energy density in the fluid, and p is the isotropic pressure, and the field equations can be rewritten as [Hawking and Ellis, 1973]

$$R_{\mu\nu} = \kappa\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right) - \Lambda g_{\mu\nu}\,,$$

where again T is the trace T^α_α and $\kappa = 8\pi G/c^4$. Now, contracting this with $u^\mu u^\nu$, using the perfect fluid stress-energy, one finds

$$R_{\mu\nu}u^\mu u^\nu = \frac{c^4\kappa}{2}\left(\rho + 3\frac{p}{c^2}\right) - \Lambda\,,$$

which, using equation 1.4, we can write:

$$\dot{\theta} - \dot{u}^\alpha_{;\alpha} + 2(\sigma^2 - \omega^2) + \frac{1}{3}\theta^2 = \frac{c^4\kappa}{2}\left(\rho + 3\frac{p}{c^2}\right) - \Lambda\,, \qquad (1.6)$$

which is the *Raychaudhuri's Equation*, which is a relativistic law of gravitational dynamics [Ellis, 2007].

We now consider additional symmetries for our solution: the *Cosmological Principle*, which states that our outlook of and place in the cosmos should probably be a *"common"* one. More formally, that means that the properties of the universe in a sufficiently large scale should look the same for any observer at any particular point, and that we expect that *fundamental observers*, i.e., imagined observers which follow the local mean motion of matter, will

experience an equivalent history of the Universe. In this case, both the shape distortions and vorticities are zero.

Now, if we at first consider only gravitational interactions, which are dominant, $u^{\mu}_{;\mu} = 0$, these assumptions will lead us to the Friedmann-Lemaitre model, with a Robertson-Walker metric (FLRW) as follows. For that, we can write the expansion as a function of the linear scale expansion by:

$$\theta = \frac{(\dot{dV})}{dV} = 3\frac{\dot{a}}{a} \, ,$$

where a is a normalised linear scale factor l/l_0 such that the volume V scale as $V \sim l^3$. Then, by using 1.6 we arrive at:

$$3\frac{\ddot{a}}{a} = -\frac{c^4\kappa}{2}\left(\rho + 3\frac{p}{c^2}\right) + \Lambda \, . \qquad (1.7)$$

The physical meaning of this *acceleration equation* is straightforward: a *change of rate* of the expansion or contraction of a small ball of radius l, relative to its size, is proportional to minus the energy density plus (3 times) the pressure, plus a constant. Since the scale a is defined as positive, if the constant Λ is zero, the rate of expansion must be decreasing (or the rate of contraction, increasing). Without the constant Λ, this is a quantitative statement which states that *gravity is attractive*.

Now, it is worthy to mention that a static universe ($\dot{a} = \ddot{a} = 0$) must have $\Lambda > 0$, and therefore is unstable: if we introduce a perturbation $l \rightarrow l + \delta l$ to a larger value, the matter density ρ increases, while Λ stays constant. Then $\ddot{a} > 0$ and the universe expands to infinity. Similarly, $l - \delta l$ will imply in $\ddot{a} < 0$ and the universe will collapse. So, *the universe should be either expanding or contracting, but not static*. From the first half of the twentieth century and on, evidence that the universe was expanding began accumulating [Slipher, 1917, Hubble, 1929, Sandage, 1958], and became the currently accepted view.

Back to our line of reasoning, if we use the conservation equation for fluids, and considering the expansion term, we have:

$$\dot{\rho} + \left(\rho + \frac{p}{c^2}\right)\theta = 0$$

$$\dot{\rho}a^2 + \underbrace{3\rho\dot{a}a}_{2\rho\dot{a}a + \rho\dot{a}a} + 3\frac{p}{c^2}\dot{a}a = 0$$

$$\dot{(\rho a^2)} = -a\dot{a}\left(\rho + 3\frac{p}{c^2}\right). \qquad (1.8)$$

Then, multiplying equation 1.7 by $\dot{a}a$, and substituting with 1.8 we finally arrive at

$$3\dot{a}^2 - c^4\kappa\rho a^2 - \Lambda a^2 = \text{const.} . \qquad (1.9)$$

This is the *Friedmann equation*, which describes the time evolution of FLRW universes. The constant term in the right hand side is a measure of the local curvature of space - that is, a 3-dimensional equal-time slice of the 4-dimensional manifold. Due to our assumption of homogeneity the curvature of space must be the same everywhere. Therefore, aspects of the global geometry of space-time can be constrained by its value. A positive curvature (*"spherical"*) everywhere will lead to a closed universe, whereas for zero (*"flat"*) or a negative (*"hyperbolic"*) spatial curvature, the topology can be either compact or infinite [Lachieze-Rey and Luminet, 1995].[11]

The Friedmann equation can be written in its most usual form as

$$\left(\frac{\dot{a}}{a}\right)^2 = -\frac{8\pi G}{3}(\rho) + \frac{\Lambda c^2}{3} - \frac{Kc^2}{a^2}, \qquad (1.10)$$

where K [12] is the constant of 1.9. For any time-slice of the universe, K/a^2 will give the spatial curvature.

If we choose a coordinate system in which any observer that sees an isotropic universe has constant space coordinate values, we can see that the left-hand side of the equation is the rate of expansion at a time t per unit distance. Since the distance from an observer to an object is $r = a(t)x$, the rate of change in the distance is $\dot{r} = \dot{a}x + a\dot{x} = (\dot{a}/a)r + a\dot{x}$. Here, $\dot{x}$ can be understood as a *peculiar velocity* of the object relative to the local isotropic frame and $(\dot{a}/a)r$ is the rate of recession of the isotropic observer frame and the isotropic object frame. In fact, the quantity $H(t) = (\dot{a}/a)$ can be directly measured if we assume

[11] Euclidean space is flat, simply connected, and infinite. A 3-torus is also (locally) flat, but multiply connected, finite and compact.

[12] *not to be confused with the Greek alphabet* κ - which will appear later as a gravitational lensing quantity.

that deviations from local mean motion should be randomly distributed, and measure the rate of recession of distant objects.

1.2   *The Standard Cosmological Model*

The above discussion led us to an equation to evaluate the background evolution of the universe. To employ it, we must now consider the properties of the contents making up the stress-energy tensor, which will be defined as fluids with some physically motivated equation of state ($p = \rho w$), and relate to present values. For convenience, the scale factor a can be normalised to one at the present time, and we define the present value of the Hubble factor $H(t)$ as $H(t_0) := H_0$, which is usually represented as $100h\mathrm{km} \cdot \mathrm{s}^{-1} \cdot \mathrm{Mpc}^{-1}$ where h, or sometimes h_{100} is called the reduced Hubble constant. It is often convenient to calculate quantities in Hubble constant independent units to compare results between different works[13].

For each type of content, the density can be written as a function of cosmological time as [Weinberg, 2008]

$$\rho \propto a^{-3(1+w)} . \tag{1.11}$$

[14]

For matter, dark or not, $w \approx 0$ and the density at a particular slice of time can be written as $\rho_m(t) = \rho_{m,0}/a^3$ respective to the current ρ_0 density. For radiation, we have $w = 1/3$ and then, $\rho_r = \rho_{\gamma,0}/a^4$ since not only the electromagnetic field density decreases but also the wavelengths (and hence photon energies) are changed due to gravitational effects [Weinberg, 2008]. Finally, we can, for Λ, define a density $\rho_\Lambda := \Lambda c^2/a^0$ which is constant.

The Friedmann equation is now, then:

$$H^2(t) = \frac{8\pi G\rho(t)}{3} - \frac{Kc^2}{a^2} , \tag{1.12}$$

with[15]

$$\rho(t) = \rho_r(t) + \rho_m(t) + \rho_\Lambda(t) .$$

If space curvature $K = 0$ we have the critical density of the

[13] This once was crucial: estimates for H_0 varied widely ($50 - 100\mathrm{km} \cdot \mathrm{s}^{-1} \cdot \mathrm{Mpc}^{-1}$). Now, it has been established as around $h = 0.7$., increased precision and different probes have found values in tension. [Riess et al., 2016, 2019]

[14] In general, $w = w(a)$, but this discussion is out of the scope of this text.

[15] Usually, the mass and radiation densities can be further divided into physical components:

$$\rho_m = \rho_c + \rho_b$$
$$\rho_r = \rho_\gamma + \rho_\nu ,$$

where ρ_c is the density of Cold Dark Matter, ρ_b is the density of baryons, ρ_γ is the density of photons, and ρ_ν is the density of neutrinos.

universe at time t given by:

$$\rho_{crit} = \frac{3H^2(t)}{8\pi G} \, . \qquad (1.13)$$

It is useful to redefine the densities of contents as a fraction of the critical density, so that we can describe the contents of the universe in *density fractions*, as the behaviour of the solution will depend on these. For any fraction χ we can define

$$\Omega_\chi(t) := \frac{\rho_\chi(t)}{\rho_{crit}(t)} \, . \qquad (1.14)$$

Even then, if we define a present time "curvature density" as

$$\Omega_K := -\frac{Kc^2}{H^2} := 1 - \Omega_M - \Omega_\gamma - \Omega_\Lambda \, , \qquad (1.15)$$

we can rewrite the Friedmann equation as a function of current densities as:

$$H^2(t) = H_0^2 \left(\Omega_\Lambda + \frac{\Omega_k}{a^2} + \frac{\Omega_m}{a^3} + \frac{\Omega_r}{a^4} \right) \, . \qquad (1.16)$$

Since $H = \dot{a}/a$, this is a first-order non-linear equation on $a(t)$, which can be solved algebraically for simple cases and numerically in general.

Until now, we have not discussed the nature of the Λ factor in the EFEs. There is no reason, at first, that the equations should contain the constant Λ, but it also does not spoil "covariant conservation". [16] Λ can be understood either as a dynamical property of empty space-time, as a part of the Einstein tensor (the left hand side of 1.1) or a universal vacuum density, as a part of the right-hand side. In this latter case, we can see that it acts as a constant density uniform fluid with negative pressure since $\rho_\Lambda \propto a^0$ implies that $p_\Lambda = -\rho_\Lambda/3$. Also, through the Raychaudhuri equation (specially, in its "Friedmann equation" form), it is clear that Λ violates the *strong energy condition* if the universe is expanding acceleratingly, since $\kappa'/2(\rho + 3p/c^2) - \Lambda < 0$. The same is thought to happen in the earliest times, during *inflation*, where the universe undergoes exponential expansion.

Recent experiments have been measuring the partial fractions of Dark Energy (DE), Dark Matter (DM), Baryonic Matter

[16] Covariant conservation, that is

$$(T^{\mu\nu})_{;\mu} = 0$$

$$\left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} \right)_{;\mu} = 0$$

does not imply energy or energy-momentum conservation. Since energy is *defined* as a current that is conserved as a result of the time-translational invariance of the laws of physics, conservation of energy must be defined along timelike vector fields. Killing vector fields are those who satisfy the equation $\xi_{\mu;\nu} + \xi_{\nu;\mu} = 0$, and are the generator of symmetries of the metric. While there are metrics which have timelike/null Killing fields, the Robertson-Walker metric that arises from the cosmological principle *does not*.[Hawking and Ellis, 1973, Misner et al., 1973]

with increasing precision, arriving at a picture that is most consistent with a universe ruled by ETG with zero curvature and dominated by DE and DM [Smoot, 1999, Hinshaw et al., 2013, Planck Collaboration et al., 2015]. These elements combine to form the *Standard Model of Cosmology*, known as ΛCDM, a universe dominated by Dark Energy(Λ) and Cold (that is, non-relativistic) Dark Matter (CDM), which has been extremely successful in explaining the structure and evolution of the cosmos. Baryonic matter contributes only with $\sim 4\%$ of the total energy density. Some of the state-of-the-art measurements of the relevant parameters are in table 1.1

Table 1.1: Measurements for ΛCDM cosmology by current CMB experiments.[Dark Energy Survey Collaboration, 2018, Planck Collaboration, 2018] Other model parameters suppressed for clarity. The explanation of σ_8 is in the next chapter.

Parameter	DESY1	Planck 2018	Parameter	Concordance
Ω_m	$0.267^{+0.30}_{-0.017}$	$0.3111(56)$	Ω_m	0.3
σ_8	$0.817^{0.045}_{-0.056}$	$0.8102(60)$		
Ω_Λ	0.733	$0.6889(56)$	Ω_Λ	0.7
$\Omega_b h^2$		$-0.052^{+(49)}_{-(55)}$	Ω_K	0
H_0	$67.2^{+1.2}_{-1.0}$	$67.66(42)$	H_0	70

1.3 *Measuring Very Large Distances*

It has been shown by [Ehlers et al., 1968] and generalised to approximations by [Stoeger et al., 1995] that if all free falling observers observe nearly isotropic background radiation, then the metric describing the background evolution of the largest scales is the Robertson-Walker metric:

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)\mathrm{d}l^2 \,, \tag{1.17}$$

where $\mathrm{d}l$ is the line element of space, which can be written in *comoving spherical coordinates* as[17]

[17] This states our $g_{\mu\nu} = \mathrm{diag}(-c^2, a^2/(1-Kr^2), a^2r^2, a^2r^2\sin^2(\theta))$ solution explicitly, finally.

$$\mathrm{d}l^2 = \frac{\mathrm{d}r^2}{1 - Kr^2} + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2 \,. \tag{1.18}$$

For photons, we can write without loss of generality a radial null geodesic as

$$\mathrm{d}s^2 = 0 = -c^2\mathrm{d}t^2 - a^2(t)\frac{\mathrm{d}r^2}{1 - kr} \,. \tag{1.19}$$

Now, since the value of Kr^2 can be absorbed into r^2 with a change of coordinates, only 3 special values of K have physical significance, namely $\{0, 1, -1\}$, representing flat, spherical and hyperbolic geometries respectively [Lachieze-Rey and Luminet, 1995]. These cases will lead to different mathematical analyses as we see in what follows.

Now, as r is a comoving coordinate, the total distance d_c is constant. The total length travelled by the photon can be calculated then as

$$\int_{t_e}^{t_0} \frac{c\mathrm{d}t}{a(t)} = \int_0^{d_c} \frac{\mathrm{d}r}{\sqrt{1 - kr^2}} = \begin{cases} \arcsin d_c & \text{if}(k = 1), \\ d_c & \text{if}(k = 0), \\ \text{arcsinh} d_c & \text{if}(k = -1). \end{cases}$$
$$(1.20)$$

where t_e is the time of emission and t_0 is the current age of the universe, in which $a(t_0) = 1$. Since The physical distance between comoving observers is $d = a(t)d_c$, the current distance is equal to the comoving distance - it will be less than that in the past, and greater in the future.

Since the FLRW universe space is neither time nor space translations invariant, the photon will not conserve neither its energy nor momentum along its path.[18] The wavelength of the photon will be directly proportional to the scale factor: $\lambda_0/\lambda_e = a(t_0)/a(t_e)$ which is usually expressed as the redshift $z = 1/a(t_e) - 1$, which is the measured change in frequency divided by the original frequency and can be directly observed from discrete line spectra.

[18] This is a direct consequence of the Noether's theorem.[Noether, 1918]

In terms of redshift, the Friedmann equation can be written, by direct substitution of $a \to 1 + z$ as:

$$H(z) = H_0 \sqrt{\Omega_\Lambda + \Omega_k(1 + z)^2 + \Omega_m(1 + z)^3 + \Omega_r(1 + z)^4}$$
$$(1.21)$$
$$= H_0 E(z)$$

For a photon emitted at a time t_e and observed at a time t_a the comoving distance between objects at redshifts z_1, z_2 can be

written, since $H(z)\mathrm{d}t = -a\mathrm{d}z$, as:

$$\left.\begin{array}{ll} \arcsin d_c & \text{, or} \\ d_c & \text{, or} \\ \operatorname{arcsinh} d_c & \end{array}\right\} = c \int_{t_e}^{t_a} \frac{\mathrm{d}t}{a(t)}$$

$$= c \int_{z_1}^{z_2} \frac{\mathrm{d}z}{H(z)}$$

$$= \frac{c}{H_0} \int_{z_1}^{z_2} \frac{\mathrm{d}z}{E(z)}\,, \qquad (1.22)$$

where $d_H = \frac{c}{H_0}$ is called the *Hubble horizon*, or the radius of the Hubble sphere. In terms of Ω_k, H_0 and c, we can write then:

$$d_c(z_1, z_2) = \frac{c}{H_0} \begin{cases} \dfrac{1}{\sqrt{\Omega_k}}\sin\left[\sqrt{\Omega_k}d_{c,0}(z_1, z_2)\right] & \text{, or} \\[2ex] d_{c,0}(z_1, z_2) & \text{, or} \\[2ex] \dfrac{1}{\sqrt{|\Omega_k|}}\sinh\left[\sqrt{|\Omega_k|}d_{c,0}(z_1, z_2)\right] \end{cases} \quad .$$

$$(1.23)$$

Naturally, neither physical nor comoving distances can be directly measured. To do so, we must rely on standard candles or some other well-behaved probe of distance.[19] In Euclidean space, the relations are $d_A = r/\theta$, for an object of known size r with an angular diameter θ and $d_L = 10^{\frac{m-M}{5}+1}$, for an object of absolute magnitude M and apparent magnitude m with $d_A = d_L = d_c$. However, as a result of the curvature of space-time, these equalities do not hold in FLRW universes. The angular distance between two different redshifts will be stretched by the relative expansion between them:

$$d_A(z_1, z_2) = d_c(z_1, z_2)\frac{a_2}{a_1} = d_c(z_1, z_2)\frac{1 + z_1}{1 + z_2}\,, \qquad (1.24)$$

whereas luminosity distance, due to flux conservation, will be related to the angular distance and the comoving distance by [Ellis, 1971]:

$$d_L(0, z) = (1 + z)^2 d_A(0, z) = (1 + z)d_c\,, \qquad (1.25)$$

which is the *Etherington's reciprocity relation* and is valid for any pseudo-riemannian description of space-time, **regardless** of the

[19] A *standard candle* uses the r^{-2} decay of flux of a source with known luminosity to measure distance.

theory of gravity. These measures of distances in cosmology will be necessary to not only determine the distance of clusters and background objects for gravitational lensing latter in this work, but also to calculate absolute magnitudes required for galaxy system observables. Using the luminosity distance, the difference between absolute and apparent magnitudes can be written as

$$DM = 5 \log \left[d_L(z) 10^5 \text{Mpc}^{-1} \right] \qquad (1.26)$$

1.4 The Big Bang

The expansion of space, when backwards extrapolated to very early times will lead us to the idea that the universe was once very hot and dense [Dodelson and Efstathiou, 2004]. In fact, one can show that for small enough ρ_Λ , the geodesics of our fluid congruence will be incomplete in the time coordinate, in other words, they cannot be extended beyond a certain proper-time or affine parameter. This is what is called a *time-like singularity* in the past, in which the curvature diverges for all points of space[Hawking and Penrose, 1970][20].

Either with or without singularities, it is now known that the early universe was dense and hot. The main observations that support this are:

- the Cosmic Microwave Background (CMB) [Smoot, 1999, Hinshaw et al., 2013, Planck Collaboration et al., 2015], which is the *relic* thermal emission of matter,

- the abundance of primordial elements [Steigman, 2007] created during the first minutes, and

- the observation of the near isotropic recession of distant objects from small to very large redshifts (Hubble's Law) e.g. [Riess et al., 1998].

The earliest phases of the Big Bang are still subject to much speculation. Although the oldest relic from the beginning directly detected is the CMB at redshift $z = 1108$, $t \approx 380000$ years, confidence in ETG allows us to formulate models of evolution that extrapolate to much earlier times.

[20] The conclusion that there must be a singularity in the past is a corollary of equation 1.6, considering $\Lambda \leq 0$. For positive Λ, it suffices to add that observations of decelerated expansion at high redshifts. The only way then to avoid an initial singularity is to assume the breakdown of the EFEs at some point, where quantum phenomena will be relevant and theory is still missing.
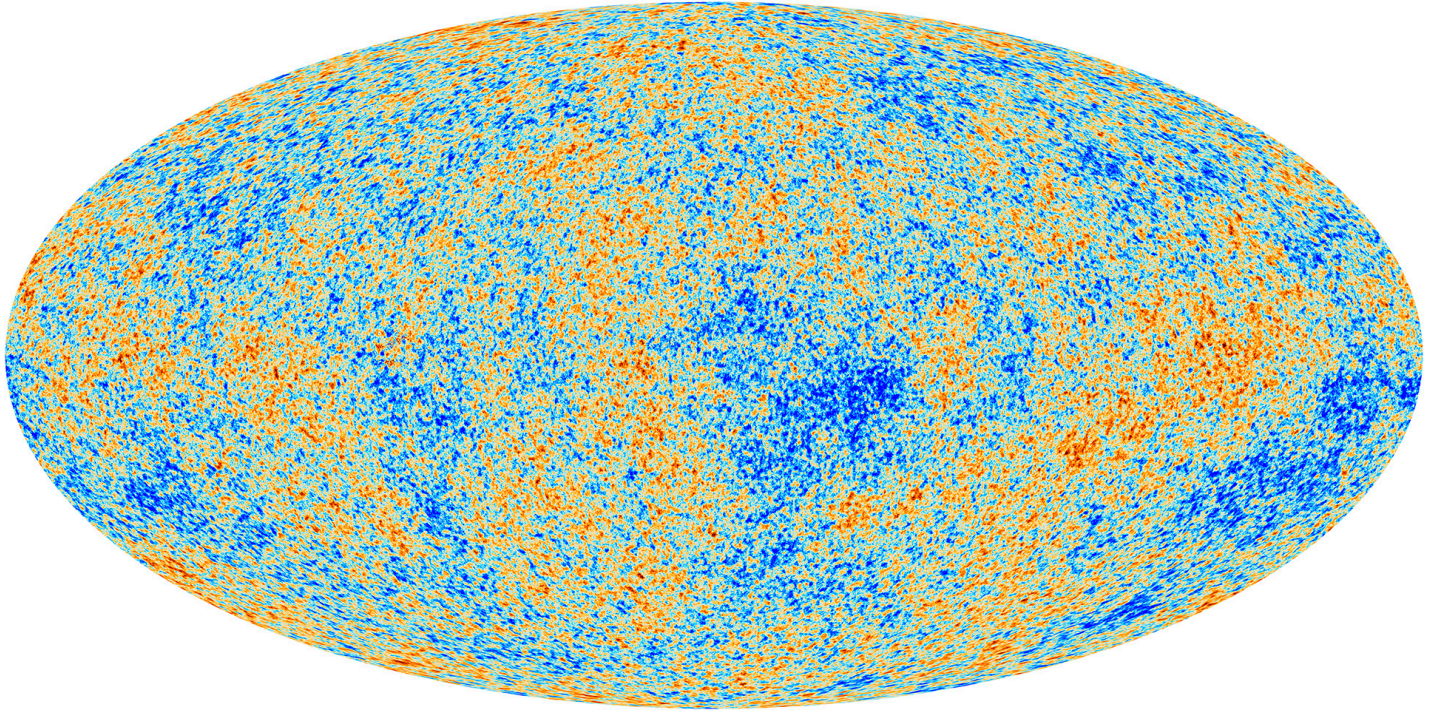
Figure 1.2: The all sky measurements of temperature deviations from the average of the Cosmic Microwave Background, by Planck Collaboration et al. [2015]. Cosmological parameters can be measured of statistics of the feature sizes of this map.

[21] An adiabatic change in background causes vacuum states to become particle states.

[22] Several processes and most of the underlying physical theory at this point are still poorly understood, but are extensively discussed in literature [Baumann, 2009, Sakharov, 1967].

In most currently discussed models, the universe began as a homogeneous quantum vacuum. At some point, a phase transition caused the universe to enter an exponentially accelerated expansion. Vacuum fluctuations then, under the rapid change of the underlying potential, became real particles[21] and density fluctuations deviating from the background homogeneity ever so slightly.[22]

As temperature, density and pressure decreased, the Electromagnetic and Weak interactions decoupled, and baryon physics arose as currently understood by physics with collision experiments. Further on, at $t = 10^{-4}$s, quarks condensed into hadrons. Neutrinos decoupled at about 1s and between the first and third minutes from the start, primordial nucleosynthesis created the primordial elements [Alpher et al., 1948]

From the end of the inflation to 10^4 years, the dynamics of the expansion were dominated by radiation content, because the a^{-4} term in the equation explodes. This is the *radiation era* and using the Friedmann equation 1.9 with all other densities being zero, we find that $a \propto \sqrt{t}$. After equipartition ($\Omega_m(t_{eq}) = \Omega_r(t_{eq})$), matter density became the dominant factor and expansion accelerated to $a \propto t^{2/3}$. The mean free path of the photon was small, as they constantly scattered off the free electrons and

protons of the hydrogen/helium plasma.

Eventually, the universe cooled to the point that the plasma condensed into neutral atoms, an epoch called *recombination*. Shortly after recombination, the photon mean free path became larger than the Hubble length, and photons travelled freely without interacting with matter. For this reason, recombination is closely associated with the last scattering surface, which is the name for the last time at which the photons in the cosmic microwave background interacted with matter.[23]

1.5 *Evolution of the Large Scale Structure of the Universe*

Small deviations from homogeneity can be studied by assuming a FLRW background and introducing local perturbations to the density field. As, in the earliest times, Ω_Λ is negligible, we consider $\Lambda = 0$ for now.

Deviations of a global average density can be written as

$$1 + \delta(\vec{x}) = \frac{\rho(\vec{x})}{\bar{\rho}}\,, \qquad (1.27)$$

where δ is called the *density contrast* with respect to the background average density $\bar{\rho}$. For now, we will consider linear perturbation, in which $\delta \ll 1$. Now we start with a simple case. A region with an overdensity $\delta(\vec{x})$ of radius $r < d_H$, in an otherwise flat ΛCDM background, has a deviation in expansion of $\theta = 3H \to 3H + \delta\theta$.

During matter domination, using the energy conservation (1.8) and the Raychaudhuri equation (1.6) with dark matter only[24], leads to background evolution equations at zeroth and first order:

$$\dot{\delta} + \delta_\theta = 0\,,$$
$$\dot{\delta\theta} + 2H\delta_\theta = -4\pi G\bar{\rho}\delta\,.$$

Eliminating δ_θ we have then:

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\bar{\rho}_m\delta = 0\,, \qquad (1.28)$$

which is an equation that governs the evolution of local matter overdensities δ with background density $\bar{\rho}$.

[25] The overall issue of introducing local inhomogeneities over background cosmology is an open problem - the "backreaction problem". Although it is solved for many particular cases [Green and Wald, 2013] it is not in general [Buchert et al., 2015].

The introduction of perturbations directly to the equations of motion as such is not actually compatible with ETG, but only with the Newtonian limit [Ellis et al., 2012][25]. That this Newtonian approximation works can be deceptively simple because its validity depends on certain gauge issues (or, more precisely, on our ability to write gauge-invariant gravitational potentials [Bardeen, 1980] in ETG, which are presented in clear form by Peacock [2003] in its section 1.2. Rigorous approaches can be read on Ellis et al. [2012] or Dodelson and Efstathiou [2004]

To work out some simple results, let us examine equation 1.28 by introducing the definition of Ω_m so that we have

$$\ddot{\delta} + 2H\dot{\delta} - \frac{3}{2}\Omega_m H^2 \delta = 0 \,. \tag{1.29}$$

Using the scale factor for a matter-dominated universe, this equation can be solved to $\delta(t) = \delta_+ t^{2/3} + \delta_- t^{-1}$ where $\delta_+ t^{2/3}$ is called the *growing mode*, since the density contrast with it grows with time (and eventually, collapses into visible structures). The evolution of the density perturbations can be written by a separation of variables as

$$\delta(\vec{x}, t) = D_+(t)\delta(\vec{x}, 0) \,, \tag{1.30}$$

where D_+ is called the *Growth Function* and is given, in the matter dominated era, as a function of the scale factor a by[Dodelson and Efstathiou, 2004]

$$D_+(a) \propto H(a) \int_0^a \frac{\mathrm{d}a'}{(a'H(a'))^3} \,, \tag{1.31}$$

and can be used normalised to $D_+(1) = 1$ for practical purposes. To account for radiation, which is not dominant after equipartition, but still relevant before recombination, the argument can be completed with a change of variables from t to $y = a/a_{eq}$, which together with the Friedmann equation gives us the *Mészaros equation* [Meszaros, 1974, Dodelson and Efstathiou, 2004]:

$$\delta'' + \frac{2+3y}{2y(1+y)}\delta' - \frac{3}{2y(1+y)}\delta = 0 \,. \tag{1.32}$$

This equation has two closed analytic solutions, the growing mode of which is $\delta(y) = y + 2/3$. Another mode is dominant if $y \ll 1$, which is $\delta \propto \ln y$. In this case, which we can extend to full radiation dominated era, and the latency in response of the density contrast due to the rapid expansion is called the *Mészaros effect*[Peacock, 2003] - a large overdensity (with size greater than d_H) that had been growing, enters the Hubble horizon and becomes nearly frozen (growing only logarithmically) before matter domination arrives.

Era	Ω	$H(t)$	$\delta(t)$
Radiation	$\Omega_m, \Omega_\Lambda \approx 0$	$\frac{1}{2t}$	$C_1 \ln t$
Matter	$\Omega_r, \Omega_\Lambda \approx 0$	$\frac{2}{3t}$	$C_+ t^{2/3} + C_- t^{-1}$
Λ	$\Omega_m, \Omega_r \approx 0$	$\sqrt{\frac{\Lambda}{3}}$	$C_1 + C_2 e^{-2\sqrt{\frac{\Lambda}{3}}t}$

Table 1.2: Sub-horizon sized matter overdensities evolution with time in different epochs.

For modes larger than the horizon d_H, Newtonian approximation breaks down, and we have to return to the full field equations.

There are many ways to define covariant potentials, analogues to the Newtonian potential Φ. Following the path of Dodelson and Efstathiou [2004] we can write:

$$3H\left(\dot{\Phi} + H\Phi\right) = 4\pi G a^2 \rho_m \delta \left[1 + \frac{4}{3y}\right] . \qquad (1.33)$$

This equation can be solved in the potentials and yields constant solutions for either matter domination or radiation domination, the latter being up to 10 times greater than the first [Dodelson and Efstathiou, 2004]. In both cases, using known relations for H and ρ, we arrive at $\delta \propto a^2$ for $a \ll a_{eq}$ and $\delta \propto a$ for $a_{eq} \ll a$. In essence, this shows that large modes will always grow.

Finally, to consider the effect of baryons - which are only relevant in sub-horizon evolution, since collisional effects are local, equation 1.28 must be changed to include pressure. This will result in

$$\ddot{\delta} + 2H\dot{\delta} + (c_s^2 \nabla^2 - 4\pi G \bar{\rho}_m)\delta = 0 , \qquad (1.34)$$

where $c_s = \sqrt{\partial p / \partial \rho}$ is the *sound speed* of the baryon fluid. This equation can be understood classically as $\ddot{\delta} - [\text{Pressure} -$

Gravity]$\delta = 0$, which is Newton's 2nd Law.

The solutions for equation 1.34 depend on the particular scale, since the gravitational forces of small overdensities cannot overcome the pressure of the photons. This is the *Jeans Length* and is expressed as:

$$\lambda_j = c_s \sqrt{\frac{\pi}{G\rho}} \qquad (1.35)$$

For a $\lambda > \lambda_j$ the structure will collapse, otherwise it oscillates due to pressure.

1.6   *The Role of Galaxies in Cosmology*

At the end of the first stages of cosmic evolution, the inhomogeneities contain the information that will shape the observed large scale structure of the universe. This structure is mostly isotropic and homogeneous on the largest scales, but its configuration displays measurable statistical properties that can be observed in the CMB, galaxy positions, and clustering. Assessing this information offers knowledge about not only the objects themselves but about the universe and the laws of physics [Weinberg et al., 2013].

The universe we see at night, on the other hand, displays a rich structure (Fig: 1.3), with field galaxies, groups, clusters and super-clusters of galaxies, filaments and large empty voids. To compare these observational results with the theory we use statistical methods over populations of features in observations, such as preferred scales in the CMB map, or number counts per interval of cluster mass, among others [Weinberg et al., 2013, Lima and Hu, 2005]. To make sense from the theory to observations, simulations, even though based in Newtonian physics and approximations for finite volumes, have been very successful to reproduce these statistical features of the large scales[Springel et al., 2005, Vogelsberger et al., 2014, Klypin et al., 2011].[26]

[26] Simulations are, in essence, numerical methods for solving the differential equations to calculate usable theoretical predictions.

Usually, one starts with the perturbed EFEs together with the Boltzmann equation for the fluids up to a certain redshift. Setting the initial conditions is problematic, since current models based on ETG will definitely fail when energy density reaches

2dF Galaxy Redshift Survey
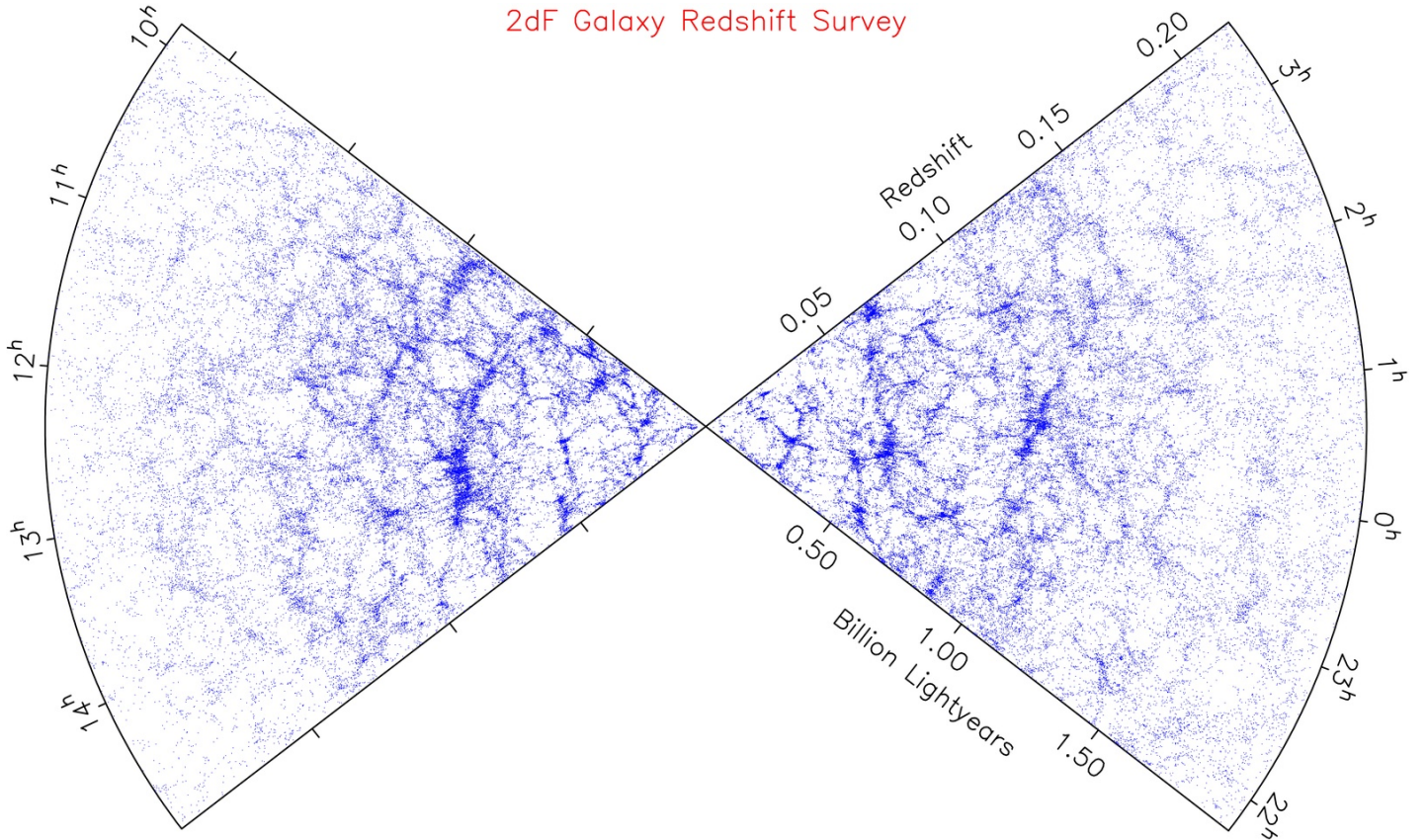
Figure 1.3: Distribution of galaxies of the 2dF Galaxy Survey [Colless et al., 2001]. From this, we can see the filamentary structure of matter distribution, as traced by galaxies. This same structure is found in cosmological simulations.

Planck-scale levels(about $t = 10^{-43}$).In these so called *Einstein-Boltzmann solver* codes, such as CAMB [Lewis and Challinor, 2002] or CLASS [Lesgourgues, 2011], one sets initial conditions after inflation, using some set of reasoned assumptions.

The statistics of the density field is encoded by n-point correlation functions.

$$\xi(r) = \langle \delta(\vec{x})\delta(\vec{x} + \vec{r})\rangle \qquad (1.36)$$

where $\xi(r)$ depends only on the distance between two points due to the statistical homogeneity and isotropy.

The density contrast can be written as an inverse Fourier transform

$$\delta(x) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\tilde{\delta}(k)e^{i\vec{k}\cdot\vec{x}} \qquad (1.37)$$

So defining the *power spectrum*[27] as

$$\langle \tilde{\delta}(\vec{k})\tilde{\delta}^*(\vec{k'})\rangle := (2\pi)^3 \delta_D^3(\vec{k} - \vec{k'})P(k) \qquad (1.39)$$

where δ_D^3 is the *Dirac distribution*, we can write the two-

[27] It is also usual to define also the dimensionless quantity

$$\Delta_k^2 := \frac{k^3 P(k)}{2\pi^2} \qquad (1.38)$$

which measures power per logarithmic scale.
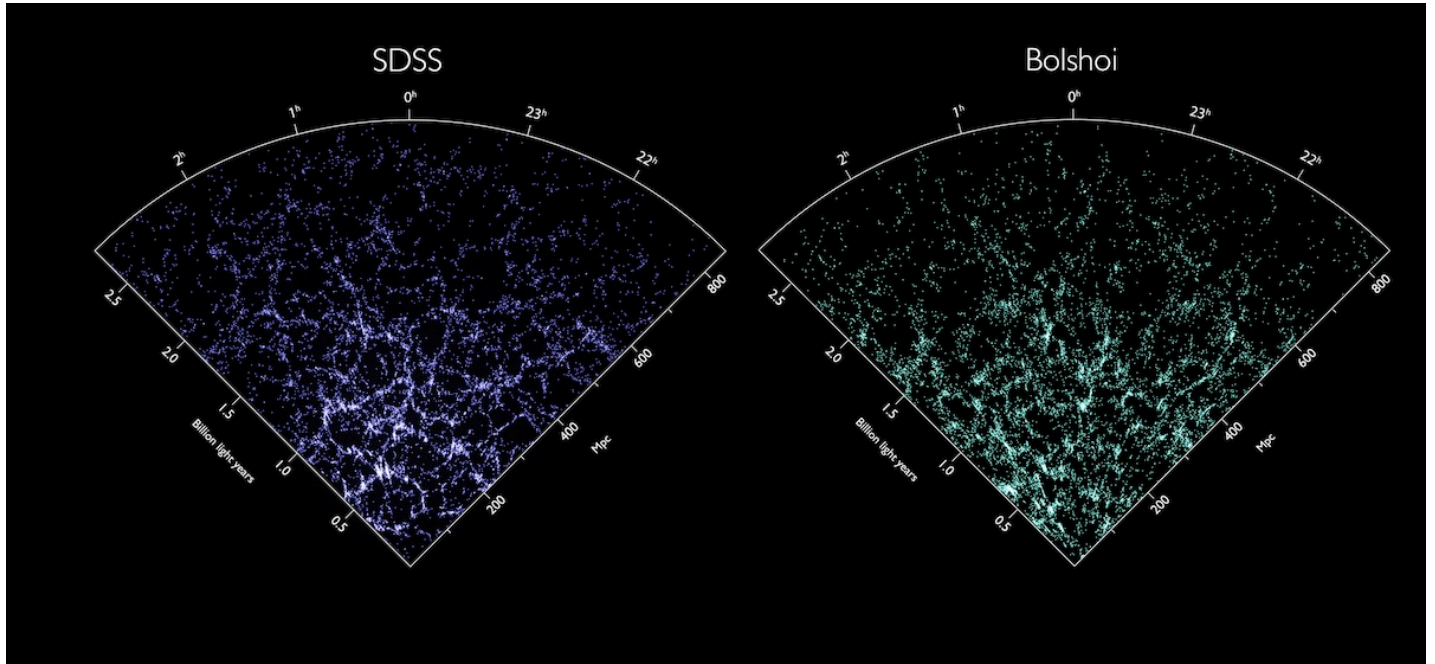
Figure 1.4: A comparison between the Bolshoi simulation and SDSS observation of the nearby universe. *Sources: Nina McCurdy/University of California, Santa Cruz; Ralf Kaehler and Risa Wechsler/Stanford University; Sloan Digital Sky Survey; Michael Busha/University of Zurich*

point correlation function as the Fourier transform of the power spectrum

$$\xi(r) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P(k) e^{i\vec{k}\cdot\vec{r}}$$
$$= \frac{1}{2\pi^2} \int \mathrm{d}k k^2 P(kr) \frac{\sin(kr)}{kr} \qquad (1.40)$$

The most appealing reason to write these quantities in Fourier space is that with translational invariance (which follows from homogeneity) we can characterise the statistics of random fields as preferred scales, or spatial frequencies, which readily show up in measured power spectra.

Up to now, we have seen how the universe evolves and how the stage is set for galaxies and galaxy clusters to act. It is time to move to non-linear scales of evolution[28], where both simplified models [Peebles, 1984, Zel'dovich, 1970] and simulations [Kravtsov and Borgani, 2012] can be used, usually providing complementary insights to understand the evolution of the large scale structure.

This final stage of collapse and formation of structures will be studied in the next chapter, as it closely relates to the dynamics and observables of galaxy systems, which is most important to our application in this work.

[28] which is characterised by $\Delta_k > 1$

2

Dynamics and Properties of Galaxy Systems

THE PHYSICAL PROPERTIES OF GALAXY SYSTEMS are, because of their complex configurations not only difficult to measure, but also convoluted. More than being "particles" bound by gravitational interactions, member galaxies fully interact, exchanging both matter content and energy, with multiple mergers of galaxies at the centre and accretion of other nearby formed associations. Feedback due to collisional processes of baryonic matter and other nonlinear physical processes, also contribute, with the gas and the star-formation also playing significant roles on final observables [Schneider, 2006]. It is crucial to understand the relationships implicated by these behaviours to correctly understand the landscape of galaxy system observables and therefore we proceed first into their formation history directly from the previous chapter, and then, on to discuss the role of the large-scale structure in their configurations. In the end of this chapter, we assess the multiple probes of cluster and group physics to place this work relative to the framework of extragalactic astronomy and cosmology.

2.1 From Overdensities to Groups and Clusters

The linear theory of evolution of density perturbation breaks down at some point. We are then forced into using simpler models for nonlinear evolution or, concomitantly, numerical simulations [Kravtsov and Borgani, 2012]. Despite its oversimplification, an overdense sphere is a very useful model,

which behaves as a small closed universe with a modified matter density due to the initial perturbation. Here, there is no need for this perturbation to be uniform, as the Poisson's equation guarantees that the evolution of any spherically symmetric perturbation is the same[Peacock, 2003, 3.2].

Far into the matter-dominated era, an overdense region can be described then by a Friedman equation with a different, enhanced local density. Then we can parametrise the solution by its proper radius and time, as functions of the development angle $\phi = H_0\eta\sqrt{\Omega_m - 1}$, where η is the conformal time as [Gunn and Gott, 1972]:

$$r(\phi) = A(1 - \cos\phi)$$
$$t(\phi) = B(\phi - \sin\phi)\,, \tag{2.1}$$

in which A and B are connected by $A^3 = GMB^2$. Expanding these relations to fifth order in ϕ gives $r(t)$ for small t as[1]:

$$r(t) \approx \frac{A}{2}\left(\frac{6t}{B}\right)^{\frac{2}{3}}\left[1 - \frac{1}{20}\left(\frac{6t}{B}\right)^{\frac{2}{3}}\right]. \tag{2.2}$$

Then, the density perturbation within the sphere will be

$$\delta \approx \frac{3}{20}\left(\frac{6t}{B}\right)^{\frac{2}{3}}. \tag{2.3}$$

Now we can examine the properties of the solution by looking at solutions with particular values for ϕ:

- if $\phi = \pi$, the radius is at maximum, and the overdensity detaches itself from the background evolution and turns around to collapse. At this point, the density contrast is $\delta_{tr} = 9\pi^2/16 \approx 5.55$. Using just linear theory, we would find that $\delta_{tr,\mathrm{lin}} \approx 1.06$,

- at $\phi = 2\pi$, $r = 0$, which is an idealisation of collapse. This occurs when $\delta_{col,\mathrm{lin}} = (3/20)(12\pi)^{2/3} \approx 1.686$.

This idealisation, however, differs substantially from reality as dissipation will act and convert the kinetic energy of collapse into random motion. By using the virial theorem, we can postulate an equilibrium at $r_{vir} = r_{tr}/2$, which occurs at $\phi = 3\pi/2$. By this

[1] it is not surprising that, at leading order $r \propto t^{2/3}$, which is just Einstein de-Sitter universe, where $r \propto a \propto t^{2/3}$

time, the density will have increased by a factor of 2^3, while the the background density will have decreased by a factor of 2^2, since $\rho \propto a^{-3}$ and $a \propto t^{2/3}$. The overdensity in the region will be

$$\frac{\rho}{\bar{\rho}} = 1 + \delta_{vir} = 1 + \delta_{tr} \times 8 \times 4 \approx 178 \,. \qquad (2.4)$$

We can define a truncation radius and the enclosed mass of a collapsed region then as

$$R_{200} = \left(\frac{3M_{200}}{4\pi\Delta_c\bar{\rho}_m} \right)^{1/3} \,. \qquad (2.5)$$

Although $\Delta_c \approx 178$ is physically motivated, the asymptotic behaviour of the profile makes a radius definition somewhat arbitrary. Additionally, including the full ΛCDM we would find a mass contrast of $\Delta_c \sim 340$ [Tasitsiomi et al., 2004], with a smaller radius than the R_{200} defined by $\Delta_c = 200$. We can, nevertheless, keep R_{200} as reference for comparing results [Johnston et al., 2007].

2.2   The Role of Cosmology in Galaxy Systems

After the collapse, overdense regions tend to develop to dynamical equilibrium [Kravtsov and Borgani, 2012], due to increase in entropy by energy dissipation. Smaller objects form first and may merge into larger objects. This characterises what is called the *hierarchical scenario* of structure evolution [Peebles, 1984]. Despite this complex landscape, dark matter particle simulations have been shown to display a characteristic shape (fig 2.1) for radial matter density distributions characterised by the logarithmic slope steepening with increasing radius (e.g. Navarro et al. [1996], Dubinski and Carlberg [1991]).

The density profiles obtained from simulations can be parametrised in a variety of ways, one of the most commonly of which is the *NFW profile*, given by:

$$\rho_{NFW}(r) = \frac{4\rho_s}{x(1+x)^2} \,, x = r/r_s \,, \qquad (2.6)$$

where r_s is the characteristic radius, where the logarithm slope of the radial density curve changes from -1 in the central region to
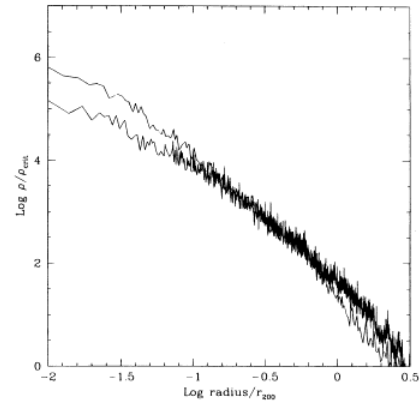
Figure 2.1: Results from simulations showing the characteristic steepening slope of dark matter density profiles. *Source: Navarro et al. [1996]*

[2] Other commonly listed profiles are the Einasto profile, given by:

$$\rho_E(r) = \rho_s \exp\left[\frac{2}{\alpha}\left(1 - x^\alpha\right)\right], x = r/r_s.$$

and the BMO[Baltz et al., 2009] profile, a modification of the NFW profile that incorporates a polynomial, smooth, truncation in outer regions

$$\rho_{BMO}(r) = \frac{4\rho_s}{x(1+x)^2}\left(\frac{y^2}{1+y^2}\right)^n,$$

where $y = r/r_t$ is a new truncation radius parameter and n is another truncation parameter. This is not even remotely an inclusive list of profile parametrisations, as there are many others.

-3 on outer radii .[2]

The self-similar shape of radial density profiles has been shown to be very general, being virtually independent of the shape of the power spectrum and background cosmology [Katz, 1991, Navarro et al., 1997], a feature usually called *universality*. However, a derivation from physical principles of this universal shape is still an open problem.

This universality of density profiles makes a useful tool for studying the matter distribution in the universe to replace the smooth field of densities (or equivalently, density contrasts) by a collection of superimposed individual *dark matter halos*. In this ansatz, these halos contain *all* matter, and therefore the density distribution is represented by an interpolation of halos scattered throughout space[Cooray and Sheth, 2002]. This is depicted in figure 2.2.
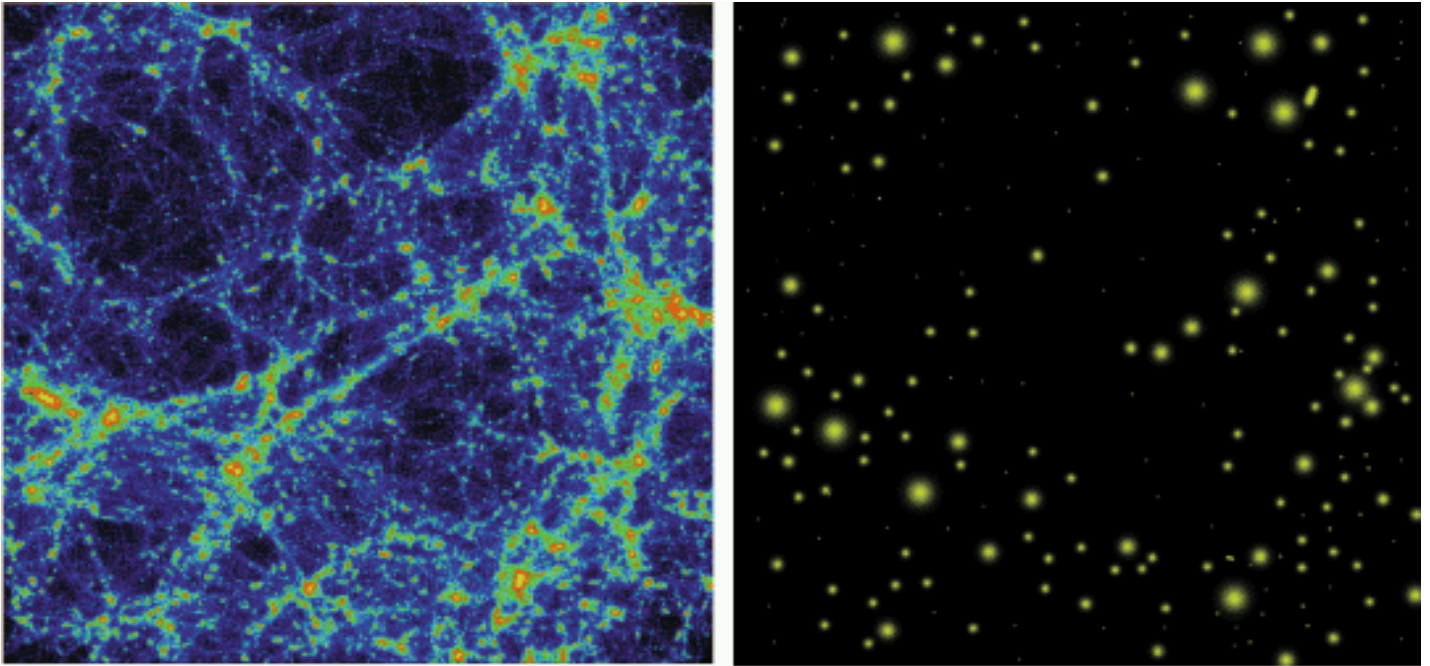
Figure 2.2: The complex, smooth, distribution of matter after the growth of structure in the universe can be understood as an interpolation of spherically collapsed halos. *Source: Cooray and Sheth [2002]*

To characterise particular cosmological models then, either simulations [eg. Tinker et al., 2008] or simple assumptions can be used to derive a prediction for the number density of collapsed halos of mass in an interval $[M, M + \mathrm{d}M]$ as we will see next. This is called the *mass function of clusters*, and is one of the most important probes of current cosmology [Weinberg et al., 2013]. The mass function is important not only to investigate cosmological models, as it is more sensitive than background

evolution [Lima and Hu, 2005], but also important to quantify effects of the large scale structure in the neighbourhood[3] of clusters, as we will see towards the end of this section.

As a first ingredient to the mass function, we must define the variance of the linear density contrast to our collapse threshold δ_c in order to quantify the fraction of collapsed halos per mass. For a given density contrast field $\delta(\vec{x})$ we can filter to halos using a window function $W(\vec{x}, R)$, normalised to unity to get a smoothed field

$$\delta(\vec{x}; R) = \int \delta(x) W(\vec{x} + \vec{x}'; R) \mathrm{d}^3 x' . \qquad (2.7)$$

This filter defines an enclosed mass $M := c_f \bar{\rho} R^3$ where c_f is some constant used to normalise the filter. Now, in Fourier space this convolution integral becomes a product $\delta(k; R) = \delta(k)\tilde{W}(kR)$. Using then a top hat model for the filter, we can write the filtered variance of the field as a function of filter radius R as

$$\sigma_R^2 = \overbrace{\langle \delta^2(\vec{x}; R) \rangle}^{\propto \xi(R)} = \frac{1}{2\pi^2} \int_0^\infty \mathrm{d}k \; k^2 P(k) \left| \tilde{W}(\vec{k}, R) \right|^2 , \quad (2.8)$$

where now $\tilde{W}$ is specifically the transform of the top hat function.[4]

If we then transform the radius into a mass scale by

$$R = \left(\frac{3M}{4\pi\rho_c} \right)^{1/3} , \qquad (2.10)$$

where $\rho_c = \bar{\rho}_m \delta_c$ is the density of the collapsed halo we can use the mass variance $\sigma^2(M)$ to calculate the mass function as a function of density peak height relative to the variance $\nu = \delta_c/\sigma(M)$.

The first statistical model for mass function was developed by Press and Schechter [1974], in which the main underlying idea is that the probability that $\delta_M > \delta_c$ at a given time is the equal to the fraction of mass contained in halos with mass greater than M at that time.

Considering that the distribution of density perturbations is a Gaussian random field, the fraction of fluctuations above the threshold δ_c correspond to collapsed regions and is given by

$$F(M) = \frac{1}{\sqrt{2\pi}\sigma(M)} \int_{\delta_c}^{\infty} \mathrm{d}\delta \exp\left[-\frac{\delta^2}{2\sigma^2(M)}\right] = \frac{1}{2}\mathrm{erfc}\left[\frac{\nu^2}{2}\right],$$
(2.11)

where $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$ is the complementary error function.

This result is, however, problematic. Since $\lim_{x\to 0} \mathrm{erfc}(x) = 1$ and $\lim_{x\to\infty} \mathrm{erfc}(x) = 0$, this model predicts that never more than $1/2$ of all matter in the universe is inside collapsed regions. The fraction of dark matter in halos above M then must be multiplied by an additional "fudge factor" of 2 in order to ensure that every particle ends up as part of some halo with $M > 0$. This is because underdense regions can be enclosed within larger overdense regions, giving them a finite probability of being included in some larger collapsed object.

Now, the number of halos with masses in the range $[M, M + \mathrm{d}M]$ per comoving volume at a time t $n(M, t)$ can be written as

$$n(M,t) = \frac{\mathrm{d}n}{\mathrm{d}M} = M\frac{\mathrm{d}n}{\mathrm{d}\ln M}.$$
(2.12)

Using the formalism idea that $\frac{\partial F}{\partial M}\mathrm{d}M$ is equal to the fraction of mass locked up in halos with masses in the range $[M, M + \mathrm{d}M]$ we find that

$$n(M,t)\mathrm{d}M = \frac{\bar{\rho}}{M}\frac{\partial F}{\partial M}\mathrm{d}M$$
(2.13)

$$= \overbrace{2}^{\text{fudge}} \frac{\bar{\rho}_m}{M}\frac{\partial}{\partial M}\frac{1}{2}\mathrm{erfc}\left[\frac{\nu^2}{2}\right]\mathrm{d}M$$
(2.14)

$$= \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}_m}{M^2}\nu \exp\left[-\frac{\nu^2}{2}\right]\left|\frac{\mathrm{d}\ln \sigma_M^{-1}}{\mathrm{d}\ln M}\right|\mathrm{d}M,$$
(2.15)

which is the PS mass function.

If we rewrite the mass function as

$$\frac{\mathrm{d}n}{\mathrm{d}\ln M} = \frac{\bar{\rho}_m}{M}\frac{\mathrm{d}\ln \sigma^{-1}}{\mathrm{d}\ln M}f(\nu),$$
(2.16)

we can recognise that the *multiplicity function* $f(\nu) = \sqrt{2/\pi}\nu e^{-\nu^2/2}$ encodes the characteristic shape of the Press-Schechter mass function.

Further advancements have been made since this first model and today, numerical simulations give us accurate mass functions from specific cosmological models. Using the original PS multiplicity function as a guide for the functional form, parametrisations are fitted to the result of simulations, which can in turn be compared to cluster counting in surveys [Tinker et al., 2008] to test new cosmological models. [5]

The relationship between the density of collapsed halos is not equivalent to the density distribution of matter, however. As it can be seen in figure 2.3, the so called *long-wavelength modes* of the density distribution interfere in spatial location, providing a higher local average $\bar{\rho}_m$ and thus enhancing the density of halos with respect to the density of matter. For that reason, collapsed objects as galaxies, groups or clusters, are *biased* tracers of the underlying matter distribution. In regions with a higher count of objects, the underlying matter density distribution will have a higher value than expected by a simple proportion as $\delta_m \propto \delta_h$ would suggest. This can be understood under the *peak-background split*, which we develop succinctly below.

[5] The Tinker mass function, which will be used in this work, is parametrised as

$$f(\sigma) = A\left[\left(\frac{\sigma}{b}\right)^{-a} + 1\right]e^{-c/\sigma^2}.$$

(2.17)

X

Figure 2.3: Regions with higher density due to long modes of mass fluctuation tend to form more collapsed structures because the collapse threshold is closer to the mean density. This, in turn, affects the mass distribution around clusters. *Source:* *Peacock [2003]*
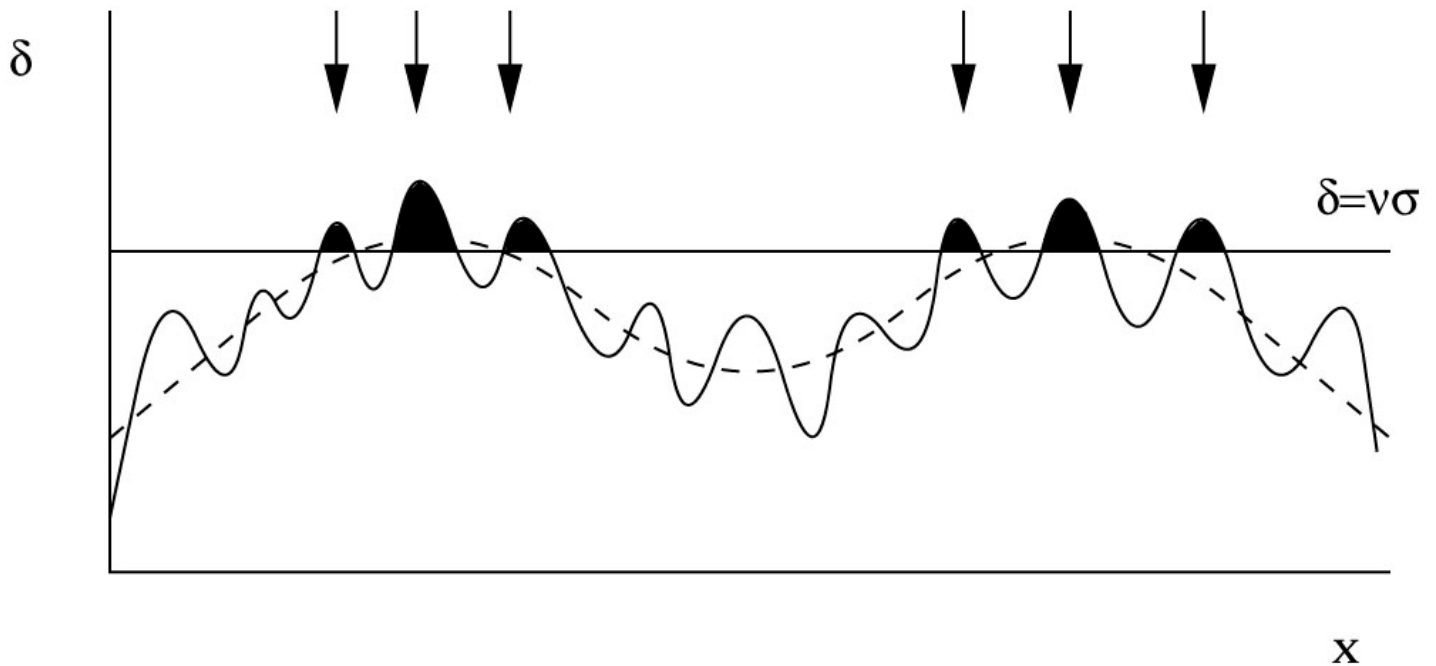
At first order, we expect that this *bias* in measuring mass

distribution due to a particular tracer is a linear function related to the height of the peak of the collapsed density perturbation[Cooray and Sheth, 2002] .

For halos, if δ_h is the contrast of halo density[6] in a small region, then:

$$\delta_m = b_h(\nu)\delta_h \, . \tag{2.18}$$

A region with an average density higher than the background can be written as a perturbation so that the local collapse overdensity is reduced to $\delta_c' = \delta_c - \epsilon$ and now $\nu = \delta'/\sigma_M$. If we expand the perturbed mass function 2.16 in a power series of ϵ we will have the number density modulated by

$$f' = f - \frac{\mathrm{d}f}{\mathrm{d}\delta_c}\epsilon \, , \tag{2.19}$$

so that the bias will be related to the mass function as

$$b(\nu) = 1 - \frac{1}{\delta_c}\frac{\mathrm{d}\ln f}{\mathrm{d}\ln \nu} \, . \tag{2.20}$$

Finally, from simulations, [?] derives a parametrical large scale bias as

$$1 - A\frac{\nu^a}{\nu^a + \delta_c^a} + B\nu^b + C\nu^c \, , \tag{2.21}$$

where the values A,a,B,b,C, and c are given by table 2.1. This is the form we will use in our analysis.

Table 2.1: Parameters of the bias equation as a function of $y = \log(\Delta)$

Param.	
A	$1 + 0.24y \exp\left[-(4/y)^4\right]$
a	$0.44y - 0.88$
B	0.183
b	1.5
C	$0.019 + 0.107y + 0.19 \exp\left[-(4/y)^4\right]$
c	2.4

Now we can estimate the effect of clustering in distorting the cluster mass radial profile due to presence of neighbouring overdense regions. To do so, we first define the halo-mass

[6] Formally, δ_h is the density contrast of collapsed halos of mass m at redshift z_1 given a mass M in a comoving volume V at a redshift z_0 and is defined by

$$\delta_h(m, z_1|M, V, z_0) := \frac{N(m, z_1|M, V, z_0)}{n(m, z_1)V} - 1 \, .$$

correlation function as

$$\xi_{hm}(r, M, z) := \langle \delta_h(x)\delta_m(x+r)\rangle\,, \qquad (2.22)$$

that quantifies the excess matter density relative to the background at a distance r from the centre of a halo of mass M at a redshift z. Its natural estimator is then

$$\xi_{hm}(r, M, z) = \frac{\overbrace{\rho_{1h}(r, M, z)}^{\rho_{NFW}}}{\bar{\rho}_m} + b^L(M, z)\xi_{mm}^L(r, z)\,, \qquad (2.23)$$

where we label b^L and ξ^L the linear bias and the linear mass correlation function.

The projected mass density due to the halo and the large scale structure can now be written as

$$\begin{aligned}
\Sigma(r|M, z) = \int \mathrm{d}z \delta\rho &= \int \mathrm{d}z \bar{\rho}_m \xi_{hm}(r, M, z) \\
&= \int \mathrm{d}z \rho_{1h} + \bar{\rho}_m b^L(M, z)\xi^L(r, z) \\
&= \Sigma_{1h}(r) + \Sigma_{2h}(r)\,, \qquad (2.24)
\end{aligned}$$

where ρ_{1h} is the cluster profile own profile, called the 1-halo term and $\rho_{2h} = \bar{\rho}_M b^L(M, z)\xi^($r, z)$ is the 2-halo term, that is, the contribution to the profile due to neighbouring halos.

2.3 The Environment of Galaxy Systems

[7]

The environment inside clusters and groups of galaxies differ substantially from the rest of the universe and that results in measurable effects not only on the morphology of galaxies that populate these overdense regions, but also on the dynamics of interactions between their contents. One first clear example of these effects is that the mixture of galaxy types inside of clusters is visibly different from the field: whereas about 70% of the galaxies are spirals in the field, clusters are dominated by ellipticals (Fig. 2.4)[Oemler, 1974].

We now turn our attention to some aspects that determine the content and dynamics of galaxy clusters and groups, to help us understand their nature and then proceed to rank their observable quantities.

Figure 2.4: The galaxy cluster Abell 2218 as imaged by Hubble, provides a display of the characteristic image of the environment of galaxy clusters, as well as their effect not only in the environment, but, as we will later see, on the images of background galaxies. The yellowish hue of elliptical galaxies and a dominant cD central galaxy show how the galaxy population of galaxy clusters stand apart from that of galaxies of the field. The thin, distorted arcs are gravitationally lensed images of background galaxies, that we will see in chapter 3. *Source: Image Credit: NASA, ESA, and Johan Richard (Caltech, USA)*

Galaxy systems have matter densities from 50 to 200 times the average density of the universe. Their masses range from $10^{13}M_\odot$ to some $10^{15}M_\odot$ and typical radii of around $1\mathrm{Mpc}/h$. Their matter content can be divided into of 4 different components:

- galaxies, which, by comparing their luminosities and the typical star populations, can only account for a tiny fraction($\sim$ 3%) of the mass, but are one of the main probes of the mass distribution, as they are expected to follow the overall density, and are readily visible to ground-based telescopes in optical,

- intracluster medium (ICM), composed mostly of a diffuse plasma spread throughout the clusters, with very high temperature (in the range between 10^7 and 10^8K) and densities

of the order of 10^{-3} particles/cm^3. The ICM is detected by X-ray thermal brehmsstrahlung emissions or by inverse Compton scattering of the Cosmic Background Radiation, which is called *Sunyaev-Zel'dovich* Effect (more on the next chapter),

- intra cluster light, that is a diffuse optical component made of stars that have been striped from member galaxies by gravitational interaction.

- dark matter, which is indirectly inferred by the amount of mass measured that cannot be accounted for with stars and the ICM and accounts for $80-85\%$ of the total, and can be quantified by methods that investigate total mass through gravitational effects.

As galaxy systems evolve they will tend to appear more spherically symmetric and to have number densities of galaxies higher in the centre [Butcher and Oemler, 1978]. The overall dynamical state of galaxy systems is varied. However, there is indication that at least a part of them is relaxed[Smith et al., 2005]. This, in turn, is complicated by the fact that two-body collisions of galaxies inside clusters are not enough to explain relaxation, since the relaxation time is greater than the age of the universe for typical galaxy systems[8].

Lynden-Bell [1967] proposed that large amplitude variations in the gravitational field, as in galaxy formation or collisions, *can* drive a *quasi-relaxation* process that is much faster than the two-body relaxation time. [9] Once the collapse of a cluster is complete, however, violent relaxation becomes ineffective, and the process must continue only through two-body interactions.

Chandrasekhar [1943] showed that massive objects moving through a a distribution of lighter objects will be dragged by an alignment of these smaller ones behind it. This alignment, in its turn, takes place exactly because the larger object pulled (See Fig. 2.5) While the original argument was on star clusters, this is readily applicable to galaxy clusters, and can be expressed as

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} \propto -\frac{M\rho\vec{v}}{|\vec{v}|^3} \,.$$

(2.27)

This effect is named *dynamical friction*. Being proportional to the mass, it is thought to cause galaxies to plunge inwards, with

[8] The dynamical relaxation time can be approximated by

$$t_R = t_{cross}\frac{N}{\log N}\,,$$

where t_{cross} is the crossing time.

[9] This *violent relaxation* time is given as

$$t_{VR} = \sqrt{\left\langle \frac{\varepsilon^2}{(\mathrm{d}\varepsilon/\mathrm{d}t)^2} \right\rangle}\,,$$

(2.25)

where ε is the energy per unit mass of the system. Using the virial relation, it can be shown that this time is of the order of the crossing time $t_{RV} \approx t_{cross} \approx 1/\sqrt{GR}$, unlike in the case of two-body relaxation, where

$$, t_{relax} = t_{cross}\frac{N}{\ln N}$$

(2.26)

which is much larger than the age of the universe.
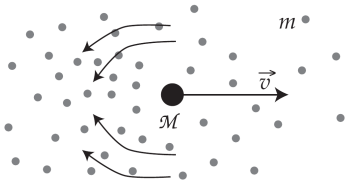
Figure 2.5: The principle of dynamical friction. The gravitational field of a galaxy of mass M and velocity $\vec{v}$ accelerates other smaller galaxies to its track, which then act as a force pulling the massive galaxy backwards. *Source: [Lima Neto, 2014]*

[10] ~ 10kpc

[11] This might be a clue to an ongoing merger [Lin and Mohr, 2004]

Figure 2.6: A typical massive cD galaxy with an active nucleus, M87. (*Source: Raniere Menezes, IAG Telescope, OPD*)

more massive galaxies sinking into the cluster centre.

The observed galaxies in clusters and groups are shown to be *redder* than the population of field galaxies, furthermore, this red color is tightly correlated with their luminosity [Baum, 1959, Bower et al., 1992]. In colour-magnitude diagrams, this will appear as a distribution around a line, called the *red sequence*. This red sequence is mostly populated by elliptical galaxies, displaying a small scatter, and being very similar for all clusters at a given redshift, regardless of mass [Koester et al., 2007, Gladders et al., 1998]. The red sequence is a powerful tool to identify cluster of galaxies because it is generally easier to measure colours than redshifts of galaxies to assess their clustering on the line of sight. Conversely, the red sequence can be used to determine redshifts of clusters, as used in [Rykoff et al., 2014] a red-sequence based cluster finding algorithm that is used in this work, and will be explained in chapter chapter 4.

Most galaxy clusters have a very bright galaxy very[10] near its center, usually called the *Brightest Cluster Galaxies* - BCGs. Here we refer to them as central galaxies (CG) only, because in some methods of cluster finding, there are clusters in which the central is not the brightest[11]. Many of these central galaxies are called type "cD" - giant ellipticals that have significant morphological differences from the rest of the elliptical population [von der Linden et al., 2007]. They have extended stellar envelopes that may exceed $R \sim 100h$kpc and have broader luminosity profiles than other elliptical galaxies. Many (from a fourth to half) cD galaxies have multiple cores, which indicate recent merging of other galaxies, in what is (perhaps fancifully) called *galactic cannibalism* [Dubinski, 1998].

Central galaxies also have larger axes roughly aligned to the same direction of the overall galaxy distribution in the cluster and even with respect to the larger scale structure, which cannot be attributed to rotation [Carter and Metcalfe, 1980].

Overall, in the innermost region of a cluster the galaxy-crowded environment makes physical collisions between galaxies important. This argument, along with dynamical friction, are the cornerstones upon which the last section of this chapter is built, where we talk about *fossil clusters*, in which these processes are

believed by some (e.g. Ponman et al. [1994], Jones et al. [2003]) to have led to overgrown central galaxies, in expense of their nearby, massive companions.

Turning now to the gas content of clusters, the ICM is sparse, but very hot $T \sim 10^8 K$, emitting brehmsstrahlung radiation on X-rays, with luminosities of the order of $L_X \sim 10^{43} - 10^{45} \mathrm{erg}/s$ and constitute the greatest fraction of cluster baryons. Instead of galaxies, which are non-collisional, the gas shows structures characteristic of fluid flows, specially on major wet mergers, as the case of the *Bullet cluster*. The separation of the baryon matter content and dark matter profiles in these systems is today one of the most pressing arguments against modified theories of gravity to account for the missing mass problem in galaxies and clusters [Clowe et al., 2006], as well as helping constrain dark matter models [Harvey et al., 2015].

Because the gas emits radiation with such an immense power, it was expected that the gas would cool and flow towards the centre of the cluster [Fabian, 1994, Croton et al., 2006]. These *cooling flows* have been inferred in the form of sharp central peaks in central emissivity, but they do not account for the total expected cooldown. The leading hypothesis for the prevention of massive cooling flows is the reheating of the gas due to feedback provided by *Active Galactic Nuclei* (AGNs) [Peterson and Fabian, 2006].[12]

Groups are the smaller systems of galaxy associations. They are composed of a few (usually $N < 50$) galaxies of luminosity $L \sim L^*$ (see the next section) and comprise slightly more than half of nearby structures in the universe. Groups of luminosity equal or less than that of the Local Group ($\lesssim L^*$) correspond to 50% of the luminosity density of the universe[Gott, 1977] .

[12] Active Galactic Nuclei are thought to be *supermassive black-holes*, which in turn are another prediction of ETG, as we have shown in chapter 1.

2.4 Observables of Galaxy Systems

The main features that are physically relevant for galaxy cluster astrophysics are the total mass, total luminosity, and relaxation status whereas for cosmology, the number counts of clusters per mass interval is, as we have previously seen, of particular importance. Unfortunately, except for luminosity, none of these constitute direct observables.

[13] analogous to the mass function for whole clusters

The characteristics of cluster galaxy distribution in luminosity are given by the distribution of galaxy counts per luminosity interval, or the *luminosity function*. The luminosity function of galaxies[13] does not differ qualitatively from galaxies in the field to those in systems, and can be represented by a Schechter function [Schechter, 1976], where $\Phi(L)\mathrm{d}L$ represents the number of galaxies in a luminosity interval $[L, L + \mathrm{d}L]$, as:

$$\Phi(L) = \left(\frac{\Phi^*}{L^*}\right)\left(\frac{L}{L^*}\right)^{\alpha}\exp\left(-\frac{L}{L^*}\right),\qquad(2.28)$$

where L^* is a characteristic Luminosity above which the number of galaxies in a population become exponentially rarer, and Φ^* is the normalisation of the distribution.

The luminosity function can also be written in terms of magnitude, as:

$$\Phi(M_\lambda) = (0.4\ln 10)\Phi^* 10^{0.4(\alpha+1)(M_\lambda^*-M_\lambda)} \times \exp\left(-10^{(M_\lambda^*-M_\lambda)}\right),$$
$$(2.29)$$

where M_λ is the absolute magnitude of the galaxy at a filter around the frequency λ.

The total luminosity of the cluster can be then calculated as

$$L_{tot} = \int_0^\infty \mathrm{d}L\; L\Phi(L) = \Phi^* L^* \Gamma(2+\alpha),\qquad(2.30)$$

where $\Gamma(x)$ is the gamma function $\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}\mathrm{d}t$.

Masses of clusters must always be measured by proxy, since they cannot be readily measured. A range of mass proxies exist, such as the temperature of the intracluster medium plasma, displacement of the CMB spectrum due to the Sunyaev-Zel'dovich effect, total luminosity, richness, [14] the velocity dispersion of member galaxies, and distortions of background galaxies due to gravity, which will be studied in detail on chapter 3.

[14] which is the number of galaxies in the cluster

The temperature of the intracluster gas correlates with the total mass because the depth of the gravitational potential is related to the mean kinetic energy of the gas particles, if we assume hydrostatic equilibrium. The mass enclosed inside a radius r is given as a function of the radial profiles of gas density and

temperature as

$$M(r) = -\frac{k_B T r^2}{G \mu m_p} \left(\frac{\mathrm{d}\ln \rho_g}{\mathrm{d}r} + \frac{\mathrm{d}\ln T}{\mathrm{d}r} \right), \qquad (2.31)$$

where k_B is the Boltzmann constant, $\mu \approx 0.63$ is the average mass of gas particle per units of proton mass m_p, and ρ_g is the density of the gas. [15]

The ICM gas temperature itself can be measured by X-ray observations[16], in space-based telescopes as Chandra and XMM-Newton [Lima Neto, 2014] or through the observation of the Sunyaev-Zel'dovich effect, as the inverse Compton scattering by the ICM electrons pushes incoming CMB photons to higher energies. One of the advantages of this method is that it can probe and find clusters of galaxies to much greater distances, since the SZ effect is independent of redshift (as the CMB is further behind any structure in the universe). The gas is isothermal and is described by a β-model, temperature of the ICM is related to the shift of the CMB temperature and luminosity distance by

$$T_X \propto \left(\frac{\Delta T}{T} \right)_{CMB}^{4/3} d_L^{-4}, \qquad (2.32)$$

where dL is the luminosity distance (Eq. 1.25). The *velocity dispersion* of the galaxies provide another way to measure the mass of a galaxy cluster. Using the virial theorem, one can show that

$$M = \frac{3\pi R_G \sigma_v^2}{2G}, \qquad (2.33)$$

where $R_G \sim 1\mathrm{Mpc}$ is the *gravitational radius* and σ_v is the velocity dispersion, measured by comparing galaxy redshifts to the cluster overall redshift.

Both gas temperature and velocity dispersion models assume dynamical equilibrium. A way to investigate the mass distribution in clusters without this assumption is through the observation of the gravitational effects on the light of background galaxies. This method will be discussed in detail in the next chapter, as it will be put to use in part II.

The radial distribution of mass, galaxies, luminosity, is easier to be investigated in the plane of the observation, not on the line-of-

[15] In current practice, a mass-temperature relation has to be calibrated, because reheating by AGNs and supernovae will alter the relationship between gas temperature and the depth of the potential well.

[16] Insert here an X-ray cluster image.

sight. To do so, all quantities must be transformed to projected quantities as

$$\Sigma(R) = \int_{-\infty}^{+\infty} \mathrm{d}z \delta\rho(\sqrt{R^2 + z^2}) = 2 \int_{R}^{+\infty} \frac{\mathrm{d}r\rho(r)}{\sqrt{r^2 - R^2}}. \quad (2.34)$$

The projected mass density can also be written as a differential mass density, which will be useful as it is directly observable by lensing, as

$$\Delta\Sigma(R) = \bar{\Sigma}(r < R) - \Sigma(R) \quad (2.35)$$

Where $\bar{\Sigma}(r < R)$ is the average density in a region inside R, that is

$$\bar{\Sigma}(r < R) = \frac{\int_0^R 2\pi r \mathrm{d}r \Sigma(r)}{\int_0^R 2\pi r \mathrm{d}r}$$
$$= \frac{2}{R^2} \int_0^R \mathrm{d}r\, r\Sigma(r) \quad (2.36)$$

In practice gravitational lensing is used to calibrate scaling relations between masses and mass proxies [Allen et al., 2011]. This is mostly because the depth in magnitude needed for deriving accurate masses for each cluster from gravitational lensing is large.

The measurement of luminosity and masses can also be combined to the masses, as M/L (mass-to-light) ratios. Since it is known that reddish K-stars that populate elliptic galaxies have mass-to-light ratios of the order $(M/L)_K \sim 3(\mathrm{M}_\odot/\mathrm{L}_\odot)$, it is interesting to compare to the mass-to-light ratio of these galaxies. The result, known as early as 1933 [Zwicky, 1933] is that the mass-to-light ratios in clusters is

$$\left(\frac{M}{L}\right) \sim 300h \left(\frac{\mathrm{M}_\odot}{\mathrm{L}_\odot}\right), \quad (2.37)$$

which displays a discrepancy of around two magnitude. This *missing mass problem* originated the idea of dark matter, an ingredient today considered crucial to explain the large scale structure of the universe and many derivative effects observed in modern surveys.

2.5 The Accretion History of Galaxy Systems

As previously discussed, the density fluctuations from the beginning of the Universe eventually merge and create larger structures in later times, in the hierarchical formation scenario [Peebles, 1984]. How this process impacts halo characteristics and how can different histories of mergers be probed by observation is important both to understand galaxies and galaxy systems, and the information they provide for cosmology.

The study of this mass accretion history from theoretical arguments is carried mostly through cosmological simulations [Bullock et al., 2001, Wechsler et al., 2002, Dariush et al., 2007, eg.]. The universality of the halo mass profile, which we have previously mentioned (Eqs. 2.6, 2, etc.) offer ways to investigate the relationship between the shape of the mass distribution and past accretion history of a halo. Here we are mostly concerned with this: given an observed system, has it been mostly formed for some time before the observation, or does it have a recent history of mergers/mass accretion?

Because inhomogeneities grow continuously since the early stages of the Universe, we have to define a formation (or collapse) time for a halo based on mass accretion rates. The easiest way to do this is to define that a system is "formed" when it has accreted a fraction (usually half) of its mass. This is of course arbitrary[17], and there are other definitions. Bullock et al. [2001] define the collapse time as the moment in which half of the halo mass was in progenitors more massive than f times the halo mass. Another way is to define a typical collapse mass (M_*) at a redshift z, given such that $\sigma[M_*(z)] = 1.686/D(z)$, where $D(z)$ is the linear growth function at the redshift - and then label a collapse time z such that $M_*(z) = FM_{vir}$, that is, a fixed fraction F of the current observed mass [Wechsler et al., 2002].

[17] And can lead to problems - a system that has accreted very recently 45% of its mass close to the time of observation, but was undisturbed for a long time before, would still count - but is probably very different, dynamically.

2.5.1 Magnitude Gaps as probes of Mass Accretion

The parametric profile proposed by Navarro et al. [1995] has two parameters: a scale radius R_S, roughly where the profile changes

from an r^{-2} to r^{-3} behaviour, and a characteristic density ρ_s. We have mentioned how can we define a mass M_{200} by imposing another radius R_{200} that defines a region where the halo has $\Delta_c = 200$ times the background density of the universe. Then, a useful alternative parameter is the ratio between $c_{200} = R_{200}/R_s$, called the *halo concentration*. In their study Navarro et al. [1995] found that, for a given cosmology, the characteristic density is inversely proportional to the scale radius - which is to say that c_{200} increases with decreasing M_{200}[18]. This follows from the fact that, in the hierarchical scenario, low mass halos form earlier, when the background density of the universe was higher - and the characteristic density reflects this.

[18] This is called the $c - M$ relation and will be revisited throughout this work.

This toy model cannot, however, explain the scatter observed in the $c - M$ relation [Bullock et al., 2001, Jing, 2000]. As halos of a given mass collapse in different times, it is then suggested that this scatters their concentrations. Bullock et al. [2001] have shown that concentrations at a given mass, for example, are lower in higher redshifts. Using the definition by Wechsler et al. [2002], they show that the characteristic density of a halo grows when the mass accretion rate is high, and approaches a constant when it is low. Because of this, the concentration of a halo evolves as $c = c_1 a_o/a_c$, where a_o is the scale factor at the observation and a_c is the scale factor at the collapse.

Other probes of the history of mass accretion have been proposed [Raouf et al., 2014, 2016]: the luminosity difference between the central galaxy and the brightest satellite (called the *magnitude gap*) and the offset between the central galaxy and the center of the system. Of these, the magnitude gap is by far the most used, as it is relatively easy to observe, and we delve into now.

The magnitude gap is quantified by the difference in magnitude of the CG and the brightest satellite inside some fraction of R_{200}. In Vitorelli et al. [2018], that is fully presented and expanded in some details here in chapter 4, we have shown that groups with larger magnitude gaps tend to be more concentrated, as it would be expected by the arguments above. Difficulties in measuring magnitude gaps properly are: correct identification of the central galaxy by the cluster finder, correct identification of member

galaxies, and flux limitation, that inevitably misses very faint galaxies and will bias the survey near the redshift limit of a survey. Galaxy systems with more extreme magnitude gaps are usually labeled *fossil groups* or *fossil clusters*. The first identification of such a group was made by Ponman et al. [1994], when they suggested that the system RX J1340.6 + 4018 was possibly the relic of a former group of galaxies. At first, it was thought to consist of a single galaxy with an extended X-ray halo, but it was later shown [Jones et al., 2000] to be actually a group of about $N \sim 10$ galaxies, out of which the central galaxy accounts for about $\sim 70\%$ of the total optical luminosity, being thus consistent with it being the merger product of the missing L^* galaxies around it.

An empirical definition was given afterwards by [Jones et al., 2003] for such *fossil groups*, and is expressed by

- a high luminosity in X-rays $L_X \geq 0.25 \times 10^{42}$erg s^{-1},

- an absolute magnitude gap between the central, usually most luminous galaxy and the second brightest galaxy greater than $\Delta M_{1-2} = M_{BCG} - M_{2BG} \leq -2$ within half the projected r_{200} radius.

Figure 2.7: NGC1132 is a typical example of fossil group. On the image, circles indicate companion galaxies and squares globular clusters of the central galaxy.

With this definition, the number of such systems is small (about $\sim 2\%$ of all systems), but not negligible, and most studies focused on tens of groups at a time. In time, however, larger systems were discovered [Cypriano et al., 2006] and have been called as fossil clusters, by analogy.

There are three main explanations suggested for fossil groups. The first one is, in line with our discussion of mass accretion, that these are systems that have been left undisturbed for a long time, and the CG had enough time to cannibalise the L^* galaxies in the inner region. The second one is that these are results of random sampling the Schechter luminosity function (eq. 2.28), which for lower richness can produce more extreme results. And finally, that these objects are dark clusters, with either less sub-halos (galaxies) or in which more of these sub-halos are dark.

Dariush et al. [2007] used simulations to show that systems with larger magnitude gaps collapse early, on average, lending support to the fossil hypothesis[19]. Another argument for an early formation is based on observations of fossil systems including

[19] These simulations show that fossil groups accreted on average only 1 further galaxy since $z = 1$, compared to the average of 3 for other groups [von Benda-Beckmann et al., 2008, D'Onghia et al., 2005]

through X-ray scaling relations [Khosroshahi et al., 2007] and morphological studies of their central galaxies [Khosroshahi et al., 2006] . Putting all these together into a coherent picture for the overall population, FGs have been repeatedly suggested (e.g. Harrison et al. [2012], Dariush et al. [2010], Khosroshahi et al. [2007] and many others) to be more relaxed systems and, as such, [Mantz et al., 2015] and Dariush et al. [2010] have also proposed to modify the criterion for the optical selection of fossil groups to $\Delta M_{1-4} \leq -2.5$, finding it to be a more efficient probe of identifying early-formed halos than the conventional definition.

Paranjape and Sheth [2012], on the other hand, have employed extreme value statistics to show that the distribution of galaxy groups as a function of magnitude gap can be consistent with the distribution resulting from a set of random draws from a global luminosity function, implying that the group mass is only related to the magnitude gap through mutual covariance with richness.

Proctor et al. [2011] have suggested, by looking to the mass-to-light ratios of fossil groups, found high dynamical masses (consistent with clusters) and extremely low richnesses. Their central galaxy luminosities were also more consistent with clusters than groups. These led them to suggest that fossil groups are cluster sized halos with very low total luminosities.

The luminosity function of the first fossil group, as shown by Jones et al. [2000], and the conclusion that it resides in a sparse environment, was suggested by D'Onghia and Lake [2004] to pose a problem for the cold dark matter models, since they did not have as much substructure as expected for such massive systems. On the other hand, simulations have shown that the luminosity function of three FGs, including the first identified by Ponman et al. [1994], RX J1340.6 + 4018, are consistent with ΛCDM predictions.

Mulchaey and Zabludoff [1999], analysing NGC 1132, have suggested that fossil groups may consist of *failed groups*, that is, local overdensities in which other bright galaxies never formed. The masses and M/L ratios of the central galaxies of FSs are also usually too large to be explained as end points of compact group evolution driven just by dynamical friction [Voevodkin et al., 2010]. The basic merger interpretation, however, remains viable,

as previously discussed here, in the context of simulations.

Finally, another possibility is that groups and clusters may go into a fossil phase, with an absence of significant mergers, with enough time for relaxation [von Benda-Beckmann et al., 2008]. A bright galaxy, or a group, may then sink to the inner region of the system and the system will return to a normal magnitude gap. Dariush et al. [2010] find that about 90% of fossil groups which were identified according to both criteria in earlier epochs become non fossils after 4Gyr and the fossil phase persists for $\sim$ 1Gyr. Using semi-analytic models based on the Millennium simulation [Springel et al., 2005], Gozaliasl et al. [2014] have shown that 80% of groups ($13 < \log M_{200} < 14$ in M$_\odot$) that would classify as fossil at redshift $z = 1$ lose their large magnitude gaps, but that 40% of the clusters ($\log M_{200} > 14$ in M$_\odot$), on the other hand, retained large gaps.

Our published results on this topic are detailed on chapter 4

2.5.2 *Central Galaxies vs Galaxy System Centres*

The displacement of the central galaxy from the system centre is often suggested as a measure of the relaxation of a galaxy group or cluster. This follows from the expectation that dynamically relaxed systems should approach spherical symmetry [Mohr et al., 1995, Maughan et al., 2008, Mann and Ebeling, 2012]. What we mean by "system centre" is, however, not a straightforward in observations. In X-rays, the emission peaks are used as proxies (eg. Dietrich et al. [2012], Oguri and Marshall [2010]) for the deepest point in the gravitational potential. Ideally, gravitational lensing could give an unequivocal location. In reality, the shape noise of lensing measurements and smoothing of current lensing studies make this not a viable option today [Dietrich et al., 2012].

The luminosity centroid of a system, quantified as the average of member galaxy positions (except the central), weighted by luminosity [20] has been used by Raouf et al. [2014] as proxies of galaxy system centers to evaluate offsets.

[20] ...and membership probabilities, if the cluster finding algorithm evaluates them.

The use of luminosity centroids as proxies of system centres is however, not without difficulties. Pre-conceptions built in cluster finding algorithms play a central role here - as they will impact the handling of superposition effects, selection of central galaxy,

selection of members, etc. Considering more than one group of galaxies as part of the same system can create highly biased systems, with great offsets and higher apparent masses.

Our results on this topic are shown in chapter 5.

Our main scientific goal with this work is to tie together these results to the concentration parameter of the NFW profile on stacks of galaxy systems. This will help us understand observables as they relate to the mass accretion history of galaxy systems and, in turn, characterise observed systems by relaxation state. In simple words, we want to create a method of dating the formation of galaxy systems, to discriminate betwen them by dynamical state. This is important as, for example, cosmology studies done with cluster counts depend on relaxation assumptions that, in reality, correspond to just a fraction of systems[Smith et al., 2003, Puchwein and Bartelmann, 2007, Comerford et al., 2010].

3

Gravitational Lensing

GRAVITATIONAL LENSING IS THE NAME GIVEN TO
THE PHENOMENON OF DEFLECTION OF LIGHT RAYS
BY GRAVITY. It is today one of the best understood and
most remarkable features of the ETG, and was among the first
empirical tests performed to validate it[Eddington, 1920]. The
idea that gravity should act on light precedes Einstein's work
by more than a century in Soldner [1804], and it is related
to the description of light as a stream of luminous particles.
Adopting this point of view, together with the equivalence
principle, Einstein re-derived the deflection angle for light
particles approaching a gravitation source with a given impact
parameter. After the discovery of the field equations (Eq. 1.1) it
was noted that the deflection angle should actually be twice the
classical result.[1]

    This provided a test of the new theory of gravity against the old
Newtonian paradigm. In 1919 a solar eclipse offered the perfect
opportunity: the position of stars in the sky was known with
enough precision to discriminate the predicted deflection [Dyson
et al., 1920] and the result supported Einstein's theoretical results.

    The study of gravitational lensing remained relatively quiet in
the next decades, however. After a running discussion between
Eddington [1920], Chwolson [1924], and Einstein [1936], it was
thought that the phenomena would be an occurrence too rare
to observe - considering only chance alignment of stars. Zwicky
[1933], however, pointed that entire galaxies could display visible
effects on other farther galaxies behind. Furthermore, having
calculated the masses of Virgo and Coma clusters to be higher
than previously thought by no less than 2 orders of magnitude, he

[1] Using the Schwarzschild metric, with r_s being the characteristic radius, and a light ray with impact parameter b, the equation for the total deflection angle is:

$$\hat{a} = \int \frac{dr}{r^2 \sqrt{\frac{1}{b^2} - \left(1 - \frac{r_s}{r}\right) \frac{1}{r^2}}} , \quad (3.1)$$

which yields

$$\hat{a} \approx \frac{2r_s}{b} = \frac{4GM}{c^2 b} . \quad (3.2)$$

The twofold factor is tied to another effect correctly addressed by the theory: the precession of orbits. Both can be understood as results of the non-conservation of the Laplace-Runge-Lenz vector in ETG.

argued that the deflection of light of distant galaxies could provide not only tests of ETG, but also allow the determination of cluster masses. Finally, Zwicky [1937] calculated the lensing probability and concluded that about one percent of distant galaxies should be significantly distorted. Nonetheless, his predictions would wait several more decades until adequate observational technologies brought Gravitational Lensing back to the spotlight.

Refsdal [1964a,b] derived the equations to formalise the practical use of gravitational lensing, and provided a method to estimate the Hubble constant H_0 by measuring time delays between two different lensed images of the same object. With the subsequent discovery of quasars by Schmidt [1963], it was proposed that gravitational lensing by galaxies could be used together with these distant, extremely bright objects to probe the masses of galaxies, and realise Refsdal's ideas. Finally, fifteen years latter, came the first observation of a doubly lensed quasar, by Walsh et al. [1979] (Fig. 3.1).

The coming-of-age for cluster astrophysics, on the other hand, came in the late 70's and early 80's. With that Narayan et al. [1984] explored in detail the possibility of clusters acting as powerful lenses. Finally, [Lynds and Petrosian, 1986] and Soucail et al. [1987] independently discovered images of "giant arcs", which have shown to be strongly distorted images of distant background galaxies near the core of forefront galaxy clusters. This was immediately interpreted by Paczynski [1987] as the effect of gravitational lensing - what would be confirmed by the measurement of the redshift of the arc in Abell 370 [Soucail et al., 1988].

The construction of the space based telescopes, advances in ground based telescopes[2], offer an unprecedented level of precision in astrometry. These steps paved the way for the systematic use of Gravitational Lensing not only as an effect of gravity, but as an integral tool to measure characteristics of the large scale structure of the universe. Furthermore, it empowered the search for very high redshift galaxies and even exoplanets. With these great and diverse results, gravitational lensing is seen today as a tool, in the same fashion as spectroscopy went from a hint of the quantum nature of the atomic structure to a basic technique to investigate
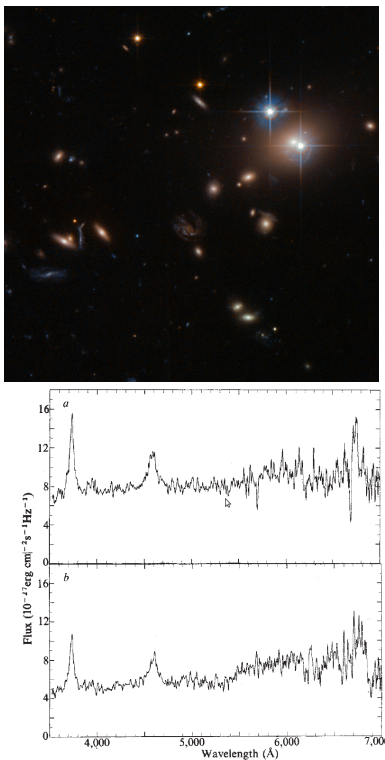
Figure 3.1: **Upper:** The double quasar QSO0957 + 561 depicted by the Hubble Space Telescope in the upper right area of the image as the "star like" objects with diffraction rays. The reddish light between them is the forefront deflector galaxy. Their spectra matches as an object at redshift $z = 1.413$, with the deflector at $z = 0.355$. The separation is of about $6''$. *Source: ESO/NASA.* **Lower:** Reproduction of the spectra of each component, from Walsh et al. [1979]

[2] in particular, adaptive optics

properties of baryonic matter.

3.1 *Gravitational Lensing Theory*

Since the environment of galaxy clusters have small curvatures, the effect of gravity can be approximated by linearisation. The deflection due to an extended mass distribution can be then written as a sum over discrete masses. If the deflection is small, we can approximate the gravitational potential along the deflected trajectory by the potential along the undeflected trajectory.[3] Using a system of coordinates in the plane of the sky $\vec{\xi} - \vec{\xi'} = b$ and the redshift in the line-of-sight z, deflection angle will be given then by a sum of the contributions of each point:

[3] similarly to the Born approximation in quantum mechanics.

$$\vec{\hat{\alpha}}(\vec{\xi}) = \frac{4G}{c^2} \int dm \frac{\vec{b}}{|\vec{b}|^2} = \frac{4G}{c^2} \int \rho dV \frac{\vec{b}}{|\vec{b}|^2} \tag{3.3}$$

$$\vec{\hat{\alpha}}(\vec{\xi}) = \frac{4G}{c^2} \int d^2\xi' \int dz \rho(\vec{\xi'}, z) \frac{\vec{b}}{|\vec{b}|^2}, \vec{b} \equiv \vec{\xi} - \xi' . \tag{3.4}$$

The first integral is calculated over the plane of the sky and the second throughout the line of sight up to the source object redshift. In the limit of a *thin lens*[4], where the distances between the source, lens, and observer are much larger than the size of the deflector, we can use the projected mass density

[4] In the context of cluster lenses, this condition holds very generally, a cluster of galaxies has a scale of few Mpc, whereas the distances of lens systems and their lensed sources are considerable fractions of the Hubble length cH_0, which is about 4.3Gpc.

$$\Sigma(\vec{\xi}) = \int \rho(\vec{\xi}, z) dz , \tag{3.5}$$

so that the deflection can be rewritten as:

$$\vec{\hat{\alpha}} = \frac{4G}{c^2} \int \frac{(\vec{\xi} - \vec{\xi'})\Sigma(\vec{\xi'})}{|\vec{\xi} - \vec{\xi'}|^2} d^2\xi'^2 . \tag{3.6}$$

With the relationship between the deflection angle and the lens mass distribution at hand, what remains is a geometrical problem. The typical situation considered in lensed systems has a deflector at a certain redshift z_d, a set of sources with redshifts z_s. The source and lens planes are defined as perpendicular to the optical axis of the observer (the dashed line in the figure).If $\vec{\eta}$ is the two-dimensional position of the source relatively to this axis, then we can write

Figure 3.2: Gravitational lens geometry. In the thin lens approximation we can consider that all the change in direction of the light rays takes place in a specific plane. Here, all distances used must specifically be angular diameter distances, since the large scale structure of space-time is not euclidean. *Source: ['Bartelmann and Schneider, 2001]*
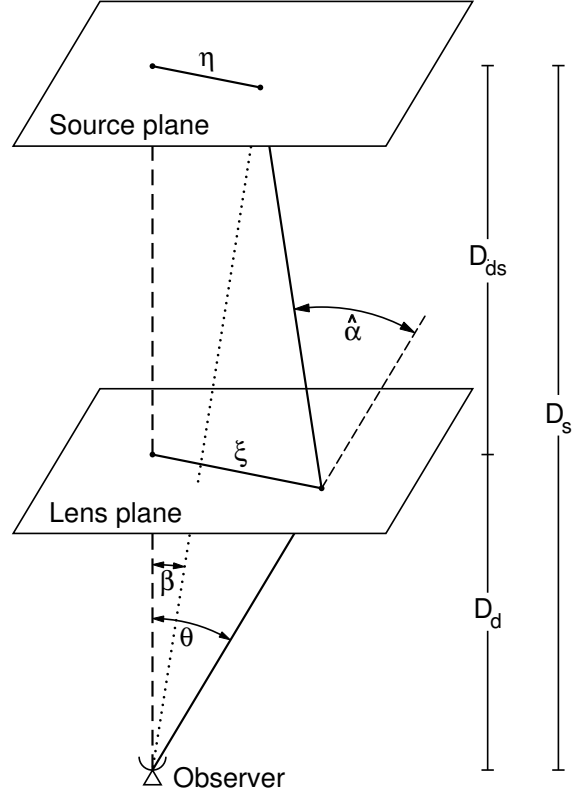
$$\vec{\eta} - \frac{D_s}{D_d}\vec{\xi} - D_{ds}\vec{\hat{\alpha}}(\vec{\xi}) \,, \tag{3.7}$$

which, using angular coordinates $\vec{\eta} = D_s\vec{\beta}$ and $\xi = D_d\vec{\theta}$ can be written as:

$$\vec{\beta} = \vec{\theta} - \frac{D_{ds}}{D_s}\vec{\hat{\alpha}}(D_d\vec{\theta}) \,. \tag{3.8}$$

This is the *lens equation* for the system and means that an object which would be observed at a "true" position in the sky $\vec{\beta} = (\beta_1, \beta_2)$ in its source plane will be seen at another position $\vec{\theta} = (\theta_1, \theta_2)$ in an "image" plane according to the deflection given by $\vec{\hat{\alpha}}$ and the given ratio of angular diameter distances.

In the original case of star images being displaced by the gravity of the Sun, we had both $\vec{\beta}$ and $\vec{\theta}$, since the angular separation of the sun and the stars vary with time. However, deep object configurations do not change significantly and then neither $\vec{\beta}$ nor $\vec{\hat{\alpha}}$ can be directly observed as $\vec{\theta}$, D_d, and D_{ds} can.

The solution to this problem comes by using the lens equation multiple times and constraining them all together to make such

system solvable. The constrain, in turn, will be set by reasonable assumptions about the sources and the deflector.

It is useful to define a function that depends on just the distances of the deflector and some source, called the *critical surface density of the lens* as

$$\Sigma_{cr} := \frac{c^2 D_s}{4\pi G D_{ds} D_d} \, , \qquad (3.9)$$

It can be shown that arrangements of lens-source pairs in which the projected surface density surpasses this critical value, the lens equation will display multiple solutions[5].Then, when the lens equation displays multiple solutions, a single source will be displayed in multiple images around the lens. [6]

Rewriting the angle equation with the critical density, we have:

$$\kappa(\vec{\theta}) = \frac{\Sigma(D_d \vec{\theta})}{\Sigma_{cr}} \, . \qquad (3.10)$$

Inserted to the deflection angle equation, this yields:

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi} \int d^2\theta' \frac{(\vec{\theta} - \vec{\theta'})\kappa(\vec{\theta'})}{|\vec{\theta} - \vec{\theta'}|^2} \, ,$$

which suggests introducing a $2D$ gravitational potential ψ for which

$$\nabla^2 \psi = 2\kappa \qquad (3.11)$$

holds. This describes a line-of-sight integrated two-dimensional "Newtonian" potential, rescaled by Σ_{cr}, called the lensing potential, such that now $\vec{\alpha}(\vec{\theta}) = \vec{\nabla}\psi(\vec{\theta})$ and $\kappa(\vec{\theta}) = \frac{1}{2}\nabla^2\psi(\vec{\theta})$.

The lens equation together with the density-deflection relations define a surjective map from the image plane $\vec{\theta}$ onto the source plane $\vec{\beta}$. The shape of images of extended objects will differ from the shape of their respective sources because the deflection of light ray bundles is locally differential. We can then investigate these local properties of this mapping in small neighbourhoods through the Jacobian matrix, given by:

$$A_{ij} = \frac{\partial \beta_i}{\partial \theta_j} = \delta_{ij} - \frac{\partial \alpha_i}{\partial \theta_j} = \delta_{ij} - \frac{\partial^2 \psi}{\partial \theta_i \partial \theta_j} \, . \qquad (3.12)$$

[5] Note that $\Sigma > \Sigma_{cr}$ is a *sufficient* but not *necessary* condition for multiple solutions:

P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*. 1992. URL http://adsabs.harvard.edu/abs/1992grle.book.....S

[6] This corresponds to a phase transition of the configuration of null geodesics around the observer-lens axis.

Evaluating each term in the Jacobian we can write the transformation matrix explicitly as:

$$A = \begin{bmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{bmatrix} , \qquad (3.13)$$

where we introduced:

$$\gamma_1 = \frac{1}{2} \left(\frac{\partial^2 \psi}{\partial \theta_1^2} - \frac{\partial^2 \psi}{\partial \theta_2^2} \right) \qquad (3.14)$$

$$\gamma_2 = \frac{\partial^2 \psi}{\partial \theta_1 \partial \theta_2} . \qquad (3.15)$$

The reason why we introduced γ is simply because the transformation A can now be expressed as two observationally different effects.

If we combine γ_1 and γ_2 into a complex quantity $\gamma = \sqrt{\gamma_1^2 + \gamma_2^2} e^{2\phi i}$ we can then write:

$$A = (1 - \kappa) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - |\gamma| \begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix} . \qquad (3.16)$$

[7]

The local distortion of the images can now be understood as a *convergence* part κ, that gives us how much an image is enlarged or diminished in area, and the traceless *shear* γ, which quantifies how much an image is laterally distorted.

There are two remarks that should be made about the local properties of the lensing transformations:

- the shear γ is *not* a vector, due to its transformation properties under rotations: the components of the shear are mapped onto themselves (an identity transformation) with a half-rotation. The reason for this behaviour due to it being the traceless part of the jacobian matrix (3.16). Hence, the shear is actually a 2-spinor.

- the convergence κ preserves surface luminosity, due to the *theorem of conservation of étendue*, which can be better understood in Hamiltonian optics as an analogue of Liouville's theorem.[8]

	< 0	> 0
κ		
Re[γ]		
Im[γ]		
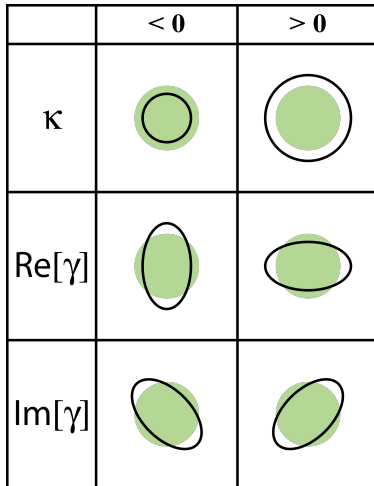
Figure 3.3: The effects of the two components of the lensing transformation. κ changes the size of images while γ distorts images laterally.(*Source: Wikipedia*)

[7] where ϕ is the angle between the shear frame and one coordinate axis

[8] P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*. 1992. URL http://adsabs.harvard.edu/abs/1992grle.book.....S

The ratio between the solid angles of a local neighbourhood in the image plane and its respective "original" source distribution is given by:

$$\mu = \frac{\theta}{\beta}\frac{\mathrm{d}\theta}{\mathrm{d}\beta} = \frac{1}{\det A} = \frac{1}{(1-\kappa)^2 - |\gamma|^2}\,, \qquad (3.17)$$

and is called the *magnification*. There may be regions in a configuration where this determinant is zero, which will correspond to a *critical* curve in the plane of images, and a respective *caustic* in the source plane. This suggests there are areas which have infinite magnification, which evidently cannot be true, which is due to approximations taken in this derivation. Still, in such cases, dramatic arcs and even ring images of background objects appear.[9]

When great distortion and multiple imaging is present, we call this *Strong Gravitational Lensing* and the method for determining masses of deflectors usually involves reconstructions of the source based on the likelihood of the multiple images or giant arcs displayed being mapped to some original configuration - this solves our geometry problem by enabling multiple uses of the lensing equation for different points. The counterpart, *Weak Gravitational Lensing*, refers to situations where light rays from distance sources are just slightly distorted in a coherent fashion. In these configurations, the effect is measured by ensembles of shapes of background objects which provide an estimator for γ as well as κ, which we will see next.

More recently, number counts of background objects have been also used for mass profile measurements. Since the effect of a gravitational lens is to spread out the image of the background sources you could expect a decrease in the projected number density of background galaxies. However, since magnification, as we have seen, enhances the flux of background objects you can also expect an increase of objects in images - which are limited by flux. The balance between these effects can be modelled in comparison to undeflected fields in regions away from cluster centres to determine projected mass distributions.

Finally, as a note for completeness, the study the region in-between the strong and weak regimes - the *flexions* - has also been

[9] A nice example is that presented by a singular isothermal sphere mass distribution. Using the an isothermal velocity distribution [Schneider, 2006] we have:

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G\xi}\,,$$

which yields

$$\kappa(\theta) = \frac{\theta_E}{2\theta}\,,$$

with

$$\theta_E = 4\pi \left(\frac{\sigma_v}{c}\right)^2 \frac{D_{ds}}{D_s}\,,$$

where θ_E is called the *Einstein radius* of the lens. In this case, we have:

$$|\vec{\alpha}| = \theta_E$$
$$\psi(\vec{\theta}) = \theta_E\theta$$
$$\mu(\vec{\theta}) = \frac{\theta}{\theta - \theta_E}\,,$$

so that when $\theta \approx \theta_E$, $\beta \approx 0$, and the magnification is very large as we can see in the incredibly near-perfect alignment below:
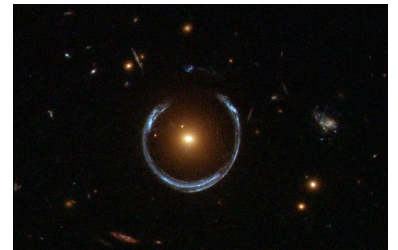
Source: NASA

employed as a complimentary technique to bridge and combine strong and weak lensing for a deeper, more complete analysis of mass distributions through lensing effects.

For strong Lensing, magnification and flexions the reader is recommended to Merten [2010, 2008], Meylan et al. [2006] and others, as we now draw our attention to weak lensing shear analysis, which will be our main tool in the application.

3.2 *Weak Gravitational Lensing*

While the presence of any amount of matter deflects the path of light rays passing around it, this will rarely show as the giant arcs and multiple images usually associated with gravitational lensing. Most regions in the sky are affected by foreground mass distributions very slightly and the deflection is impossible to quantify by a single background source. In such cases, the presence of the foreground mass can be detected by measuring the systematic alignment of background sources around the lensing mass. Weak gravitational lensing is an intrinsically statistical measurement, one that overcomes the minuteness of the effect by combining a large number of individual measurements.

In regions that $\kappa, \gamma \ll 1$, the Jacobian of the lensing transformation - and an estimator for Σ - can be mapped by observing the combined effect of the shear γ on the distribution of projected ellipticities of background galaxies. As field galaxies are (mostly) randomly oriented, systematic alignments can be used to probe the lensing of a foreground source. Even if galaxy shapes are not perfect ellipses, their ellipticities can be measured by fitting models to images, or by measuring the second moments of a galaxy image.

For a given image, if the distribution of surface brightness in a small neighbourhood around a source is given by a function $I(\vec{\theta})$, the centroid of the object can be estimated by the average of the distribution

$$\langle \vec{\theta} \rangle = \frac{\int d^2\theta I(\vec{\theta})\vec{\theta}}{\int d^2\theta I(\vec{\theta})}, \tag{3.18}$$

whereas the second moment tensor is given by

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta I(\vec{\theta}) \left(\theta_i - \langle\theta_i\rangle\right)\left(\theta_j - \langle\theta_j\rangle\right)}{\int \mathrm{d}^2\theta I(\vec{\theta})}\,. \qquad (3.19)$$

If we define the ellipticity of the object as[10]

$$\epsilon := \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}\,, \qquad (3.20)$$

we can write then the second moment tensor of the unlensed source with the observed tensor Q and the transformation A as

$$Q^s = AQA^T = AQA\,, \qquad (3.21)$$

which gives us the original ellipticities of the source as

$$\epsilon^s = \begin{cases} \dfrac{\epsilon - g}{1 - g^*\epsilon}, & |g| \leq 1 \\[2em] \dfrac{1 - g\epsilon^*}{\epsilon^* - g^*}, & |g| > 1 \end{cases}$$

with the *reduced complex shear* g given by

$$g(\vec{\theta}) = \frac{\gamma(\vec{\theta})}{1 - \kappa(\vec{\theta})}\,. \qquad (3.22)$$

Now, the ellipticity expression can be inverted to give

$$\epsilon = \begin{cases} \dfrac{\epsilon^s + g}{1 + g^*\epsilon^s}, & |g| \leq 1 \\[2em] \dfrac{1 + g\epsilon^{s*}}{\epsilon^{s*} + g^*}, & |g| > 1 \end{cases}\,. \qquad (3.23)$$

If we consider many galaxies in a small enough neighbourhood of the image, using our ansatz that the intrinsic ellipticity ϵ^s is randomly distributed we have $\langle\epsilon^s\rangle = 0$. Thus, the averaged

[10] Ellipticities are usually defined by the ratio

$$\varepsilon := 1 - q = 1 - \frac{b}{a}\,,$$

where b/a is the axis ratio of the ellipse. In gravitational lensing it is customary to define quantity

$$\epsilon := \frac{1 - q}{1 + q}e^{2i\phi}\,,$$

where ϕ is the same we have seen before in the shear definition - ellipticity, as shear, is transformed into itself by a half-rotation. Joining this with the elliptical parameters

$$Q_{11} = a^2\cos^2\phi + b^2\sin^2\phi$$
$$Q_{22} = a^2\sin^2\phi + b^2\cos^2\phi$$
$$Q_{12} = (a^2 - b^2)\sin\phi\cos\phi\,,$$

gives the ellipticity definition a practical meaning.

measured ellipticity should be

$$\langle \epsilon \rangle = \begin{cases} g, & |g| \leq 1 \\ \\ \dfrac{1}{g^*}, & |g| > 1 \end{cases} . \tag{3.24}$$

So, in the weak lensing limit the averaged ellipticities in a small region provides an estimator for the local reduced shear g. However, g is a single quantity, but κ and γ are two and κ is the quantity most directly related to the mass distribution, through $\Sigma = \kappa \Sigma_{cr}$. Any transformation of the form

$$1 - \kappa' = \lambda(1 - \kappa)$$
$$\gamma' = \lambda \gamma$$

leaves g unaltered. This is the so called *mass-sheet degeneracy*, because it amounts to the fact that a sheet of uniform surface density does not produce any lensing effects. This is a serious problem to the calculation of masses of individual clusters and can be dealt with in some ways, the most straightforward of them is to consider magnification effects. Since the quantity

$$\mu = \frac{1}{(1 - \kappa)^2 - |\gamma|^2}$$

is *not* invariant through the aforementioned transformation, it can be used to calibrate the shear/convergence-reduced shear relation.

Another useful way to constrain the degeneracy, used in cluster surveys and in this work, is to assume a spherical distribution of the lens mass, this immediately breaks the degeneracy by introducing a model for the line-of-sight mass distribution, with a one-to-one correspondence between κ and γ.

3.3   *Parametric Modelling*

The universality we previously discussed of the halo mass profile provides us a method to use gravitational lensing to investigate observables of galaxy systems. An axisymmetric matter

distribution can be described by a radial function $\Sigma(\vec{\theta}) = \Sigma(\theta)$ so that the deflection angle yields

$$\vec{\alpha}(\vec{\theta}) = \frac{\vec{\theta}}{\theta^2} 2 \int_0^\theta \mathrm{d}\theta'\theta'\kappa(\theta')$$

$$\hat{\vec{\alpha}}(\vec{\xi}) = \frac{\vec{\xi}}{\xi^2} \frac{4G}{c^2} 2\pi \int_0^\xi \mathrm{d}\xi'\xi'\Sigma(\xi')$$

$$= \frac{4GM(<\xi)}{c^2\xi} \frac{\vec{\xi}}{\xi},$$

where $M(<\xi)$ is the total mass enclosed inside radius ξ.[11] Since all directions $\vec{\beta}$, $\vec{\theta}$, and $\hat{\vec{\alpha}}$ are collinear we can write the lens equation with scalars

$$\beta = \theta - \alpha(\theta)\,, \tag{3.25}$$

where the deflection angle is given by

$$\alpha(\theta) = \frac{m(<\theta)}{\theta} = \bar{\kappa}(<\theta)\theta\,, \tag{3.26}$$

where again $m(<\theta)$ is the dimensionless mass enclosed inside the radius θ and $\bar{\kappa}(<\theta)$ is the averaged convergence in the same region. The Jacobian matrix can be then calculated rewriting the lens equation as

$$\vec{\beta} = \left[1 - \bar{\kappa}(<|\vec{\theta}|)\right]\vec{\theta}, \tag{3.27}$$

then applying the definition (eq. 3.16) we arrive at

$$A(\theta) = [1 - \bar{\kappa}(<\theta)]\,\mathbb{1} - \frac{1}{\theta}\frac{\mathrm{d}\bar{\kappa}}{\mathrm{d}\theta} \begin{bmatrix} \theta_1^2 & \theta_1\theta_2 \\ \theta_1\theta_2 & \theta_2^2 \end{bmatrix}. \tag{3.28}$$

If we now transform θ into a polar coordinate system around the centre of the distribution, we can write $\vec{\theta} = \theta(\cos\phi, \sin\phi)$ which, together with the former presentation of the jacobian matrix in eq. 3.16 yields:

$$\gamma(\theta) = (\bar{\kappa}(<\theta) - \kappa(\theta))\,e^{2\phi i}\,. \tag{3.29}$$

If a given ellipticity γ is rotated by ϕ the result amounts to a multiplication $\gamma e^{-2\phi i}$, so we can define a polar coordinate

system for the ellipticities such that $\gamma_p := \gamma e^{-2\phi i}$. Like the Cartesian ellipticity, the polar ellipticity admits a separation in two components which are the tangential γ_t and cross $\gamma_\times$ components that can be calculated by:

$$\gamma_t = -\Re\left[\gamma_p\right] = -\left[\gamma_1 \cos(2\phi) + \gamma_2 \sin(2\phi)\right] \tag{3.30}$$

$$\gamma_\times = -\Im\left[\gamma_p\right] = -\left[\gamma_1 \cos(2\phi) - \gamma_2 \sin(2\phi)\right] \tag{3.31}$$

Tangential alignment   Cross alignment

$\epsilon_t > 0$ $\epsilon_t < 0$ $\epsilon_t = 0$ $\epsilon_t = 0$

$\epsilon_\times = 0$ $\epsilon_\times = 0$ $\epsilon_\times > 0$ $\epsilon_\times < 0$
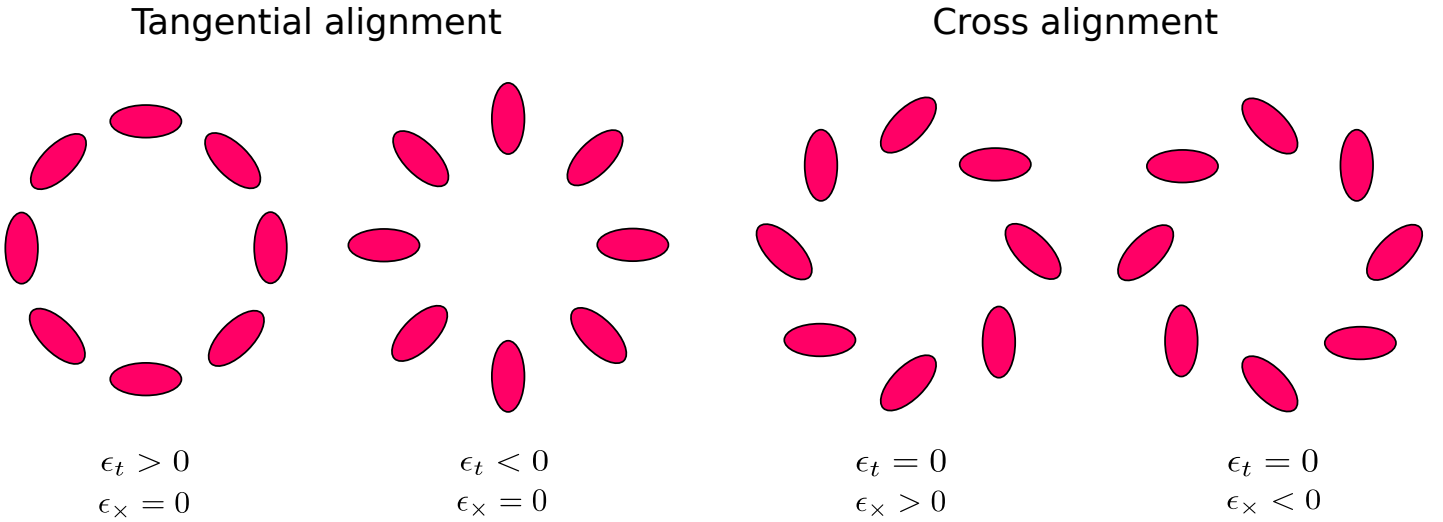
Figure 3.4: Considering a spherical symmetry, we can write the ellipticity of an object as a sum of a tangential component and a cross component. *Source: Mirian Castejon Molina*

The negative sign in this equation can be understood as follows: consider a circular mass distribution and a point on the θ_1 axis outside the Einstein radius. The image of a circular source will be mapped into a stretched image along the θ_2 axis. In this case, $\phi = 0$, the shear is real and negative and in order to have tangential shear positive, and thus to define the tangential component according to intuitive understanding, this minus sign is introduced. Owing to the spherical symmetry, we expect only the tangential projection of γ_t effects from gravitational lensing, whereas $\gamma_\times$ can be used to investigate possible systematic effects.

Now, using that

$$\gamma_t(\theta) = \bar{\kappa}(<\theta) - \kappa(\theta), \tag{3.32}$$

we can rewrite this equation as a more convenient form:

$$\Sigma_{cr}\gamma_t(\theta) = \bar{\Sigma}(<\theta) - \Sigma(\theta), \tag{3.33}$$

which will suit better to practical cases when we combine clusters into stacked halos, because since each lens-source pair z_d and z_s is well determined, we can use the left hand side of the equation (in the limit that $\gamma \sim g$) as data and treat the right hand side as model.

Effectively, we will have many sources with different gravitational signals: a source point contribution due to baryonic mass concentration of the central galaxy, the halo profile, effects due to the miscentring of the profile, and a contribution from the large scale structure of the universe. The proper treatment for each of these effects will be discussed in detail in chapter 4, but all of them will always be modelled into a surface distribution Σ and into effect by the equation above (3.33).

Finally, to measure tangential ellipticities on large areas, we have to address the proper conversion from g_1, g_2 types of ellipticities to g_t and $g_\times$. On small patches of the sky, it is straightforward to convert from both systems, as the angle ϕ can be calculated by transformation from image coordinates, in pixels (x,y) to r, ϕ when a centre is defined. However, on larger scales, we need to take into account spherical geometry.

Let us consider, as an example, a lens centered on some right ascension α and declination δ. We want to calculate ϵ_t of a galaxy, having ϵ_1 and ϵ_2, and its position α_g, δ_g. Therefore, we need to use Equation 3.30, and to do so, we need the angle ϕ. Now, suppose our coordinate system for ϵ is aligned with α[12]. Then the separation between the two can be well approximated by Vincenty's formula:

Figure 3.5: We need to find not only separations between the source galaxy and the cluster center in the sky, but also the angle between their separation grand circle and the right ascension line, to have our tangential ellipticities.

[12] that is, positive ϵ_1 corresponds to elongation in the same direction of the constant right ascension line

$$\rho = \arctan \frac{\sqrt{(\cos\delta_2 \cdot \sin(\Delta\alpha))^2 + (\cos\delta_1 \cdot \sin\delta_2 - \sin\delta_1 \cdot \cos\delta_2 \cdot \cos(\Delta\alpha))^2}}{\sin\delta_1 \cdot \sin\delta_2 + \cos\delta_1 \cdot \cos\delta_2 \cdot \cos(\Delta\alpha)} .$$

If we then take a third point (α, δ_1), we can use Vincenty's formula between all three and apply the spherical law of sines[13] to have polar (ρ, ω) coordinates for our galaxy.

[13] Suppose we find, for our three points, σ_1, σ_2, σ_3 separations. Then, the azimuthal angle between σ_1 and σ_2 will be

$$\cos\omega = \frac{\cos\sigma_3 - \cos\sigma_1 \cos\sigma_2}{\sin\sigma_1 \sin\sigma_2} \quad (3.34)$$

Part II

Application

4

Magnitude Gaps & Concentrations of Galaxy Systems

The following chapter is intended to be an exposition of the work published in Vitorelli et al. [2018]. Here, we intend not to quote the published work verbatim, but to expand on details of some of the procedures taken and tests that built confidence in the results.

4.1   *Introduction*

As we have previously seen on section 2.5.1, the study of the relationship between the central galaxy (CG) and its host system properties is important to understand the evolution of these structures and their value as cosmological probes [Dubinski, 1998, von der Linden et al., 2007, Kravtsov and Borgani, 2012]. The predominance of the CG over the system can be represented in several ways, but is most commonly measured by the difference in magnitude of the CG, which is usually the most luminous, to the second brightest in a central region (BSG - Brightest Satellite Galaxy).

The Jones et al. [2003] criteria single out extreme cases of systems where satellite galaxies are orders of magnitude less luminous than the CG. The extended diffuse X-Ray halo, in turn, indicates a potential well which would be expected to be filled with more bright galaxies.

To recall, these are essentially the points made about fossil groups and clusters in the past:

- They might be indeed fossil systems, that is, end products of galaxy groups and clusters evolution, where the central galaxy has absorbed all central bright galaxies. [Jones et al.,

2003, Khosroshahi et al., 2004, 2006, 2007, D'Onghia et al.,
2005, Dariush et al., 2007, von Benda-Beckmann et al., 2008,
Dariush et al., 2010, Raouf et al., 2016]

- They could be statistical fluctuations that are more likely
  to happen with low richness systems, or just have different
  formation histories.[Hearin et al., 2013, Paranjape and Sheth,
  2012, Kundert et al., 2017]

- They could also be *dark clusters*, that is, systems that have masses
  in line with galaxy clusters but have low richness.[Mulchaey and
  Zabludoff, 1999, Voevodkin et al., 2010, Proctor et al., 2011,
  Harrison et al., 2012]

- All in all, they span from groups to clusters [Mendes de Oliveira
  et al., 2006, Cypriano et al., 2006, Mendes de Oliveira et al.,
  2009], and have similar characteristics to other galaxy clusters
  (Fossil Group Origins - FOGO - Aguerri et al. [2011], Méndez-
  Abreu et al. [2012], Girardi et al. [2014], Zarattini et al. [2014,
  2015], Kundert et al. [2015], Zarattini et al. [2016])

In order to better understand the relationship between the
magnitude gap and formation history, we have investigated its
relation to the mass distribution, using cross-correlation (also
called *stacking* of clusters) weak-lensing methods with parametric
mass profiles. Previous analysis both in observations [Khosroshahi
et al., 2007] and simulations [Deason et al., 2013] indicate
that larger magnitude gap samples will have, on average, more
concentrated systems, as is expected for the lower-mass earlier
formed systems [Navarro et al., 1997, Wechsler et al., 2002].
However, previous observational studies measured concentrations
on a handful ($n \sim 10$) of systems at a time. We want to go further
by using a survey with a large number of systems to study the
average concentration in populations ranked by ΔM_{1-2}.

Instead of focusing on optically selected Fossil Systems
specifically as a separated population, we define ranked *stacks* by
magnitude gaps, after dividing our whole population in redshift
slices, and measure their mass concentrations through stacking
their weak lensing signal, using a parametric profile. Throughout
this chapter we use the standard Flat ΛCDM cosmology with
$h_{100} = 0.7$, $\Omega_m = 0.3$, when h is not made explicit.

4.2 *Data*

The data used in this work come from two different surveys: a catalogue of galaxy clusters built by the Rykoff et al. [2014] algorithm on SDSS images [Annis et al., 2014] and the CFHT Stripe-82 survey (CS82; Moraes et al. [2014]). The so called Stripe-82 is an equatorial region about 2° wide in latitudes between $-40° < \text{RA} < 60°$, which has been extensively investigated in many different bands from various instruments [Viero et al., 2014, LaMassa et al., 2013, Durret et al., 2014]. The CS82 was then specifically designed to take profit of the synergies in this abundance of data, adding shear measurements for weak lensing analysis of the large scale structure of the universe.

4.2.1 *The* REDMAPPER *catalogue*

The catalogue of galaxy clusters and groups used was obtained by the REDMAPPER (**red**-sequence **M**atched-filter **P**robabilistic **Per**colation) algorithm for cluster identification. redMaPPer is a photometric cluster finding algorithm based on the optimised richness estimator λ of Rykoff et al. [2012], which is designed be related with cluster masses with minimum scatter.

The REDMAPPER algorithm identifies galaxy clusters as overdensities of *red-sequence galaxies* around central galaxy candidates. First it uses spectroscopic training sets to characterise the evolution of the red sequence as a function of redshift. It then uses the resulting red sequence model, together with a radial aperture filter and a luminosity function filter based on the Schechter function to estimate the probability that any given observed galaxy belongs to some cluster. The cluster richness is then defined as the sum of the probabilities of galaxies considered, as an estimator for the expected value of the number of galaxies in the cluster,

$$\lambda = \sum_i p_i. \tag{4.1}$$

In addition, by identifying a red-sequence for each cluster, it can estimate a cluster redshift by simultaneous fitting all possible member galaxies to its red-sequence model. The aperture filter

also defines a percolation radius R_c that is related to the obtained richness λ by

$$R_c = 1 \text{ Mpc} \times \left(\frac{\lambda}{100}\right)^{\beta}, \qquad (4.2)$$

with $\beta = 0.2$ [Rykoff et al., 2014, 2012].

We use this radius to define the central region of the system as $R < 0.3 R_c$, which translates into about $\sim 0.4 R_{200}$ [Rykoff et al., 2012]. We have attempted more restrictive radii considerations, to compensate for the large errors in magnitude gaps (due to the probabilistic nature of the cluster/group identification) but we found no qualitative differences in the final results. However, it should be noted that the number of fossils change as shown in figure.
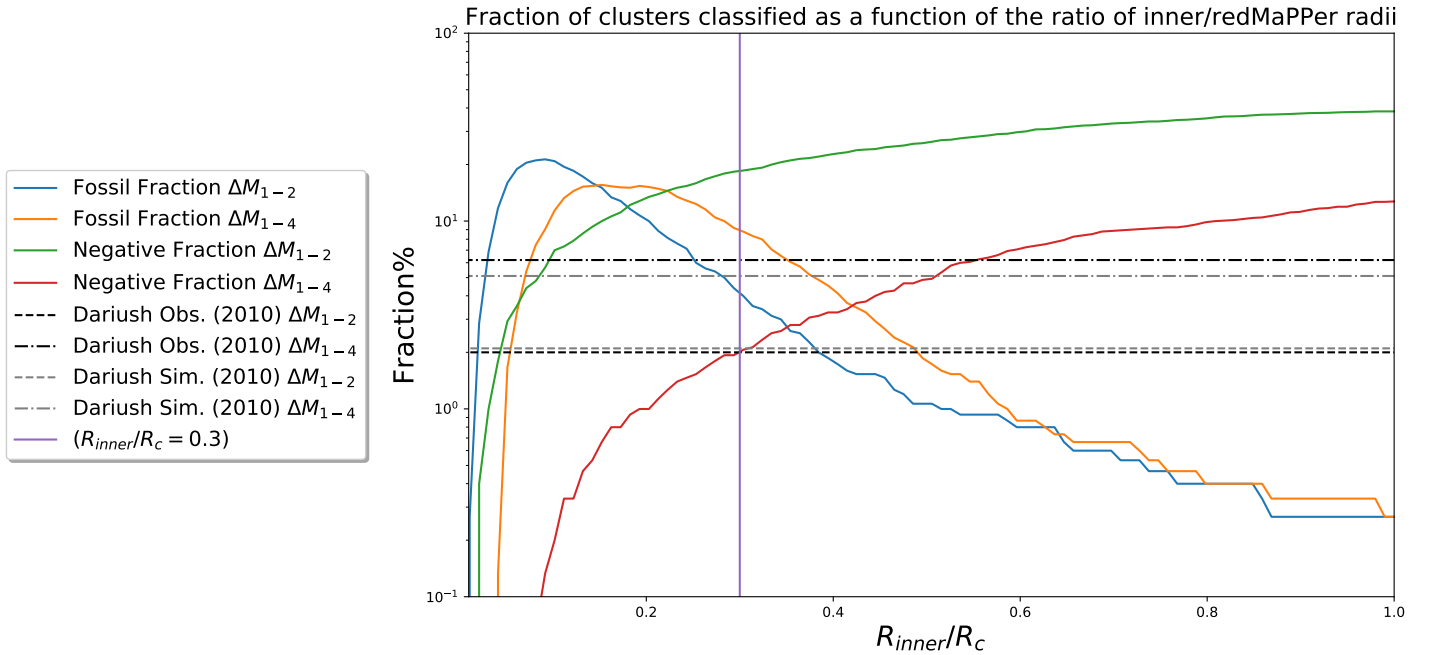
Figure 4.1: The different census of fossils/non fossils and systems with non-brightest central galaxies, as a function of the ratio between the central radius and the redMaPPer defined R_c for magnitude gaps in the r band. For the way we calculate magnitude gaps, please refer to the text below.

4.2.2   CS82 Data

The CS82 survey consists of 173 pointings of the MegaCam instrument, using the $i.MP9702$ filter ($\sim$ SDSS i band). The MegaCam is a large (1deg^2) field of view camera with an angular scale of 0.187arcsec/pixel. The completeness magnitude limit achieved was $m_i < 24$, with an excellent median seeing of $\sim 0.6''$. The image reduction process profits from the CFHTLS pipeline.
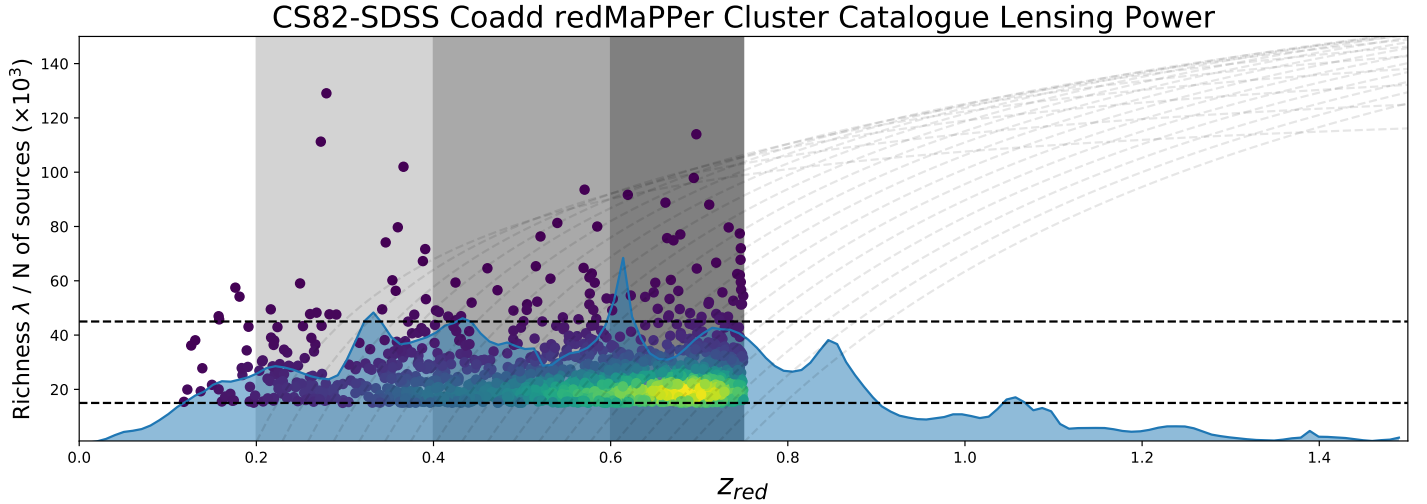
Figure 4.2: In purple to green the dots indicate system photometric redshift and richness distribution both given by REDMAPPER. The gray bands behind are the three redshift slices we consider, and the horizontal dashed lines are the richness selection we restrict in this work. Overlayed, we put the distribution of CS82 galaxy sources, in light blue, together with dashed light gray curves that are in the shape of weak lensing power as a function of the distance between the source galaxy and the lensing cluster. In this picture we see that although high redshift systems are way more numerous, their weak-lensing signal is much lower.

The total effective area after masking and de-overlapping the images corresponds to about $124 \deg^2$ of the sky. This magnitude limit is defined as a safe limit to guarantee homogeneity for all the 173 tiles, with an average galaxy density per image area of $\sim 10 \mathrm{gal/arcmin}^2$.

The classification and measurements of shapes of objects has been done by the LENSFIT algorithm [Miller et al., 2007, Kitching et al., 2008, Miller et al., 2013], and the details of the calibration and its systematics are discussed in Erben et al. [2013]. In this work, all objects with magnitudes $i_{AB} < 23.5$, with LENSFIT weight[1] $w > 0$ and FITCLASS[2] $= 0$ are used. Together, this criteria result in a total of $4,450,478$ galaxies used for weak-lensing, distributed as in Figure 4.2. The lensing products of this survey have been used previously in published results by Shan et al. [2014], Liu et al. [2015], Shan et al. [2015], Leauthaud et al. [2017], Niemiec et al. [2017], Li et al. [2014, 2016], Hand et al. [2015], Pereira et al. [2018], Niemiec et al. [2017].

Individual photometric redshifts for galaxies that were used in the shear analysis were taken from Bundy et al. [2015], and were calculated from SDSS photometric data by applying the BPZ algorithm of Benítez [2000], the redshift distribution of sources is overlayed into Figure 4.2. We present here briefly the ellipticity distribution of galaxies in Figure 4.3

[1] The LENSFIT weight is defined in Equation 6.2

[2] FITCLASS is a star/galaxy identification parameter where 0 corresponds to galaxies and 1 corresponds to stars.

4.3 Method

In this section specially, we expand on [Vitorelli et al., 2018] to include nuanced detail about the selection of methods and objects in this study.

4.3.1 Determination of Magnitude Gaps
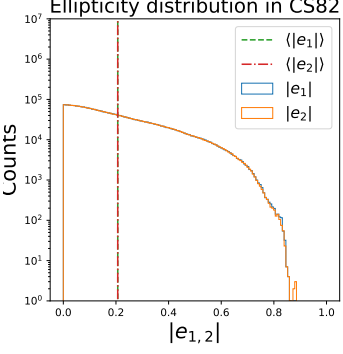
Figure 4.3: The distribution of ellipticity measurements of the CS82 galaxies.

The calculation of magnitude gaps require that we address two important particularities of the redMaPPer catalogue: first the algorithm does not always identify the central galaxy as the brightest, which is in agreement with the literature [Hoshino et al., 2015, Skibba and Macciò, 2011] that central galaxies need not to be the brightest. Therefore, we refer to the central galaxy as CG, instead to the literature usual BCG, and the magnitude gaps are calculated between the most probable central galaxy and the brightest satellite galaxy ($\Delta M_{1-2} := M_{BSG} - M_{CG}$). Secondly, redMaPPer calculates *probabilities of memberships* for galaxies - that is, we should not calculate ΔM_{1-2} straightforwardly by subtracting magnitudes of the CG and the possible BSG in the catalogue, since we do not know with absolute certainty which galaxies are system members. Instead, we calculate *expected values* of magnitude gaps, by simply computing the the expected value from the *probability distribution* of the probable members catalogue of each system.

Our first step is to calculate absolute magnitudes for the galaxies. Since redMaPPer galaxies are already de-reddened[3], we proceed to calculate the absolute magnitudes by calculating the Distance Modulus μ by:

[3] That is, extinction by foreground dust is already taken into account.

$$\mu = 5 \log \left(\frac{D_L}{10\text{pc}} \right) = m - M + k_i \,, \qquad (4.3)$$

where D_L is the luminosity distance in parsecs, and the k-correction as given by O'Mill et al. [2011] for the band i magnitude, using the other filters g, r, and z:

[4] To calculate the uncertainties, we use uncertainty propagation:

$$\delta_{k_i}^2 = \left(\delta_g \frac{\mathrm{d}k_i}{\mathrm{d}g} \right)^2 + \left(\delta_r \frac{\mathrm{d}k_i}{\mathrm{d}r} \right)^2 + \left(\delta_z \frac{\mathrm{d}k_i}{\mathrm{d}z} \right)^2$$

With:

$$\frac{\mathrm{d}k_i}{\mathrm{d}g} = 0.538z - 0.027$$

$$\frac{\mathrm{d}k_i}{\mathrm{d}r} = -0.538z + 0.027$$

$$\frac{\mathrm{d}k_i}{\mathrm{d}z} = 0.538(g - r) - 0.075$$

$$k_i = (0.538(g - r) - 0.075)z + (-0.027(g - r) - 0.120) \quad (4.4)$$

4
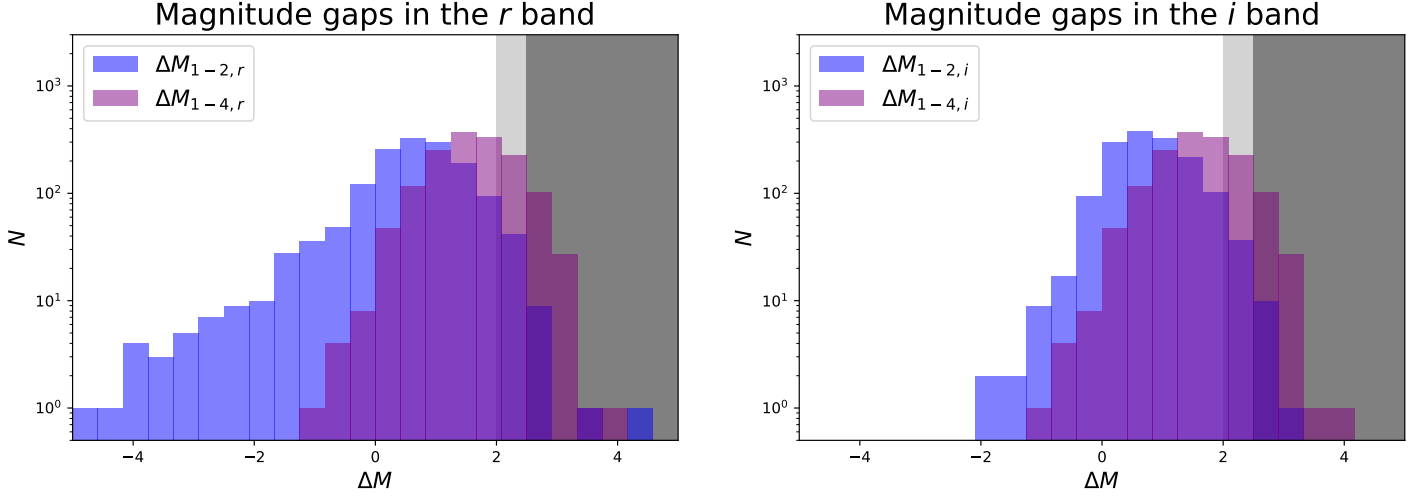
Figure 4.4: The distribution of magnitude gaps in r and i bands, for both ΔM_{1-2}, ΔM_{1-4}. Light gray area are ΔM_{1-2} optical fossil candidates, dark gray area per Dariush et al. [2010] ΔM_{1-4} criterion.

With these, we find some galaxies that are extremely blue, but these cases will be addressed by some of the following procedures.

Now, find the expected value for the magnitude gap, we consider a list of all galaxies inside an inner region (as previously mentioned) of the system except for the central, with N redMaPPer-identified possible members, so that their membership probabilities are given by $\vec{p} = p_1, p_2, p_3...$ ordered by decreasing brightness (increasing absolute magnitude). Then, the probability of the first galaxy to be the brightest in this group is just p_1 as no other galaxy can possibly be brighter than it, So that $p_1 = P_1$ is also the probability of the magnitude gap ΔM_{1-2} to be $M_{CG} - M_1$. Now, for the second galaxy we need to ensure that the first is not present $(1 - p_1)$ *and* that the second is, that is, $(1 - p_1)p_2$. By simple iteration, the probability of the n-th galaxy of being the BSG is given by:

$$P_n = p_n \prod_{i=2}^{n-1} (1 - p_i) \,. \qquad (4.5)$$

Collecting these results the expected values are the sum weighted by individual probabilities for each possible gap:

$$E\left[\Delta M_{1,2}\right] = M_{CG} - \sum_{n=1}^{N} P_n M_n \,, \qquad (4.6)$$

and the errors adopted follow directly from the definition of the

variance also,

$$\mathrm{Var}\left[\Delta M_{1,2}\right] = E\left[\left(\Delta M_{1,2}\right)^2\right] - \left[E\left(\Delta M_{1,2}\right)\right]^2 . \qquad (4.7)$$

We have also tested a criterion proposed by Dariush et al. [2010] as the difference in magnitude between the CG and the third brightest satellite galaxy ($3SG$) that we label as ΔM_{1-4}. The result of both, in the r and i bands are in the Figure 4.4. We tested this alternative throughout, but reached the same conclusions with the most common definition of magnitude gaps.

We have checked those systems where either the magnitude gap is reversed or have abnormal $r - i$ member galaxy colors and found, by visual inspection, that at least some of these cases correspond to either systems with foreground galaxies or fake detections.

Figure 4.5: Distribution of ΔM_{1-2} in both i and r filters of all systems in the catalogue. We discard all those in which the central galaxy is not the brightest for both filters. That also eliminates all systems with galaxies with anomalous colours. The blue shaded area are optical fossil candidates in both filters.
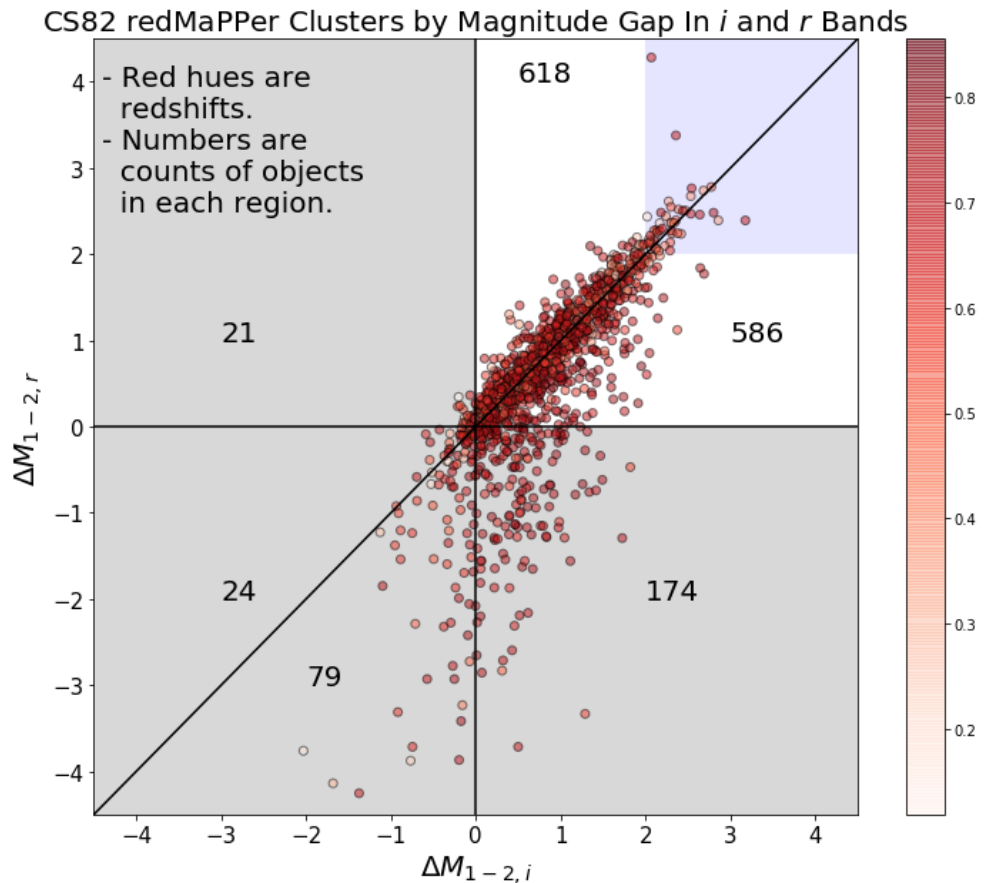
In conclusion, we use $\Delta M_{1-2,r}$ magnitude gaps in the r band excluding galaxy systems in which the central galaxy is not the brightest.

4.3.2 *Parametric Modelling of Mass Profiles*

We will now develop the details of the parametric mass profile model, reviewing previous work from several groups [Mandelbaum et al., 2008b, Johnston et al., 2007, Ford et al., 2015, Shan et al., 2015] to make any model choices explicit. Parametric modelling has the advantage of breaking the mass-sheet degeneracy through postulating spherical symmetry, which, even if not true in many galaxy systems, will be a good description for stacked data taken from randomly oriented systems.

The main component of the mass distribution of a galaxy system is the radial profile of the dark-matter dominated halo, which has been shown to follow roughly a universal form [Huss et al., 1999, Katz, 1991, Cole and Lacey, 1996, Navarro et al., 1997]. Among several proposed expressions for the radial mass density, as we have previously mentioned, we use the simulation-derived result by Navarro et al. [1997], given by Equation 2.6, which we will write as:

$$\rho_{\mathrm{NFW}}(r) = \frac{\delta_{\mathrm{NFW}}\,\rho_{\mathrm{crit}}(z)}{(r/r_{\mathrm{s}})(1 + r/r_{\mathrm{s}})^2}\,, \qquad (4.8)$$

with

$$\delta_{\mathrm{NFW}} = \frac{\Delta_c}{3}\frac{c_{200}^3}{\ln(1 + c_{200}) - c_{200}/(1 + c_{200})}\,. \qquad (4.9)$$

where c_{200} is the concentration parameter at $\Delta_c = 200$, which is the assumed overdensity at collapse.

Since the mass of the NFW profile diverges if the density is integrated to infinite radius, we choose a cut-off radius at R_{200}, defined as the radius in which the average matter density is 200 times the average density of matter in the universe.

Using these definitions, the mass is given by

$$M_{200} = \Delta_c \Omega_M \bar{\rho}_c \frac{4}{3}\pi R_{200}^3\,. \qquad (4.10)$$

Together with the concentration parameter defined above, labelled accordingly as c_{200}, we can write analytic expressions for the lensing shear [Wright and Brainerd, 2000] for the NFW profile as

$$\Sigma_{NFW}(x)_z = 2r_s\delta_{NFW}\rho_c \begin{cases} \frac{1}{(x^2-1)}\left[1 - \frac{2}{\sqrt{1-x^2}}\text{arctanh}\sqrt{\frac{1-x}{1+x}}\right] & (x < 1) \\[3mm] \frac{1}{3} & (x = 1) \\[3mm] \frac{1}{(x^2-1)}\left[1 - \frac{2}{\sqrt{x^2-1}}\arctan\sqrt{\frac{x-1}{1+x}}\right] & (x > 1) \end{cases} , \qquad (4.11)$$

where $x = R/r_s$.

Beyond the shape of the halo mass distribution, there are practical effects which we need to consider when measuring the masses from the weak lensing signal of stacked systems. First, we will refine our model by including the probability of a fraction of the systems having significant offsets between the true barycentre and the CG. This has the effect of damping the signal in regions of small radii, as they will be further away from the true barycentre. And then we will include the effect of clustering of clusters.

4.3.3   *Treating miscentred systems*

When using radially symmetric profiles to model a stack of mass distributions, it is important to consider how the overall profile is affected when a fraction of the observed systems may be incorrectly centred. It is known that some fraction of the CGs may be offset from the true centre of the gravitational potential it inhabits [Girardi et al., 1997, Krempec-Krygier and Krygier, 1999], and also that cluster finder algorithms may identify a wrong galaxy as the CG [Johnston et al., 2007]. When we combine several clusters, the effect of this *miscentring* is to produce lower levels of shear in the inner radii, as the density peak of the miscentred systems will be spread away from the centre, which may bias results towards lower halo masses [Johnston et al., 2007].

If the $2D$ offset in the lens plane of a single profile is given by R_{off}, the azimuthally averaged surface profile will be given by a shift of the centre and an integral around the correct centre as

[Yang et al., 2006]:

$$\Sigma_{\text{off}}(R|R_{\text{off}}) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\theta \, \Sigma \left(\sqrt{R^2 + R_{\text{off}}^2 - 2RR_{\text{off}}cos(\theta)} \right) .$$

(4.12)

with the distribution of offsets being modelled as

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

(4.13)

where the parameter σ_{off} is the peak of the offset distribution.

The resulting mean surface mass profile for incorrectly centred combinations of clusters can be written then as

$$\Sigma^s(R) = \int_0^\infty \mathrm{d}R_{\text{off}}P(R_{\text{off}})\Sigma_{\text{off}}(R|R_{\text{off}}) .$$

(4.14)

4.3.4 *Contributions from the large scale structure of the universe*

The NFW profile is expected to be a good representation of halo profiles only to a certain scale, at most $\sim 2\mathrm{Mpc}/h$. To go further to outer radii, the contribution of the large scale structure must be accounted for. To do so, we can write the two halo mass contribution as

$$\rho_{2h} = b(\nu) \overbrace{\Omega_m \rho_{c,0}(1+z)^3}^{\bar{\rho}_m(z)} \xi^L(r,z) ,$$

(4.15)

where $b(\nu)$ is the linear halo bias calculated at the density peak height of the halo $\nu = \delta_c/\sigma(M)$, where $\sigma(M)$ is the linear matter variance in the Lagrangian scale of the halo, that is, with $R = [3M/(4\pi\bar{\rho}_m)]^{1/3}$.

The linear matter correlation function at redshift z can as

$$\xi^L(r,z) = D(z)^2 \, \sigma_8^2 \, \xi_n^L\left[(1+z)r\right] ,$$

(4.16)

where $\xi_n^L(r)$ is the linear correlation function at redshift zero, taken from a CAMB-calculated [Lewis and Challinor, 2002] linear power spectrum, normalised to $\sigma_8 = 1$, and $D(z)$ is the growth function.
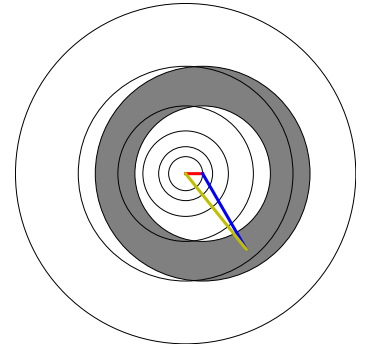
Figure 4.6: The effect of miscentring of galaxy clusters can be modeled as a shift (red) in the real centre relative to the centre being considered. Here the blue line represents R, whereas the yellow is the true radius from the centre of the mass distribution. This, together with an integral over the distribution of miscentring radii, accounts for the distribution of miscentrings in a stack of galaxy systems.

Now, to calculate the projected density due to the large scale structure, we can use the bias for a halo of mass M at redshift z as

$$B(z, M) := b(z, M)\,\Omega_{\mathrm{M}}\,\sigma_8^2 D(z)^2\,, \qquad (4.17)$$

so that we can write the projected 2-halo term as

$$\Sigma_{2h}(R) = B(z, M)\Sigma_l(R)\,, \qquad (4.18)$$

with

$$\Sigma_l = (1 + z)^2 \rho_{c,0} W((1 + z)R)\,, \qquad (4.19)$$

and

$$W(R) := \int_{-\infty}^{\infty} \mathrm{d}y\; \xi_l\left(\sqrt{y^2 + R^2}\right)\,. \qquad (4.20)$$

4.3.5 The Full Model

Figure 4.7: Testing the model with parameters taken from [Johnston et al., 2007] for comparison. Here we see each contribution for the total model as a function of radius for the scales that we use in our analysis.

$M_{200} = 44 \times 10^{12} h^{-1} M_{\odot}$, $c_{200} = 4$, $p_{cc} = 0.4$, $\sigma_{off} = 0.42 h^{-1} \mathrm{Mpc}$ & $M_0 = .32 \times 10^{12} h^{-1} M_{\odot}$

Legend:
- $\Delta\Sigma$ Total
- BCG point mass
- $\Delta\Sigma_{NFW}$
- $\Delta\Sigma_{NFW}^{off}$
- $\Delta\Sigma_{2h}$

Axes: $\Delta\Sigma(R)$ [$h M_{\odot} \cdot \mathrm{pc}^{-2}$] versus R [Mpc/h]
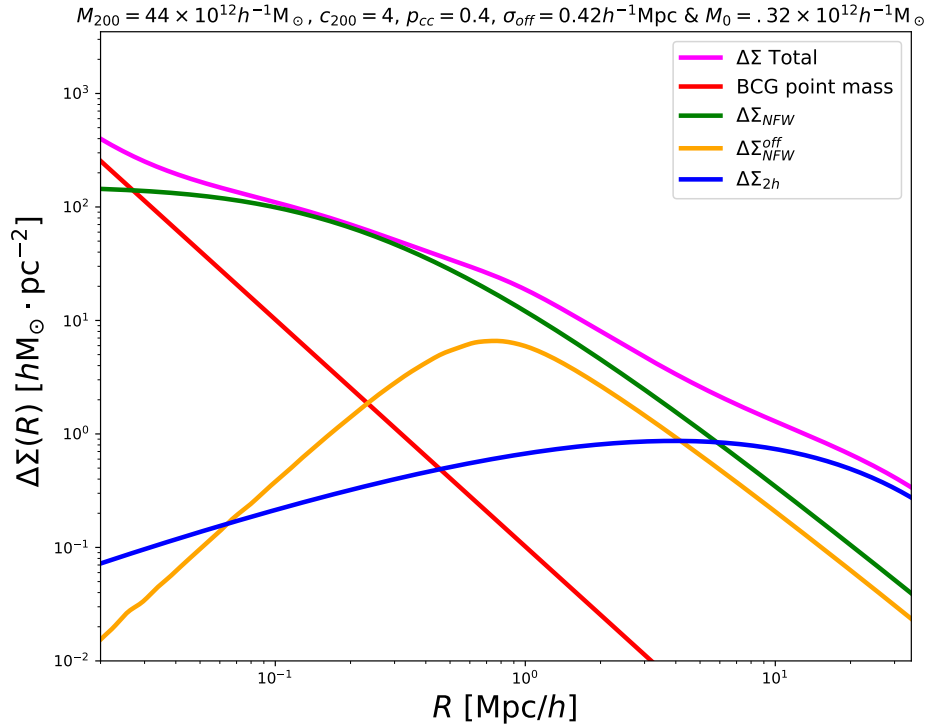
We have now all ingredients to our parametric model, which can use up to 5 parameters, namely $M_0, M_{200}, c_{200}, \sigma_{\mathrm{off}}, p_{\mathrm{cc}}$ - which represent, respectively, the baryonic mass of the BCG,

the halo mass and concentration inside R_{200}, the peak of the miscentring offset radius distribution, and the fraction of systems that were miscentred in the stack.

Using the lensing differential surface mass density (Eq. 3.33) [5], where $\bar{\Sigma}(R)$ is the average mass density inside a radius R, we can parametrise the total lensing signal as

$$\Delta\Sigma(R|M_0, M_{200}, c_{200}, p_{\rm cc}, \sigma_{\rm off}) = \frac{M_0}{\pi R^2} + p_{\rm cc}\Delta\Sigma_{\rm NFW}(R) + (1 - p_{\rm cc})\Delta\Sigma_{\rm NFW}^{\rm off}(R) + \Delta\Sigma_{2ht}(R) \,.$$

(4.21)

The partial contribution of each of this terms is represented in figure 4.7 using fiducial values for a typical stack of clusters of $\sim 10^{14} \, \mathrm{M}_\odot$.

4.3.6 Stacking Galaxy Systems

To address the low SNR from shape measurements, mostly due to intrinsic shape dispersion of galaxies, we use the cross-correlation lensing [Johnston et al., 2007], which consists of stacking the signal of a number of systems selected by a set of properties, such as redshift and richness. We define three redshift *slices*, one at a lower redshift interval $(0.2 < z < 0.4)$, another in an intermediate range $(0.4 < z < 0.6)$, and a last one at $(0.6 < z < 0.7)$. Furthermore, we restrict ourselves to an interval of cluster richness $15 < \lambda < 45$, and then divide into sets of different ΔM_{1-2}. We also reject systems with poorly-defined central galaxies by requiring that the probability that the CG is correctly identified (which is an output from redMaPPer) is greater than $p_{cen} > 0.8$. We find a total of 941 systems that fit these criteria, divided by stacks as shown in Table 4.1.

To compare stacks of different median magnitude gaps optimising our signal, we did not divide into "Fossil" and "non Fossil" systems, as the FS count was too low. Instead we define roughly equal-sized partitions of each redshift slice ranked by ΔM_{1-2}.

In order to test the stability of our results to different partitions, we used both terciles and quartiles cuts in each redshift slice. Additionally, we tested the higher magnitude gap tercile/quartile against a stack made from the combined rest of the systems in

that redshift slice, to have a simpler picture from which we can understand our results. The resulting median values for the magnitude gap (ΔM_{1-2}), richness (λ), and redshift (z) for each stack can be seen in Table 4.1, for the quartiles case.

		ΔM_{1-2}			
redshift		Large	Interm. I	Interm II	Small
	N	40	41	41	42
$0.2 < z < 0.4$	ΔM_{1-2}	$1.7^{+0.5}_{-0.1}$	$1.1^{+0.2}_{-0.1}$	$0.8^{+0.1}_{-0.1}$	$0.5^{+0.2}_{-0.2}$
	λ	20^{+10}_{-4}	22^{+6}_{-4}	22^{+11}_{-6}	22^{+10}_{-5}
	z	0.32	0.31	0.33	0.34
redshift		Large	Interm. I	Interm II	Small
	N	82	82	82	84
$0.4 < z < 0.6$	ΔM_{1-2}	$1.7^{+0.4}_{-0.2}$	$1.2^{+0.2}_{-0.2}$	$0.9^{+0.2}_{-0.2}$	$0.5^{+0.2}_{-0.3}$
	λ	22^{+7}_{-6}	20^{+12}_{-4}	22^{+12}_{-5}	19^{+10}_{-3}
	z	0.50	0.51	0.51	0.53
redshift		Large	Interm. I	Interm II	Small
	N	111	112	111	113
$0.6 < z < 0.7$	ΔM_{1-2}	$1.6^{+0.5}_{-0.4}$	$1.0^{+0.3}_{-0.2}$	$0.6^{+0.5}_{-0.3}$	$0.5^{+0.4}_{-0.5}$
	λ	21^{+8}_{-4}	20^{+12}_{-4}	20^{+11}_{-4}	21^{+10}_{-4}
	z	0.67	0.68	0.68	0.69

Table 4.1: Stacks, with the number of systems N, and the respective median of magnitude gaps ΔM_{1-2}, richness λ, and redshift z when the redshift slices are further separated into quartiles of different magnitude gaps.

Having defined the stacks of similar systems, we combine the lensing signal of each stack by first defining radial annuli, and calculating the estimator for the lensing signal for each ring as:

$$\widehat{\Delta\Sigma}(R) = \frac{\sum_{d,s} w^{d,s} \Sigma^{d,s}_{crit} \gamma_t^s}{\sum_{d,s} w^{d,s}} \tag{4.22}$$

where γ_t is the tangential ellipticity of each background source galaxy, $\Sigma^{d,s}_{crit}$ is the lensing critical density for each pair system (d) galaxy (s) and $w^{d,s} = w^s \left(\Sigma^{d,s}_{crit}\right)^{-2}$ is the lensing weight, which quantifies the measurement error in each source galaxy.

Errors in the lensing signal estimator were calculated by bootstrapping each stack $N = 200$ times, from which we also calculate the full covariance matrix between each pair of rings i and j.

$$C_{i,j} = \left[\frac{N}{N-1}\right]^2 \frac{1}{N} \sum_{k}^{N} \left[\Delta\Sigma_k(R_i) - \overline{\Delta\Sigma_k}(R_i)\right] \times \left[\Delta\Sigma_k(R_j) - \overline{\Delta\Sigma_k}(R_j)\right] . \tag{4.23}$$

We have tested several values ($N = 50, 150, 200, 300$) for the bootstrap number and arrived at roughly similar error bars in all of them, without any clear trend of diminishing errors with higher N. Also, bootstrap errors were compared to simple standard deviation of the shear signal in each bin, resulting in the standard deviation being a less conservative (smaller error bars) approach.

4.4 *Analysis*

We have calculated the lensing signal $\Delta\Sigma(R)$ for 6 logarithmically spaced rings, spanning from about $150 \mathrm{kpc}/h$ to $6\mathrm{Mpc}/h$[6], and using the geometric centre of each radial bin as values for R. We use this large extent of radii so that the typical scale of the centring offset distribution is well within our fitting range, minimising the effect of that choice, as discussed by Mandelbaum et al. [2008a].

[6] More exactly: 0.1584, 0.3311, 0.6918, 1.4454, 3.0199, and 6.3096 Mpc, with the bin frontiers defined so that each of these radii are geometrical centres of each bin.

Using the EMCEE code of Foreman-Mackey et al. [2013], we find posterior distributions for the parameters in our model by performing an MCMC fit of a multivariate Gaussian likelihood given by

$$
\ln\mathcal{L} = -\frac{1}{2}\left[\left(\widehat{\Delta\Sigma} - \Delta\Sigma(R|M_{200}, c_{200}, p_{\mathrm{cc}})\right)^T \times C^{-1} \times \left(\widehat{\Delta\Sigma} - \Delta\Sigma(R|M_{200}, c_{200}, p_{\mathrm{cc}})\right)\right],
$$
(4.24)

where $C^{-1} = \mathcal{H}C^{-1}$ is the inverse of the covariance matrix corrected by the method of Hartlap et al. [2007][7]. We use flat priors for concentration and mass, and a Gaussian prior on p_{cc} as shown in table 4.2. The prior on the correctly centred fraction comes from the redMaPPer algorithm, which gives probabilities for the correct identification of the CG. We discarded the baryonic component of the central galaxy and used a fiducial value for the peak of the centre offset distribution after tests fitting the full model did not affect our results other than resulting in less precision in $\widehat{\Delta\Sigma}$ errors. This has no significant effect on measuring concentrations, as the inclusion of the central galaxy masses found in the full model (eq. 4.21) changed only the innermost radial bin and only by less than 10%.

[7] The Hartlap factor is given by

$$
\frac{N}{N - p - 1}
$$
(4.25)

where N is the number of samples made to calculate the covariance matrix and p is the size of the covariance matrix.

Parameter	Prior
$\log(M_{200})$	$[12, 15]$
c_{200}	$[0.1, 20]$
p_{cc}	$\mathcal{N}(\overline{p_{\mathrm{cen}}}, \sigma_{p_{\mathrm{cen}}})$
σ_{off}	$0.41\ \mathrm{Mpc}/h$
M_0	0

Table 4.2: Priors used with the MCMC fitting process. The $\log(M_{200})$ and c_{200} are flat priors. The p_{cc} prior is determined by a Gaussian with centre and width determined by the redMaPPer given p_{cen} of the highest probability central galaxy for each system. We have also fixed the miscentring offset parameter to a fiducial value and the BCG baryonic mass to zero, as the data have not enough SNR to constrain a full parameter model.

The results of the $\Delta\Sigma(R)$ estimation and the fitting process are shown together in figure 4.8, where we plot the result for a stack

of the quartile with the largest ΔM_{1-2} systems against a stack of the others in the three redshift slices. This plot shows both the tangential and cross ellipticity components, the latter being a tool to diagnose any systematic effect.

We can see in figure 4.8 that the slope of the mass profile is steeper in the centre for the stacks with larger magnitude gaps, especially in the lower redshift slice. Additionally, we see that in the higher redshift slice the profile is poorly constrained, which can be attributed to the lack of observational depth.

Figure 4.8: Lensing signal data points combined with lines that represent model fitting. The thinner, diffuse, lines represent a random selected sample of chain links of the posterior probability chain obtained by the MCMC - which serves as a picture of uncertainty, while the thicker line represents the median-derived best values for the model. The cross-ellipticity shown below is consistent with no systematic effects.
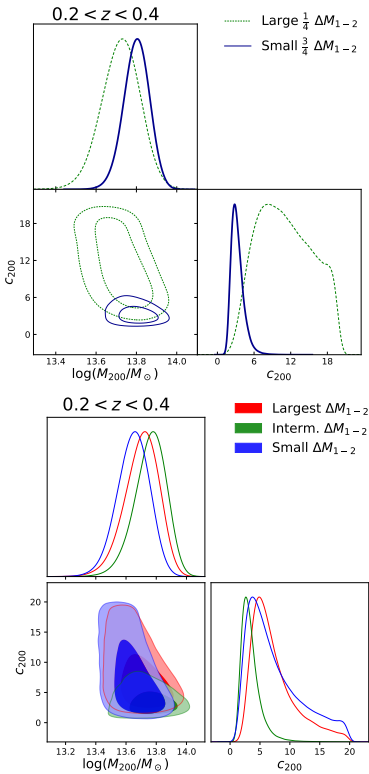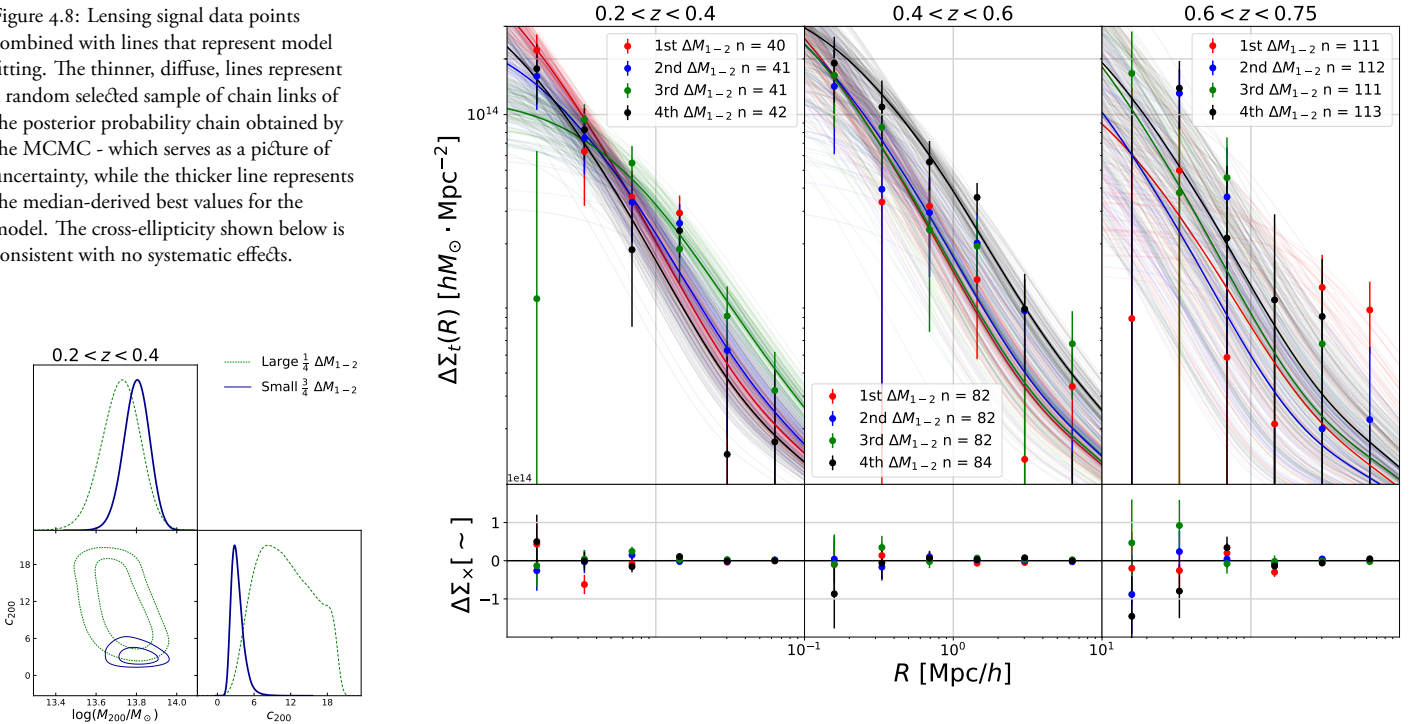
Figure 4.9: MCMC-computed 2D posterior distributions for lower ($0.2 < z < 0.4$) redshift stacks with peak-normalised 1D posteriors of mass and concentration, taken from the quartile partition, comparing the higher magnitude gap quartile with a stack with the other three. The contour lines progressively define 68%, 95% and 99% confidence level regions.

4.4.1 Posterior Distributions for Masses and Concentrations

Here we present the $2D$ posterior distributions marginalised over p_{cc} for the lower Redshift slice, in figure 4.9. In the intermediate redshift slice ($0.4 < z < 0.6$) we have been unable to constrain the concentration for the quartiles case, with the posterior spreading flat as the prior. In the third (highest) redshift slice, both masses and concentrations are poorly constrained by the data - as we can see in Figure 4.8, and we restrain from further analysis.

The mass posteriors of each stack in the two lower redshift slices have medians of the NFW M_{200} that are in agreement with

expectations for the richness λ of each stack [Rykoff et al., 2014].

4.4.2   *Correcting for the mass-concentration relation*

In the hierarchical scenario of the formation of the large scale structure of the universe, more massive halos assemble latter and thus, are expected to be less concentrated [Navarro et al., 1997]. This gives rise to a mass-concentration relation where more massive objects are generally less concentrated. Therefore, to compare stacks of different masses properly, we must first correct the concentrations to compare them on an "equal mass" reference. The mass-concentration relation of dark matter halos has been studied by several authors over many studies. We have used the relation given by Duffy et al. [2008] (Eq. 4.26) as a scaling relation to offset the concentrations to of the lower ΔM_{1-2} to the higher one.

$$c_{200}^{i,\text{corr}} = c_{200}^{i} \left(\frac{M_{200}^{j}}{M_{200}^{i}} \right)^{\beta} \text{ with } \beta = -0.091 \, . \qquad (4.26)$$
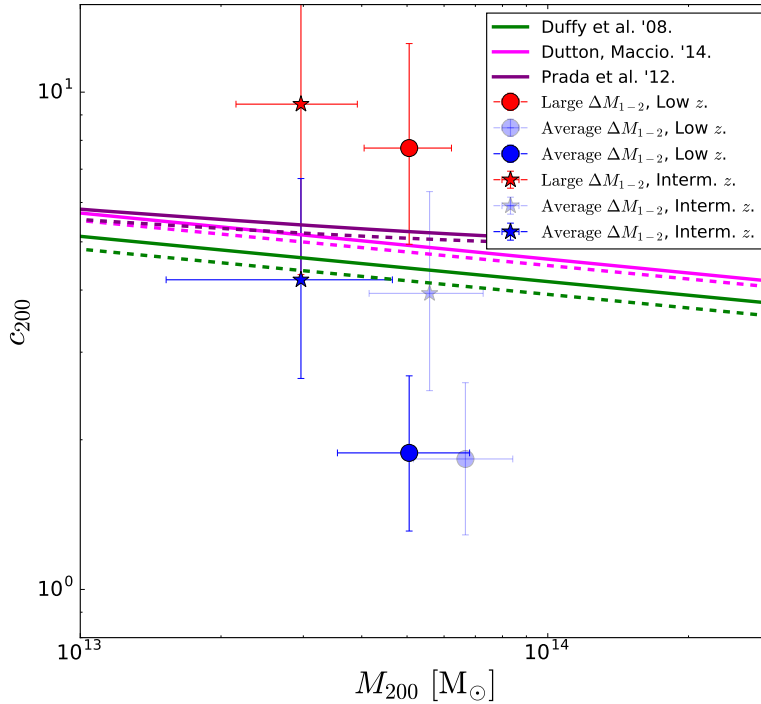
Figure 4.10: Adjusting concentrations to compare in an "equal-mass" footing, we slide them to the left using the relation of Duffy et al. [2008]. The solid markers are corrected mass-concentrations of the MCMC-calculated translucent translucent ones. The solid lines represent several $c - M$ relations for $z = 0$ and the dashed for $z = 0.5$. This graph corresponds to the comparison between the largest ΔM_{1-2} stack against the stack of the rest three quartiles of systems in each redshift slice.

In order to apply this correction taking into account the full probability distribution in mass and concentration, we calculate

c_{200}^{corr} by choosing random pairs of the MCMC chain links from the reference stack (the largest $\Delta M_{1,2}$ one) and the one to be corrected, and apply the scaling relation from the Duffy relation.

We have then a probability distribution for the corrected concentrations. The resulting shift in concentration can be seen in figure 4.10, where we show the correction in mass and concentration for the Large ΔM_{1-2} quartile against the stack of the other 3 quartiles, in both the low and medium redshift slices.

	Terciles						Quartiles							
	$0.2 < z < 0.4$			$0.4 < z < 0.6$			$0.2 < z < 0.4$				$0.4 < z < 0.6$			
ΔM_{1-2}	Large	Interm.	Small	Large	Interm.	Small	Large	Interm. I	Interm II	Small	Large	Interm. I	Interm II	Small
M_{200}	$5.1^{+1.4}_{-1.2}$	$5.8^{+1.5}_{-1.3}$	$4.5^{+1.2}_{-1.0}$	$4.5^{+1.2}_{-1.0}$	$2.7^{+1.0}_{-0.8}$	$4.4^{+1.5}_{-1.3}$	$5.1^{+1.2}_{-1.0}$	$5.3^{+1.5}_{-1.4}$	$7.7^{+2.0}_{-1.8}$	$3.9^{+1.3}_{-1.1}$	$3.1^{+1.1}_{-1.0}$	$3.7^{+1.5}_{-1.3}$	$3.3^{+1.6}_{-1.4}$	$10.0^{+2.5}_{-2.2}$
c_{200}	$6.3^{+4.6}_{-2.3}$	$3.1^{+1.7}_{-1.1}$	$6.2^{+6.4}_{-3.0}$	$11.4^{+5.7}_{-5.6}$	$5.3^{+6.5}_{-2.9}$	$4.6^{+3.8}_{-1.8}$	$11.0^{+5.5}_{-4.4}$	$4.5^{+4.0}_{-1.9}$	$1.7^{+0.6}_{-0.5}$	$8.1^{+6.9}_{-4.3}$	$10.0^{+6.4}_{-5.5}$	$7.2^{+7.4}_{-4.3}$	$8.7^{+7.1}_{-5.3}$	$4.8^{+3.1}_{-1.7}$
c_{200}^{corr}	$6.3^{+4.6}_{-2.3}$	$3.2^{+1.7}_{-1.1}$	$6.1^{+6.1}_{-3.0}$	$11.4^{+5.7}_{-5.6}$	$5.6^{+6.6}_{-3.0}$	$5.1^{+4.1}_{-2.0}$	$11.0^{+5.5}_{-4.4}$	$4.5^{+3.8}_{-2.0}$	$1.8^{+0.7}_{-0.5}$	$7.7^{+6.5}_{-4.2}$	$10.0^{+6.4}_{-5.5}$	$7.4^{+7.4}_{-4.4}$	$8.7^{+7.0}_{-5.3}$	$5.3^{+3.4}_{-1.9}$
p_{cc}	0.96	0.95	0.92	0.95	0.94	0.90	0.96	0.95	0.95	0.92	0.96	0.95	0.93	0.90

Table 4.3: Posterior medians and 68%CL intervals for the masses, concentrations, corrected concentrations and miscentred fraction of systems, each calculated by marginalising other parameters. Masses are given in units of 10^{13} M$_\odot$, and the 68CL intervals around the median p_{cc} values are of about $\sim \pm 0.05$.

To summarise the results of the MCMC fitting we present in table 4.3 the median and 68%CL intervals for the halo mass (M_{200}), concentration c_{200}, corrected concentration c_{200}^{corr}, and p_{cc} for each tercile and quartile from the two lower redshift slices.

4.5 Results

We have quantified the differences in corrected concentrations between different stacks by building, for each pair in a redshift slice, a chain of differences between randomly selected chain links of each corrected concentration chain. This resulting chain samples the probability distribution for the differences Δc_{200} and the results can be seen in figure 4.11 again for the case of 1 quartile of large ΔM_{1-2} against the other three, in both low and medium redshift slices.

By integrating the positive side of these distributions (which can be simply done by counting the fraction of chain links with $\Delta c_{200} > 0$) we calculate probabilities that one be greater than the other, given the previous concentration results. We display these results in table 4.4.

For the lower redshift slice ($0.2 < z < 0.4$) we find that the larger magnitude gap stack has a high probability of being more concentrated than the stack of other systems both when divided in terciles ($P = 0.84$), and quartiles ($P = 0.98$).
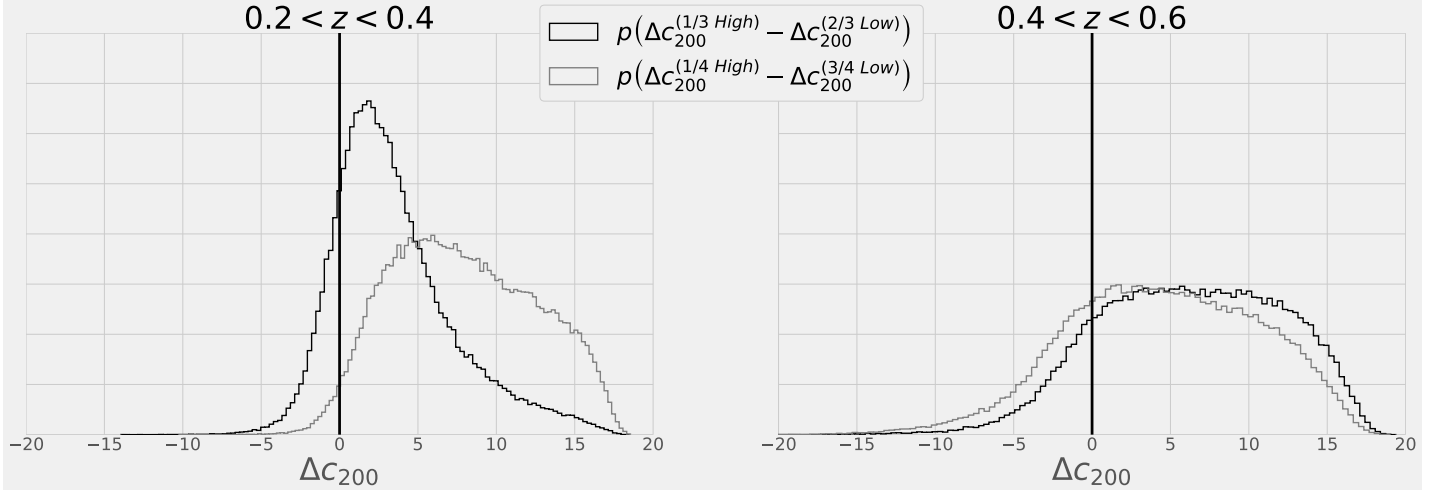
Figure 4.11: Probability distribution for the difference in concentration (Δc_{200}) between larger and smaller ΔM_{1-2} stacks, when each redshift interval is divided into terciles (black) and quartiles (grey) of different magnitude gaps, using the 1 versus the rest comparison. The probabilities that the larger magnitude gap stack is more concentrated than the smaller is then given by the portion of the distribution left to the $\Delta c_{200} = 0$ line.

In the intermediate redshift slice ($0.4 < z < 0.6$) we again see the same tendencies, where the first terciles (quartiles) have probabilities of $P = 0.85(P = 0.76)$ of being more concentrated than the other systems stacked together. However, we have to be cautious to consider this result, as the concentration for the higher magnitude gap quartile was not well defined, with its probability distribution being mostly flat as the prior. In the terciles, however, the posterior was much better defined.

redshift	Terciles	Δc_{200}	$P\left(\Delta c_{200} > 0\right)$
	"Large-Intermediate"	$3.1^{+4.5}_{-2.8}$	0.86
$0.2 < z < 0.4$	"Large-Small"	$0.3^{+4.9}_{-6.0}$	0.52
	"Intermediate-Small"	$-2.8^{+3.4}_{-6.1}$	0.20
$0.2 < z < 0.4$	"Large-Others"	$2.8^{+4.3}_{-2.8}$	0.84
$0.4 < z < 0.6$	"Large-Others"	$6.3^{+6.2}_{-6.1}$	0.85
redshift	Quartiles	Δc_{200}	$P\left(\Delta c_{200} > 0\right)$
	"Large-Interm. I"	$5.9^{+5.9}_{-5.4}$	0.86
	"Large-Interm. II"	$9.1^{+5.5}_{-4.5}$	> 0.99
$0.2 < z < 0.4$	'Large-Small''	$2.8^{+6.5}_{-6.8}$	0.67
	"Interm. I-Interm. II"	$2.6^{+3.9}_{-2.1}$	0.92
	"Interm. I-Small"	$-2.9^{+5.1}_{-6.7}$	0.29
	"Interm. II -Small"	$-5.9^{+4.2}_{-6.5}$	0.04
$0.2 < z < 0.4$	"Large-Others"	$7.6^{+5.5}_{-4.5}$	0.98
$0.4 < z < 0.6$	"Large-Others"	$4.5^{+6.6}_{-6.1}$	0.76

Table 4.4: Median values with 68%CL intervals for the difference in concentration between quartile stacks by magnitude gap in the two lower redshift slices.

4.6 *Discussion of the Results*

In general, our data indicate that populations of galaxy systems with larger magnitude gaps have, on average, more concentrated

mass profiles. In our analysis we did not attempt to segregate Fossil Systems into a particular stack, to optimise our lensing SNR. However, we still managed to see a trend of larger magnitude gap systems having more concentrated halos. The difference in concentrations between stacks was most significant when comparing the higher magnitude gap stack against the rest of the systems. Some of this trend still appears in the intermediate redshift slice $0.4 < z < 0.6$ though, unlike in the low redshift slice, the results were not statistically significant to the 95%CL.

We have found some pairs of stacks in which this trend is reversed. This, however, only happens in smaller magnitude gap pairs, where the difference in median ΔM_{1-2} of each of the stacks is smaller. We expect that these systems are less different in nature and these particular results can be attributed to random scatter.

Apart from statistical fluctuations, it is possible that the smallest ΔM_{1-2} stack of each redshift slice have more systems with misidentified CGs, and ongoing mergers. This not only complicates the choice of centre, but also can bias parametric mass estimates [Hoekstra et al., 2002, 2013].

Possible problems in our analysis can arise due to biased sampling, since cluster finding algorithms can have preconceptions on what a galaxy cluster/group *should* look like. The redMaPPer algorithm, for instance, assumes a Schechter-like luminosity function, which can undercount more extreme ΔM_{1-2} systems. In fact, we find a very low fraction of systems as Fossils($\sim 5\%$), in comparison with the literature ($\sim 10 - 20\%$) which may be a failure of the algorithm to identify systems with large magnitude gaps. Further analysis with other cluster finding algorithms can, in the future, bring new light to this question.

A few systems have, due to the nature of the cluster finding algorithm, central galaxies that are not the brightest system member. All these are included in the smallest magnitude gap stacks, both in terciles or quartiles. As such, removing these from our analysis just moves the terciles/quartiles breaks up in magnitude gaps - and gives similar results.

Nevertheless, we can support our conclusions simply by comparing only the largest ΔM_{1-2} stack against the intermediate one(s). Both stacks have well behaved systems, in the sense that

there is very little doubt on the central galaxy as shown, for instance, of the very high and similar values of the p_{cc} parameter and we do see significant differences on the concentration points that large ΔM_{1-2} systems tends to be more concentrated.

In conclusion, our results push further the argument that the magnitude gap is an indicator of early forming systems, as there is theoretical basis for the earlier formation of more concentrated systems [Navarro et al., 1995].

This shows that, despite the evolution of magnitude gaps and contamination by younger systems, that is expected from simulations, some information of the accretion history is still present in the concentration. A deeper analysis will become possible when more statistically complete populations of galaxy systems become available in future surveys.

Regarding the claim that FGs are dark clusters, that is, halos with low occupation numbers, data shows no evidence to support this, since all our stacks have essentially the same richness.

4.7 *Summary and Conclusions*

We have performed the first weak lensing mass distribution analysis of galaxy systems ranked by the magnitude gap of the central galaxy (CG) and the brightest satellite galaxy (BSG). Using cross-correlation lensing of stacked systems, divided by redshift and rank of ΔM_{1-2} we calculated thee shear signal in radial bins and, to that data, applied a MCMC procedure to calculate posterior distributions for the mass and concentrations of these stacks. After correcting the concentration for mass differences by a $c - M$ relation, we calculated the probability distribution of differences in concentrations between the stacks in each redshift slice, dividing them both in terciles and in quartiles. Finally, we integrate the positive side of these distributions to calculate a probability of the larger ΔM_{1-2} stack having a more concentrated halo according to our data.

We have found significant evidence in our data indicating that populations of systems with larger magnitude gaps are more likely to have more concentrated mass distributions. Assuming that our model describes the mass distributions sufficiently well, we
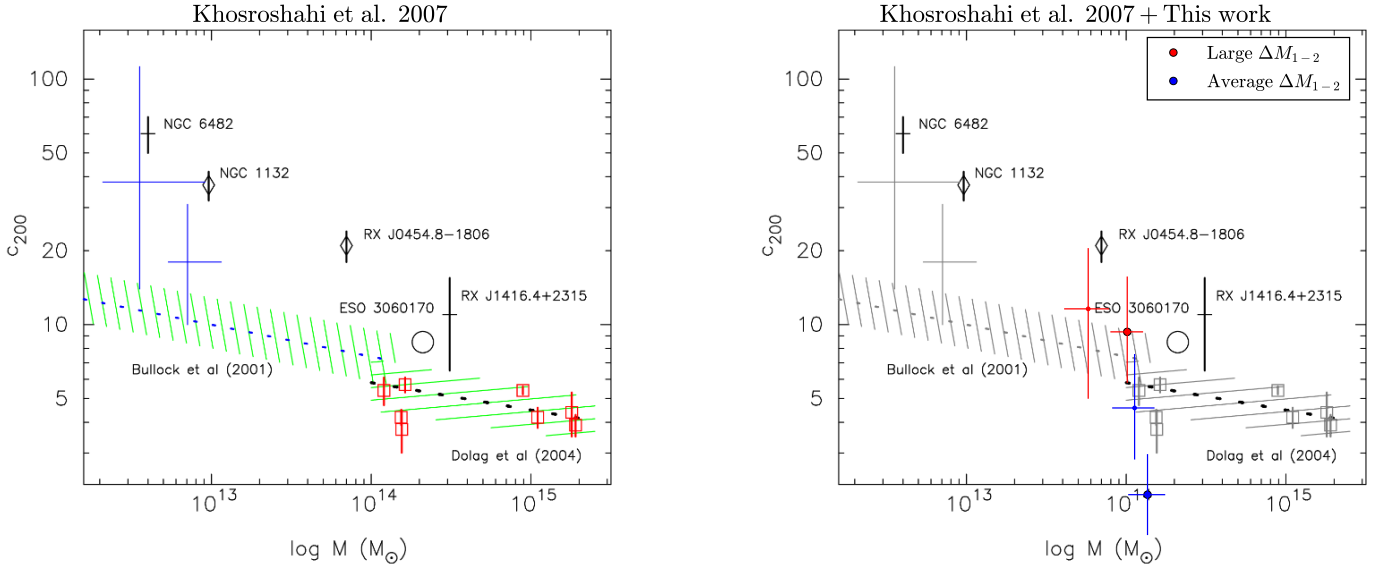
Figure 4.12: A comparison with Khosroshahi et al. [2007] shows us where we land on the mass-concentration relation with different selections by magnitude gaps.

find that, in the lower redshift slice ($0.2 < z < 0.4$), our larger magnitude gap quartile stack is more concentrated than the rest at more than 95%CL. We also find the same trend in the intermediate redshift slice ($0.4 < z < 0.6$) although with less confidence.

These results agree with and strengthen the claim of larger magnitude gaps being correlated to more concentrated halos [Khosroshahi et al., 2007],as we see in Figure 4.12. Together with simulation studies by Deason et al. [2013], Gozaliasl et al. [2014], this supports the idea that a substantive fraction of systems with larger magnitude gaps form earlier. On the other hand, we do not discard differences in the recent mass accretion history. As both simulation and observational arguments are are all based in the ΛCDM model of cosmology, these results are compatible with the same assumptions. Recent work [Gozaliasl et al., 2014] with simulations further indicate that magnitude gaps can be used together with other tracers to efficiently discriminate virialised populations of galaxy systems. These more relaxed systems offer, in turn, more insight into Cosmology and the formation of the large scale structure of the universe [Mantz et al., 2015].

Finally, large upcoming optical surveys will present more complete populations of galaxy systems at low redshifts in great areas of the sky such as JPAS [Benítez et al., 2015], Euclid [Laureijs et al., 2011], and others. With more precise photometric

redshifts as in the case of JPAS, weak lensing will be at an advantage point to look for more evidence of the dynamical nature of large magnitude gap systems.

5
Luminosity Centroids & Concentrations

As mentioned in section 2.5.2, the displacement between the galaxy system centre and the centremost galaxy, be it the brightest one or not, is an indicator to the dynamical state of the system [Rasmussen et al., 2012]. Besides that, it has become increasingly clear that magnitude gaps alone cannot guarantee that the system is early-forming [Kundert et al., 2017, D'Onghia et al., 2005, Raouf et al., 2014].

Therefore, we now turn to investigate other galaxy system observables that are thought to correlate to earlier histories of mass accretion. Recalling that Raouf et al. [2014] has proposed, from simulations, that offset between CG and luminosity-weighted centre of the system galaxies, and the r-band luminosity can be alternative indicators for early mass accretion histories, our first attempt focus on the luminosity centroid offset and its relationship with mass concentrations.

At the time of this study, we did not have a catalogue of X-ray luminosity centres, which are most commonly used for comparisons with the CG position [Skibba and Macciò, 2011]. It should be noted, however, that merging processes are known to disturb the X-ray emitting gas and move their centres and peaks out of the barycentre of the system [eg. Markevitch and Vikhlinin, 2007]. A better measure for the true center of the system would be, naturally, the centre of the total mass, that could be derived from mass-mapping with weak lensing. The determination of barycentres with weak lensing would be, however, exceedingly noisey and imprecise, given the data available from Moraes et al. [2014].

Finally, this work is also to be seen as a preparation for expanding these questions into future surveys, such as J-PAS,

in which the much greater count of galaxy systems can better constrain the parameters we are interested in.

5.1   Methods

The description of the methods in this section is intended to be detailed and extensive, much more so than the upcoming paper. We do so not only to provide an accurate description of the research, but also to help upcoming students in continuing this line of research.

The luminosity centre of a system of galaxies from the REDMAPPER catalogue can calculated by:

$$\vec{X}_{cen} = \frac{\sum \vec{X}_n p_n L_n}{\sum p_n L_n} \, , \tag{5.1}$$

where p_n is the probability of the galaxy being part of the system, L_n is the SDSS i-band luminosity, and $\vec{X}_n = (\alpha_n, \delta_n)$ is the WCS coordinates.

The offset can then calculated by taking the angular separation between the centre $(\vec{X}_{cen})$ and the CG $(\vec{X}_{CG})$, and multiplying by the angular diameter distance:

$$R_{off} = \mathcal{V}(\vec{X}_{CG}, \vec{X}_{cen}) \, d_A(0, z_{cluster}) \, , \tag{5.2}$$

where $\mathcal{V}$ is the Vincenty's formula given by Equation 3.3, and d_A can be found in this work at Equation 1.24.

We have attempted first to calculate the offsets between the redMaPPer derived CG and the luminosity centroid by computing with all galaxies, that is, including the central galaxy and all members. Again, we do that with the same sampling pre-conceptions detailed in the last chapter: we remove systems with non-brightest central galaxies and - given the previous results - restrict ourselves to systems with lower redshifts ($0.2 < z < 0.4$) and richness ($15 < \lambda < 50$). Doing this procedure considering the central galaxies, we get much smaller offsets, which are also slightly correlated with CG properties. As such, offsets considering the central galaxies are also more correlated ($r = 0.48$) with magnitude gaps than offsets between the CG and luminosity centroid of all *other* galaxies ($r = 0.18$).

Therefore, the relationship between the CG and the system is better described by the luminosity centroid of all *satellite* galaxies, as there are groups (such as fossils) in which the dominant luminosity of the central galaxy would draw the centroid much closer to the itself.

When using these offsets, we found that the concentration was larger for the *larger* offsets, what proved to be difficult to explain. However, by inspecting some of the members (specially those with the largest offsets, it was proposed that at least some of the large offset systems had projection effects between two systems, that were in turn misclassified as a single system by REDMAPPER.
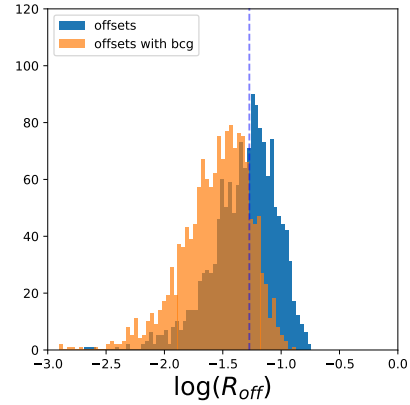
Figure 5.2: Distribution of $\log R_{off}$ of the systems used (those with $z < 0.4$, see below) both including and excluding the CG. The median of the offsets without the cg is depicted ($\log \tilde{R}_{off} = -1.2$).
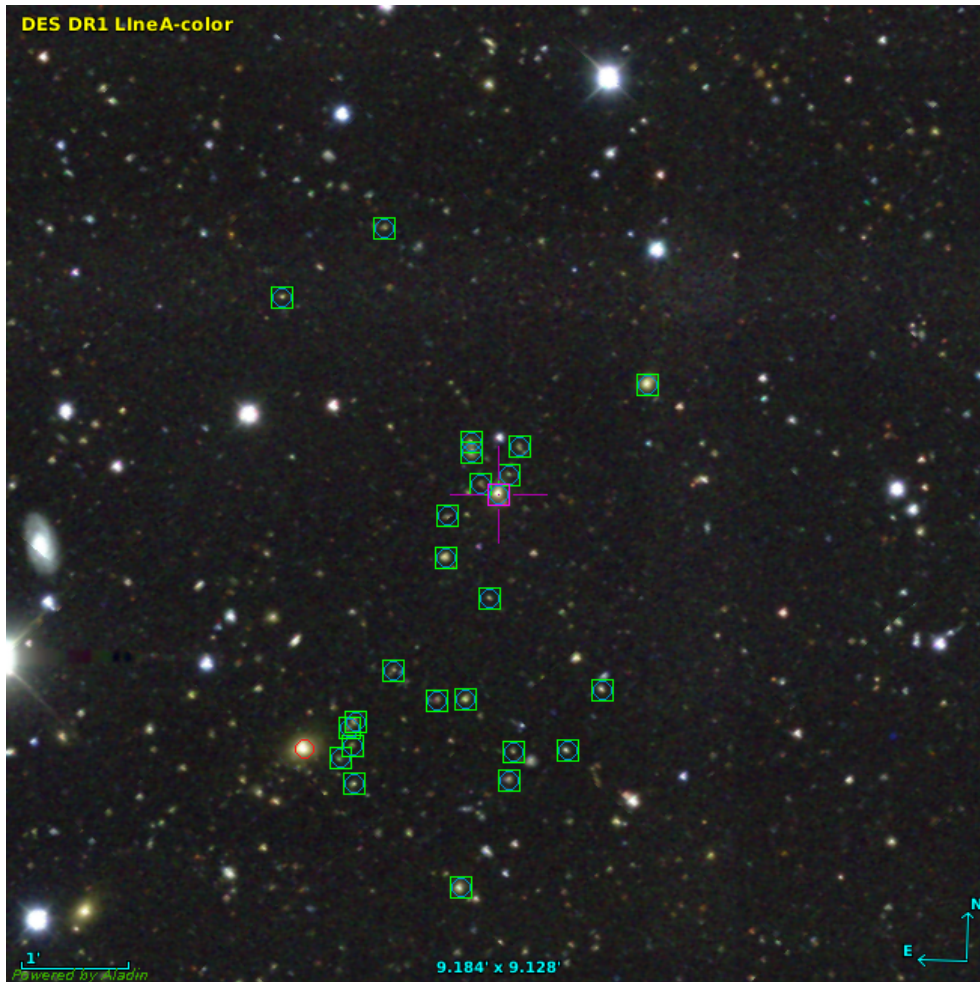
Figure 5.3: One of the systems with highest offsets, at RA = 23 : 37 : 30.22 DEC = −00 : 38 : 18.2. Green boxes indicate CS82 redMaPPer database for cluster members. The redMaPPer central galaxy is under the cross, and has a spectroscopic redshift of $z = 0.2772$. The major elliptical galaxy at the bottom of the image, slightly to the right, that is not identified as a member of the system, has a spectroscopic redshift of $z = 0.285$ [SDSS Collaboration, 2015] and is given as a BCG to many of the same galaxies, in other cluster finders [Hao et al., 2010, Rykoff et al., 2014, Wen et al., 2012, Oguri, 2014] .*Image source: DES DR1 LineA color image.*

This is because separating two red sequences of clusters with small redshift differences is difficult. This can be illustrated by the system with the highest offset found in redMaPPer (Fig: 5.3). In this system, the member galaxy distribution is highly asymmetric and one galaxy, not in our CS82 redMaPPer catalogue, has a very

close, but different redshift. In many other cluster catalogues, including the SDSS DR-8 redMaPPer, many of these galaxies (to the south in the figure) are part of a group centered in the non-assigned galaxy, instead of the one given.

These projection effects lead to systems with much lower masses and low offsets being actually under consideration, which in turn have much higher concentrations. For a time, this plagued us until we considered the same cut ($0.3R_c$), the same as in chapter 4. It is also important to highlight that this is a problem in red-sequence cluster finders in general, as the width of the red-sequence can make systems that are near in sky coordinates to be indistinguishable.

Figure 5.4: The distribution of clusters offsets over the same selection by magnitude gaps. Here we see that some of the systems with the largest offsets (the large circles) have average magnitude gaps and that their magnitude gaps in the r bands tend to be slightly higher.
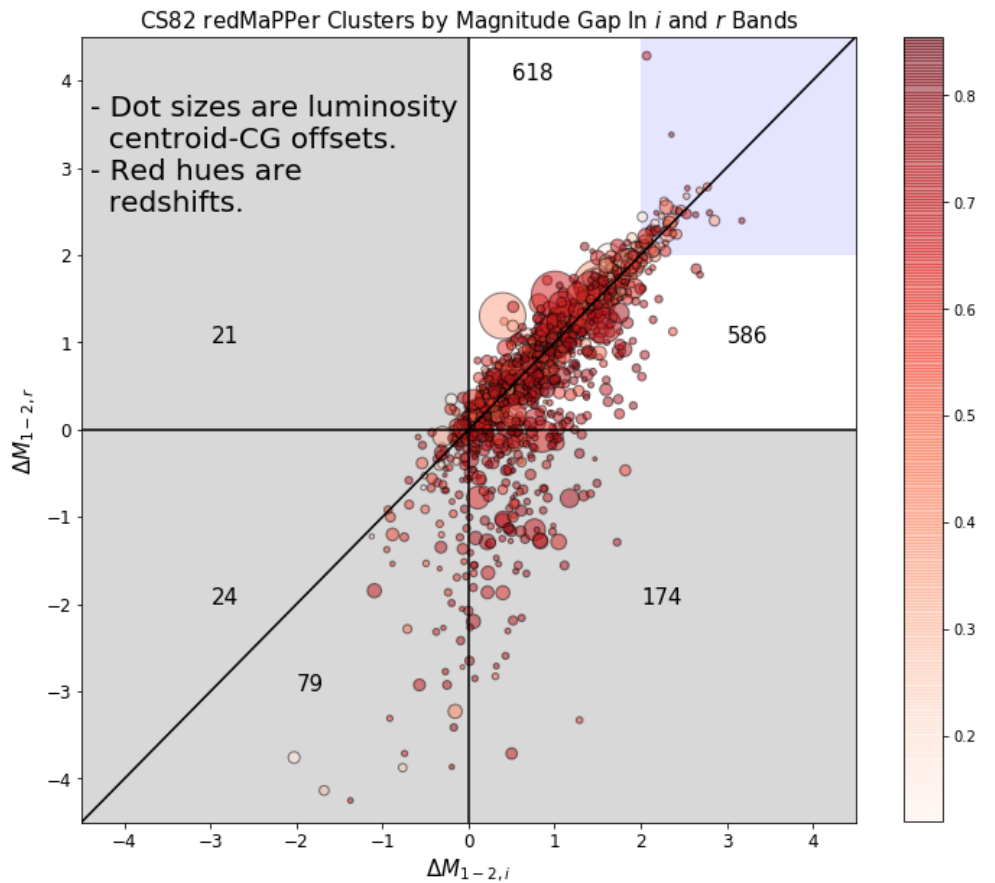
5.2   Results

We repeat the analysis processes given in chapter 4, however, we here use 8 radial bins. Concentrations listed here are corrected by the mass-concentration relation in the same manner of the

previous chapter. We first used the cross-correlation code on stacks divided by both offsets and magnitude gaps. Using the previously mentioned richness and redshift range, we have found $N = 158$ objects to analyse with cross-correlation lensing. We proceed similarly with our previous study by dividing them into terciles and quartiles ranked by each quantity we want to compare with concentrations.

We have also studied these systems first with all 5 parameters of the model to understand if the miscentring fraction (p_{cc}) or the miscentring characteristic distance σ_{off} where higher in greater offset systems, to which our result was that they were not.

Dividing arbitrarily into 4 groups by partitioning in ΔM_{1-2} and R_{off} (Fig 5.5), we found no correlation between redshift, magnitude gaps, and offset luminosity centroid offsets, concluding that they are independent observables.
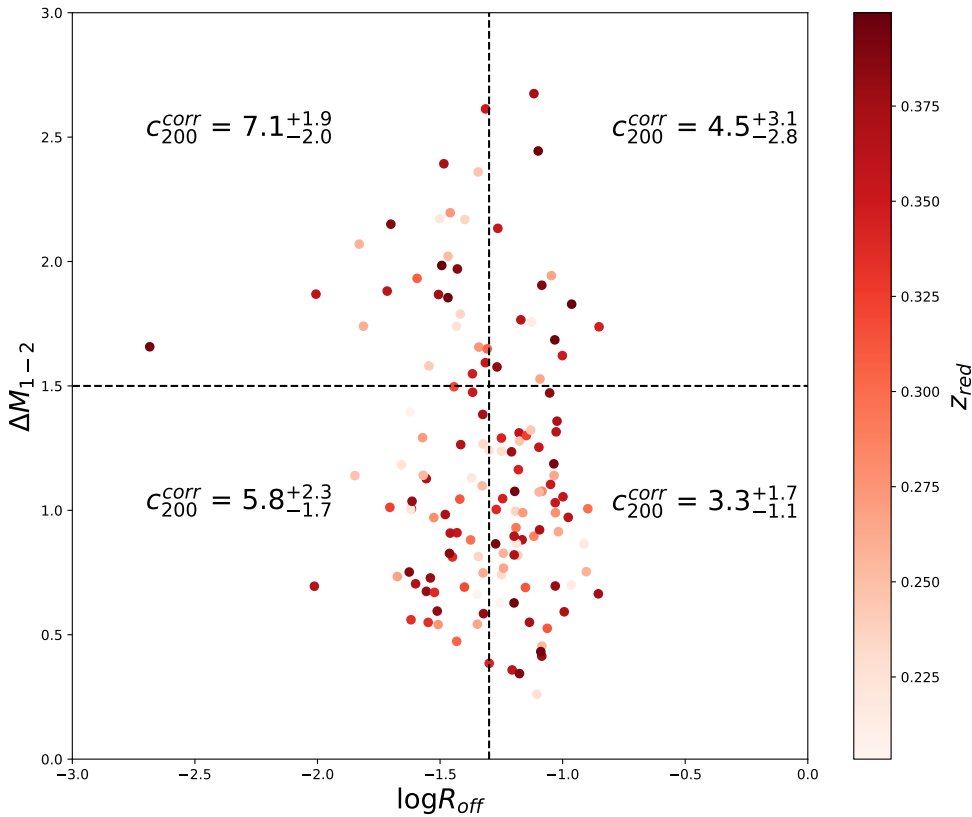
Figure 5.5: A comparison between each stack divided by magnitude gaps and offsets of systems with richnesses $15 < \lambda < 50$ and redshifts $0.1 < z < 0.4$. The concentration results are already corrected for the mass-concentration relation as in the previous chapter.

The full 5-parameter posteriors indicate that the catalogue is relatively well-centered. How this possibly relates to the off-centering of the CG and the luminosity centre is will not be discussed here. A full discussion of the centres will need,

Table 5.1: Posterior medians and 68%CL intervals for the masses, concentrations, corrected concentrations and miscentred fraction of systems, each calculated by marginalising other parameters. Masses are given in units of 10^{13} $M_\odot$, and the 68CL intervals around the median p_{cc} values are of about $\sim \pm 0.05$. These results indicate that concentrations rise with ΔM_{1-2} as previously shown, and decline with greater offsets R_{off}. Also, the A metric is an indicator for the concentration.

Small ΔM_{1-2}				Large ΔM_{1-2}			
Small R_{off}		Large R_{off}		Small R_{off}		Large R_{off}	
M_{200}	$7.0^{+1.4}_{-1.2}$	M_{200}	$5.5^{+1.4}_{-1.2}$	M_{200}	$7.8^{+1.7}_{-1.6}$	M_{200}	$2.7^{+2.0}_{-1.3}$
c_{200}	$5.8^{+2.3}_{-1.7}$	c_{200}	$3.3^{+1.8}_{-1.1}$	c_{200}	$7.0^{+2.0}_{-2.0}$	c_{200}	$4.9^{+3.4}_{-3.0}$
P_{cc}	$0.97^{+0.02}_{-0.02}$	P_{cc}	$0.97^{+0.02}_{-0.03}$	P_{cc}	$0.99^{+0.01}_{-0.01}$	P_{cc}	$0.99^{+0.01}_{-0.01}$
σ_{off}	$0.76^{+1.1}_{-0.51}$	σ_{off}	$1.0^{+1.2}_{-0.71}$	σ_{off}	$0.7^{+1.1}_{-0.44}$	σ_{off}	$0.7^{+1.0}_{-0.44}$
$\log M_0$	$10.0^{+1.4}_{-1.3}$	$\log M_0$	$10.1^{+1.3}_{-1.4}$	$\log M_0$	$10.0^{+1.4}_{-1.4}$	$\log M_0$	$10.0^{+1.4}_{-1.4}$
c_{200}^{corr}	$5.8^{+2.3}_{-1.7}$	c_{200}^{corr}	$3.3^{+1.7}_{-1.1}$	c_{200}^{corr}	$7.1^{+1.9}_{-2.0}$	c_{200}^{corr}	$4.5^{+3.1}_{-2.8}$

eventually, better weak lensing data, more objects, and matching X-ray luminosity distributions.

Our results can be summarised by the medians and CL68 regions drawn from the posteriors marginalising other parameters, and presented in the table 5.1.

Next, we have combined the distributions of magnitude gaps and offsets to create a tentative metric "A", given by difference between each quantity around the average, scaled by their respective standard deviations, that is,

$$A = \frac{\Delta M_{1-2} - \langle \Delta M_{1-2} \rangle}{\sigma \left[\Delta M_{1-2} \right]} - \frac{R_{off} - \langle R_{off} \rangle}{\sigma \left[R_{off} \right]}, \qquad (5.3)$$

where σ stands for the standard deviation of each variable.

We have analysed partitions of terciles and quartiles ranked by A and the results show no improvement in respect to correlating to concentrations over either offsets or magnitude gaps being used alone. A comparison of all the results given by separating in terciles ranked by ΔM_{1-2}, R_{off}, and A is shown in table 5.2

We have repeated these tests using the redMaPPer fiducial values for $\sigma_{off} = 0.17$ Mpc and $p_{cc} = 0.75$ from Zhang et al. [2019] to marginalise over the nuisance parameters, using the baryonic mass of the central galaxy of 10^{10} $M_\odot$. The results were not qualitatively different of those previously obtained, but had larger errors. We display the posterior distributions in Fig. 5.7.

All in all, our results remain quantitatively very sensitive to choices in selection for the stacks, be it redshift or richness, and separating into terciles or quartiles. This is because at the present number of systems, our stacks are very close to the limit that we can minimally constrain mass concentrations[1]. However, in all tests, the results follow the expected trends, even if with low significance. Improvements will depend on having larger surveys.

[1] We have tested stacks with fewer than 35 systems, and the concentration posteriors start to fill the priors uniformly.
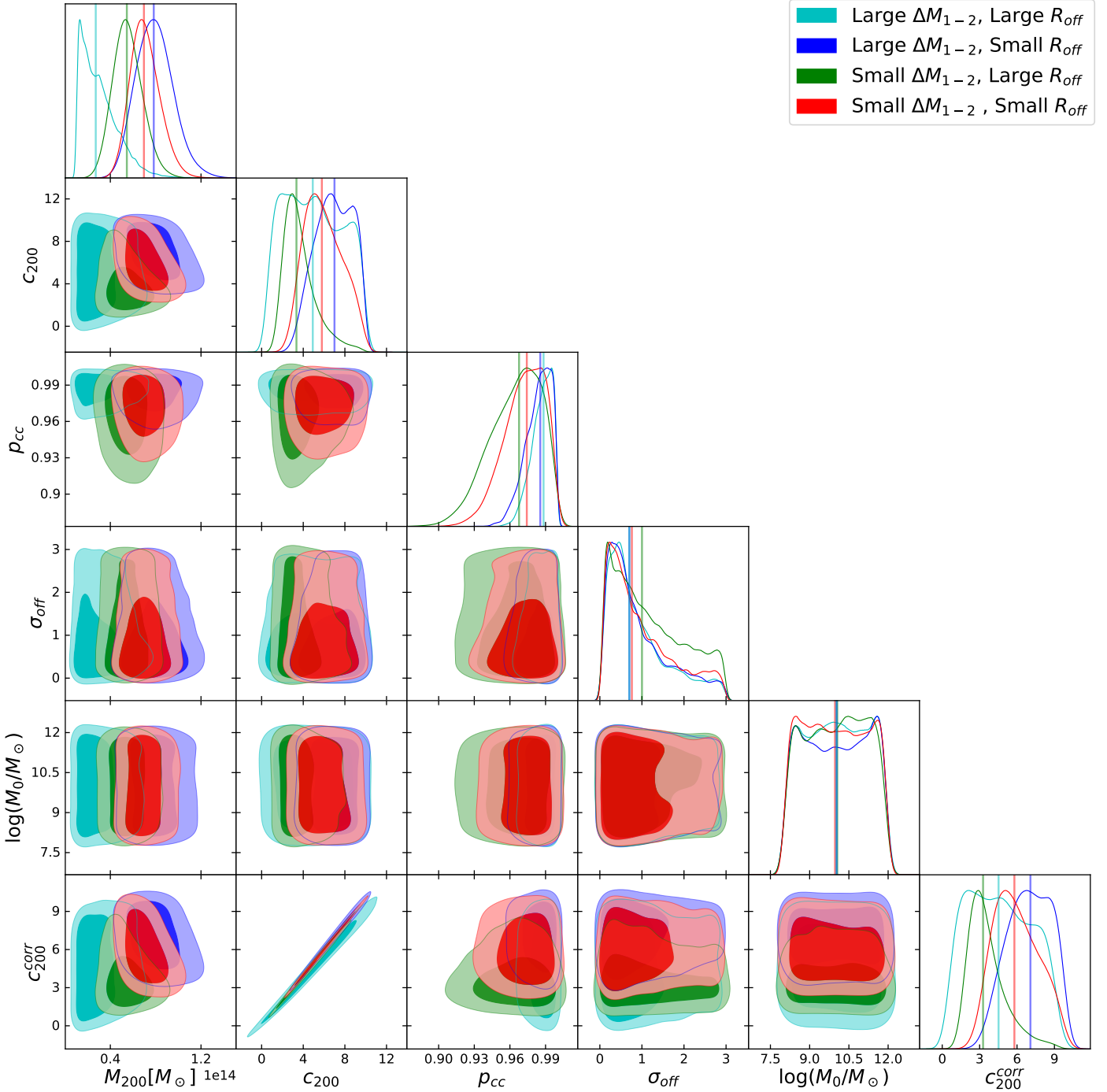
Figure 5.6: Posteriors over partitions of $R_{off} \times \Delta M$ with medians (vertical lines). Large ΔM_{1-2} & small R_{off} stacks (blue) have the highest concentrations and small ΔM_{1-2} & large R_{off} (green) have the lowest. The large ΔM_{1-2} & high R_{off} stack has lower masses, and may have more problematic identifications. We find σ_{off} higher than that off [Zhang et al., 2019], because of the long tailed posterior, as we have less constraining power with our data.

We have also investigated the signal-to-noise of the data across the 8 radial bins by using

$$(S/R) = \left[\Delta\Sigma^T C^{-1} \Delta\Sigma\right]^{1/2} \qquad (5.4)$$

where $\Delta\Sigma$ here represents a vector of the signal in each of the 8 radial bins and c^{-1} is an estimator of the inverse of the covariance matrix of the data.

Selected by	ΔM_{1-2}		R_{off}		A	
Small	M_{200}	$6.04^{+1.26}_{-1.16}$	M_{200}	$5.62^{+1.05}_{-0.98}$	M_{200}	$5.58^{+1.43}_{-1.26}$
	c_{200}	$4.39^{+2.09}_{-1.42}$	c_{200}	$7.01^{+1.85}_{-1.75}$	c_{200}	$3.54^{+1.82}_{-1.22}$
	P_{cc}	$0.97^{+0.02}_{-0.02}$	P_{cc}	$0.98^{+0.01}_{-0.01}$	P_{cc}	$0.97^{+0.02}_{-0.02}$
	σ_{off}	$0.88^{+1.19}_{-0.60}$	σ_{off}	$0.73^{+1.13}_{-0.49}$	σ_{off}	$0.88^{+1.19}_{-0.62}$
	$\log M_0$	$9.98^{+1.37}_{-1.35}$	$\log M_0$	$10.10^{+1.35}_{-1.42}$	$\log M_0$	$9.91^{+1.38}_{-1.32}$
	c_{200}^{corr}	$4.31^{+1.99}_{-1.41}$	c_{200}^{corr}	$7.01^{+1.85}_{-1.75}$	c_{200}^{corr}	$3.54^{+1.82}_{-1.22}$
Intermediate	M_{200}	$6.09^{+1.35}_{-1.22}$	M_{200}	$9.51^{+1.43}_{-1.42}$	M_{200}	$5.86^{+1.27}_{-1.12}$
	c_{200}	$5.32^{+2.13}_{-1.59}$	c_{200}	$4.79^{+1.53}_{-1.13}$	c_{200}	$4.58^{+2.15}_{-1.67}$
	P_{cc}	$0.97^{+0.02}_{-0.02}$	P_{cc}	$0.97^{+0.02}_{-0.02}$	P_{cc}	$0.97^{+0.02}_{-0.02}$
	σ_{off}	$0.88^{+1.16}_{-0.61}$	σ_{off}	$0.91^{+1.25}_{-0.63}$	σ_{off}	$0.78^{+1.21}_{-0.52}$
	$\log M_0$	$9.98^{+1.36}_{-1.37}$	$\log M_0$	$9.97^{+1.34}_{-1.34}$	$\log M_0$	$9.99^{+1.34}_{-1.37}$
	c_{200}^{corr}	$5.23^{+2.03}_{-1.55}$	c_{200}^{corr}	$5.02^{+1.55}_{-1.17}$	c_{200}^{corr}	$4.61^{+2.16}_{-1.68}$
High	M_{200}	$7.57^{+1.22}_{-1.17}$	M_{200}	$5.09^{+1.07}_{-0.96}$	M_{200}	$7.08^{+1.11}_{-1.05}$
	c_{200}	$6.67^{+2.01}_{-1.85}$	c_{200}	$4.26^{+2.47}_{-1.51}$	c_{200}	$6.96^{+1.83}_{-1.62}$
	P_{cc}	$0.98^{+0.01}_{-0.02}$	P_{cc}	$0.97^{+0.02}_{-0.02}$	P_{cc}	$0.98^{+0.01}_{-0.01}$
	σ_{off}	$0.77^{+1.07}_{-0.48}$	σ_{off}	$0.86^{+1.22}_{-0.59}$	σ_{off}	$0.95^{+1.11}_{-0.63}$
	$\log M_0$	$10.01^{+1.41}_{-1.40}$	$\log M_0$	$10.03^{+1.32}_{-1.37}$	$\log M_0$	$9.92^{+1.46}_{-1.33}$
	c_{200}^{corr}	$6.67^{+2.01}_{-1.85}$	c_{200}^{corr}	$4.23^{+2.42}_{-1.49}$	c_{200}^{corr}	$7.11^{+1.84}_{-1.62}$

Table 5.2: Summary of median and CL68 margins for all 5 parameter model posteriors using 3 ranked stacks of 3 different quantities: the magnitude gap (ΔM_{1-2}), the offset between the CG and the luminosity centroid (R_{off}), and the tentative "Age" (A).

With $N = 52$, the (S/N) is about ~ 9. When trying to use quartiles, some of our stacks drop to below 7 and a proper analysis becomes more difficult.

5.3 Discussion

We investigated the claim that offsets between the central galaxy and the luminosity centroid of the system member galaxies are correlated, proposed by Raouf et al. [2014] to be an indicator of early mass accretion histories.

We are, at this preliminary stage, able to affirm that galaxy systems with smaller offsets between the CG and the luminosity centroid of inner galaxies tend to have higher concentrations, for redshifts below $z = 0.4$, which, in turn, are believed to be more relaxed, earlier formed systems, on average [Wechsler et al., 2002]. However, these results are based on fairly low signal-to-noise measurements of stacks of galaxy groups.

A combined analysis of ΔM_{1-2} and R_{off} has also shown to be an indicator of higher concentration. However, this still does not provide any improvements into correlations with concentration
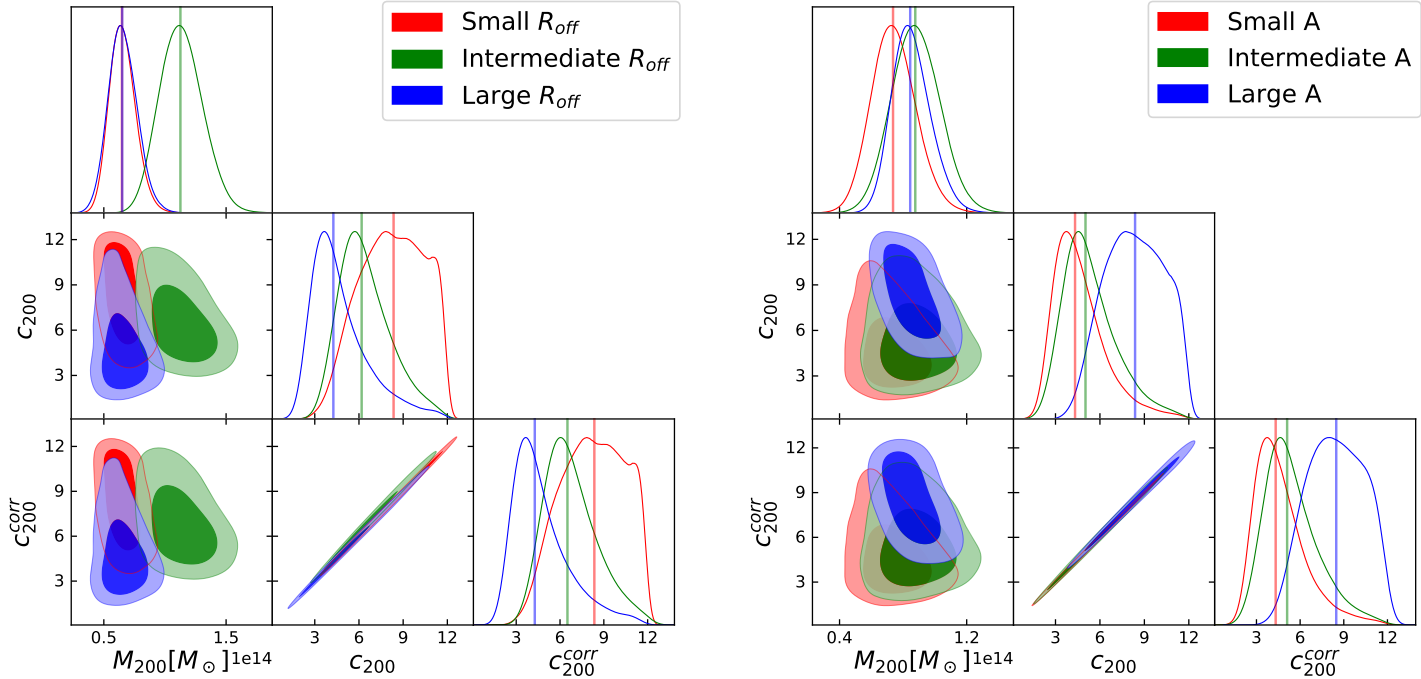
Figure 5.7: Posterior distributions for the 2-parameter (M_{200}, c_{200}) model for stacks ranked by offsets and A. We do not show here our results for magnitude gaps as they are similar to those in the previous chapter.

at this point. We are still investigating these results. This is something that will inevitably improve with larger surveys, that permit better constraints to masses and concentrations by weak gravitational lensing. When working with the dynamical state of systems from the point of mass profile concentration, higher numbers may be needed to properly separate the concentration posterior from each stack.

Further interesting tests can be done using new cluster finder catalogues, and models that take int account each galaxy system individually, by using a hierarchical Bayesian model that constrain a joint likelihood of all the systems in a partition.

A more interesting point, expanding on the subject of galaxy systems centres in general, will be to compare the weak lensing mass barycentre, the X-ray luminosity peak, and the position of the central galaxy. This can offer further information about the dynamical state of galaxy systems.

As a next step, we intend to use the r-band absolute magnitude of the CG to look for a 3 parameter space that better separate more concentrated systems, and differences in the steepness of red-sequence of each system, which can be correlated with cluster relaxation [Gladders et al., 1998]. With this, we may be able to discriminate better the more relaxed systems. In the near future,

also, surveys using machine learning over X-ray gas distributions may offer additional measurements of cluster substructure and relaxation processes.

6

Shear Measurements for Large Galaxy Surveys

THE PREVIOUSLY REITERATED NEED FOR GREATER NUMBER OF GALAXY CLUSTERS to understand their formation processes and dynamical characteristics takes us into the final chapter of this work. As a preparation for upcoming surveys, having a computationally scalable and well understood pipeline for shear measurements will be paramount.

The weak lensing regime is studied mainly in three different situations, in increasingly weaker signal: cluster lensing, galaxy-galaxy lensing, and cosmic shear. Cluster lensing refers to the regime around galaxy clusters, a few Mpc from their centers and has been the main tool in this work. Galaxy-galaxy lensing refers to the minute lensing effect on background galaxies by other galaxies - something that also can only be verified by adding the signal of many sources. Finally, cosmic shear looks for correlation between galaxies across many scales to probe the overall distribution of the mass in the universe. These regimes are separated from each other by roughly one order of magnitude in shear (cluster lensing displaying a shear of around 10% of the intrinsic ellipticity of galaxy images) .

This final chapter will cover the work done to offer a full, reliable, and usable pipeline for large galaxy surveys. We have adopted a three way strategy using overlaps of new observations with CFHTLenS shear measurements [Erben et al., 2013]: using second moments to find image ellipticities, forward fitting models of luminosity distribution, and neural networks over galaxy images. As an ongoing work, the results presented are not final and will probably change in the upcoming months.

6.1 Introduction

We are currently entering the era of the so-called Stage IV[1] dark energy experiments such as LSST [LSST Dark Energy Science Collaboration, 2012], Euclid [Laureijs et al., 2011], and WFIRST [Akeson, 2019], which investigate two of the most basic aspects of this mysterious, but dominant, content of the universe: its density and equation of state. All of these experiments will use weak gravitational lensing to map the mass distribution in the universe. Their main focus is cosmic shear - the correlation of shearing on neighbourhoods of the sky. This is a very minute effect, that requires extremely precise shear measurements of galaxy images. Additionally, properties of the population of galaxy clusters can offer another probe to cosmology, and as shown in [Loureiro et al., 2019, Dvorkin et al., 2019] to measure neutrino masses. Cluster masses from weak lensing coming from these surveys will also be used to find scaling relations with other observables, such as numbers of galaxies (ie. *richness*), total luminosity, and distribution of the intra-cluster gas temperature. These other observables that correlate well with total mass are then more easy targets to weigh an even greater number of galaxy clusters. This, in turn, become probes of statistical properties of the mass distribution in the LSS, such as the distribution of number of galaxy clusters per mass interval, which are correlated to fundamental properties in the standard model of Cosmology.

The precise measurement of weak lensing by the LSS for cosmic shear will need accurate shape measurements of $+10^9$ galaxies. The observation of each of these galaxies will be carried through many different filters, with several exposures each - and the combined data will be used to achieve much better precision in shear measurements [Jarvis et al., 2016]. This will make calculating ellipticity for each galaxy a process done on tens of images, greatly enhancing the computation costs.

The need for additional images in different filters comes from the stringent requirements of the upcoming surveys to surpass the current and previous ones [Dark Energy Survey Collaboration, 2018, Hildebrandt et al., 2017, Erben et al., 2013]. However, detailed care must be taken as it has been shown that improving

the precision of shape measurements can also introduce biases as systematic errors become the main contribution to lack of the error budget [Mandelbaum et al., 2018] .

After more than a decade of tests (STEP - Heymans et al. [2006]; STEP2 -Massey et al. [2007]; GREAT08 -Bridle et al. [2010]; GREAT10 - Kitching et al. [2012]; GREAT3 - Mandelbaum et al. [2015]), many implementations of different methods of shear measurement exist (eg. Kaiser et al. [1995], Miller et al. [2007], Hirata and Seljak [2003], Bernstein and Jarvis [2002], Sheldon [2015]) and have been verified. However, several of those are not open source, and out of many that are, most are not thoroughly documented. As such, it is important that we investigate some of the open source codes to offer a repeatable and usable solution for the future.

Finally, the existence of previously conducted weak lensing surveys that have been widely checked deals us a golden opportunity to test our practices to conduct shear measurements against real images. This, obviously, do not substitute the need for simulations, which will aid us in our first steps.

6.2   *Methods of Shear Measurements*

The shape of galaxies in images can vary greatly, from the very smoothly distributed luminosity of some elliptical galaxies, to clumpy, irregular blobs of stars. To do weak gravitational lensing, however, we are not interested in a detailed description of galaxy appearance, but in measuring an overall departure from circularity, as a proxy for local gravitational shearing effects on light rays - as explained before in chapter 3. There are two main types of methods used to measure the shapes of galaxy images: calculating image moments from the pixel values directly [eg. Kaiser et al., 1995], or fitting a *luminosity profile* to the galaxy image [eg. Miller et al., 2007]. This luminosity profile does not have to be a perfect description of the galaxy shape: we want to know how much this two dimensional profile is elliptical instead of circular.

The advantage of measuring image moments is that it is a relatively inexpensive computational task. However, it tends

to be more susceptible to image artefacts and the presence of other nearby objects in the image. This is the rationale of fitting parametric models, which in turn, are much more computationally demanding.

A more recent idea is to use convolutional neural networks on galaxy images - when surveys have overlaps with available shear data. This, in practice, amounts to expand previously obtained shear data to wider surveys and has the advantages of being practical (especially with novel machine learning libraries that are easy to use, as KERAS [Chollet et al., 2015], and backed up by GPU computing, with TENSORFLOW [Abadi et al., 2015]). We will be also demonstrating this in a practical application, together with a single application of one of each of the two methods mentioned before.

The Forward Process.

Galaxies: Intrinsic galaxy shapes to measured image:

Intrinsic galaxy (shape unknown) → Gravitational lensing causes a *shear (g)* → Atmosphere and telescope cause a convolution → Detectors measure a pixelated image → Image also contains noise
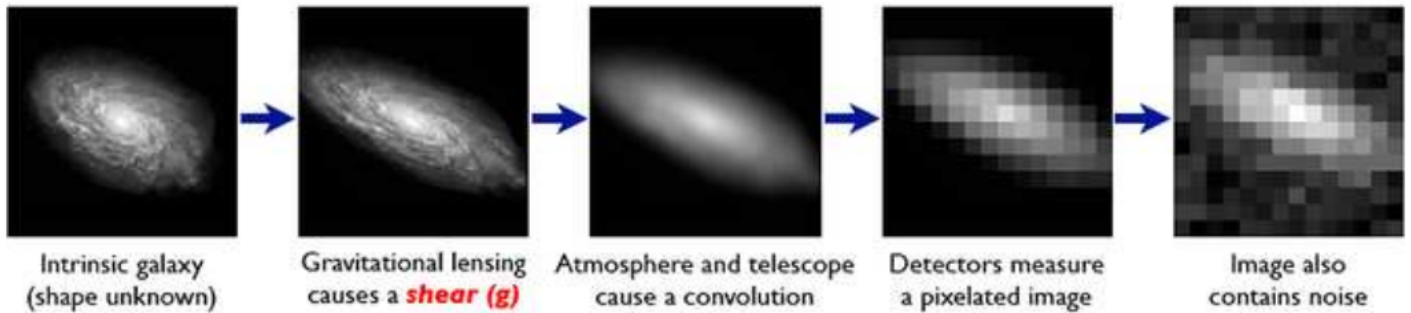
Stars: Point sources to star images:

Intrinsic star (point source) → Atmosphere and telescope cause a convolution → Detectors measure a pixelated image → Image also contains noise
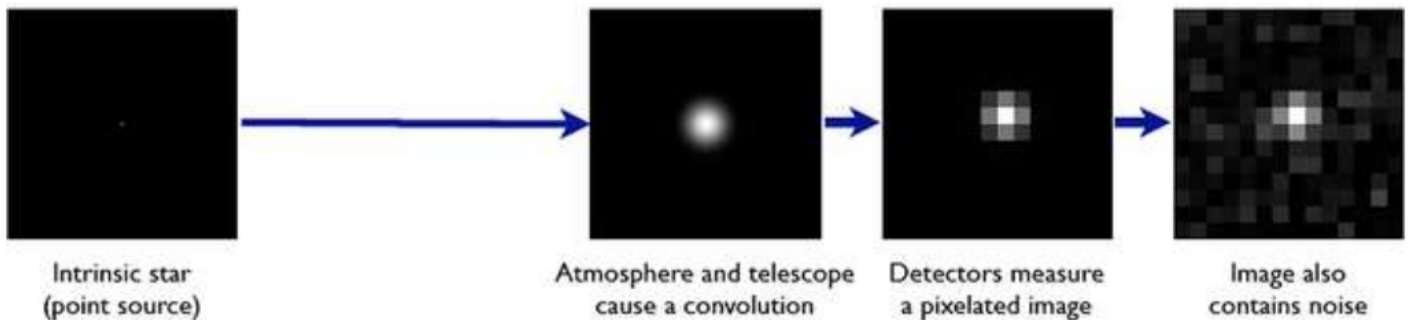
Figure 6.1: The now common place used image from the GREAT08 shear measurement challenge illustrates the forward process that composes galaxy and star images found in surveys. The goal of this work is to find a way to reverse this process and find ellipticities that, averaged, can display local shearing of light rays due to gravity. *Source:* [Bridle et al., 2010]

Before that we must address errors in shape measurements. As other measurement errors, there is a statistical and a systematic contribution to the error budget. Any measure of shear cannot be free of the intrinsic shape distribution of galaxies, nor of detector

noise. Systematic errors are also unavoidable, as telescopes have imperfections, and even the orientation of the actual shape of galaxies is not totally random [Erben et al., 2001, Mandelbaum et al., 2008b, Heavens et al., 2000]. Furthermore, because of needed simplifications, models cannot accurately describe the galaxy shape - this, in turn, becomes model biases.

To consider the effect of errors, we must address both accuracy and precision. To do so, consider the true ellipticities ϵ of a number of observed galaxies. Then, the departure from the observed ellipticities ϵ_o can be quantified as:

$$\epsilon_o = (1 + m)\epsilon + c \qquad (6.1)$$

where m is a multiplicative bias, and c is the additive bias. Also, we will have a scatter of ϵ_o.

In an ideal measurement, we want to have both m and c as close to zero as possible. It is against a certain level of m and c that we classify requirements for lensing surveys - because we can evaluate the effect of a certain level of these biases in the strength of the constrains of the quantities we want to measure, be it parameters of a cosmological model, or masses and concentrations of galaxy clusters.

For Stage IV surveys, shear measurement criteria have been defined by Mandelbaum et al. [2015] that set a goal of reaching $m < 2 \times 10^{-3}$ and $c < 2 \times 10^{-4}$, motivated by estimated requirements for Euclid [Massey et al., 2013].

6.3 *Data*

The main data we use for building the pipeline comes from the J-PAS pathfinder survey, which is a rehearsal for the future full-scale J-PAS survey. It consists of a single-CCD camera capable of using all filter trays of the full survey, handling a total of 5 wide band and 55 narrow band filters. The pathfinder survey has in total 4 adjacent (and overlapping 7.5%) fields. Each field has a 900 arcmin^2 area for a total of 0.967^o area.

Additionally, we use shear measurements from the CFHTLenS, to which we compare our results. A weak-lensing survey aimed at providing a large weak lensing survey, to be used in cosmology,
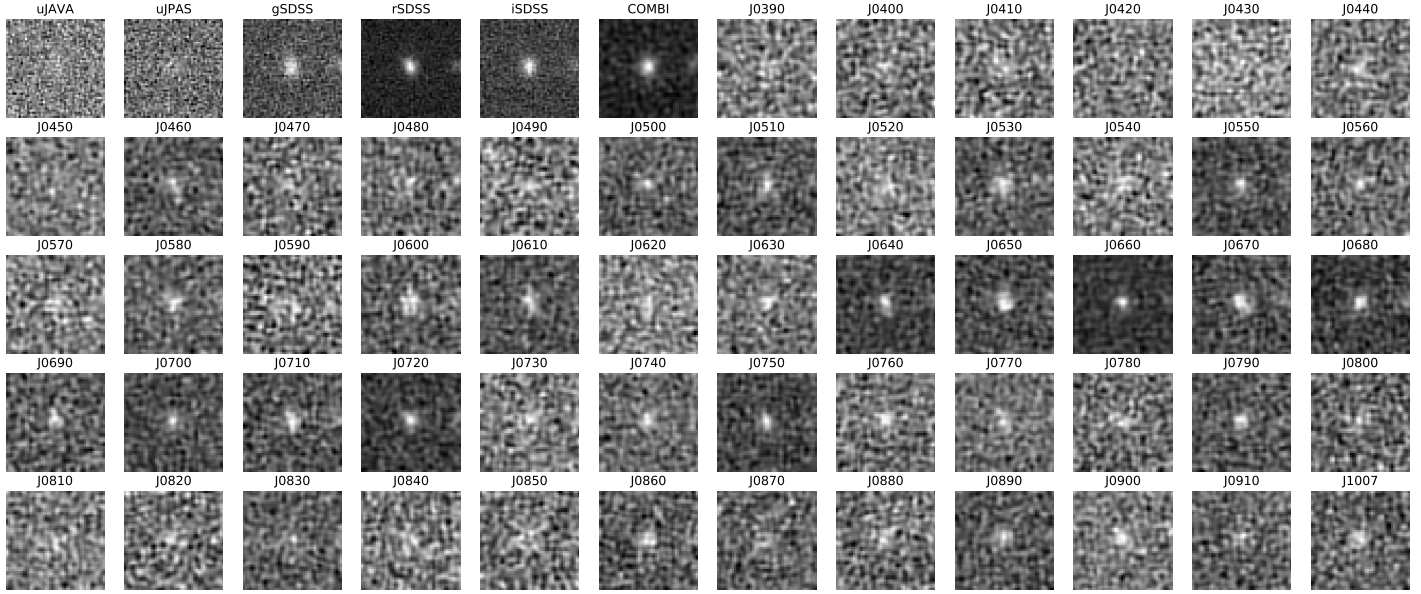
Figure 6.2: Example of a high signal-to-noise ratio galaxy $(S/N) = 26.8$, with images from all but one of the 60 filters of the upcoming J-PAS survey. The first 5 images are of the broadband filters, and those after "COMBI", 54 of the narrow-band (all except the bluest). The filters are labeled over each image, with the exception of "COMBI", that represents a coadded image from all narrow-band filters.

[2] LensFit ellipticity weights were mentioned in chapter 4

galaxy. The CFHTLenS uses the same MegaCam instrument described in chapter 4, and consists of deep and wide fields that have millions of galaxies with shear measurements done by LensFit - an example parametric fitting lensing code. As the CFHTLenS observations are deeper than the current large surveys (eg. DES), it has been used to provide trustable results of shear measurements to other surveys.

From the mini-J-PAS pipeline, using SExtractor, and matching galaxies from the CFHTLenS catalogue, we get a total 2873 objects with ellipticities and weights[2], that are given by

$$w = \frac{1}{\sigma_i^2 + C_{1,1} + C_{2,2}} \qquad (6.2)$$

where σ_i is the intrinsic dispersion of galaxy true shape ellipticities, and $C_{i,i}$ are the diagonal values of the covariance matrix of the measurement.

A display of the identified galaxies used in respect to the (S/N) of the detection in the J-PAS pathfinder survey and the quality of the lensing measurements is given in figure 6.3.

Finally, in order to test basic procedures with controlled data, we have GALSIM simulation of galaxy images, which we use, not to calibrate any external input, but just to check the procedures qualitatively.

6.4 Sources of systematic error in galaxy shape measurements

The incoming light from distant galaxies is subject to a number of transformations that result in the galaxy image recorded by the CCD. Passing through the atmosphere, the wavefront is scattered and enters the telescope not as a plane wave, or a bundle of parallel rays. Although this effect can be greatly minimised in modern telescopes by adaptive optics [Hardy, 1998], some scattering will remain. Furthermore, the characteristics of the optics of the telescope will add distortions to the image. Finally, and most important, the very finiteness of the telescope aperture makes the projected image a transformation of the incoming image. The way to encapsulate these additive effects that results in the galaxy image is to consider how a single ray of light, coming from a point in the sky, gets transformed into the image on the detector. The function that transforms this point-source to a

distribution of intensities over the detector plane is called the *Point Spread Function*, or PSF. Then, we can understand how an object is represented as an image created by a convolution between the original distribution of light rays and the PSF:

$$I_d(x, y) = I_{sky}(\alpha, \delta) * PSF(\vec{x}, \vec{\theta}) \times J, \qquad (6.3)$$

where $I_d(x, y)$ is the distribution of light flux over a coordinate system (x, y) on the detector, I_{sky} is the flux coming into the detector from directions (α, δ) in the sky, PSF is the point spread function, a convolution filter that tells how much the intensity from a direction $\vec{\theta} = (\alpha, \delta)$ affects a point (x, y) on the detector, and J is a jacobian representing the change of coordinates between the detector and sky, the *world coordinate system* (WCS).

The PSF can impact both multiplicative and additive bias. It has a multiplicative effect because the spread of light blurs the source, making shapes rounder, less elliptical. The additive effect happens because the geometry of the detector can create anisotropic spread of the light ray, which will sum to the ellipticity of the source.

To measure the effect the PSF has over the image, we can consider a star as a point source - the largest star in the sky - 0.044 arcsec in diameter (Betelgeuse) - is far smaller than the resolution of almost all telescopes. The image of a star will, then, be an estimator of the PSF at that point. If we suppose that the PSF varies smoothly across the detector, then at any point, the PSF shape will be an interpolation of the PSF measured by the stars. There are several schemes to carry this interpolation over the whole detector area, but most take a form of averaging the measured shape of the stars near the desired point, under some radially decreasing filter.

As direct deconvolution is a not a trivial process, several processes have been developed to address how to remove PSF distortion from images. There are, again, two different procedures to deal with the correction of the PSF: subtraction, and forward modelling. In the first, the basic hypothesis is that the PSF can be accurately described as a two dimensional profile of some sort (normally a Gaussian) and then the second moments of this distribution provide an estimate of the PSF contribution
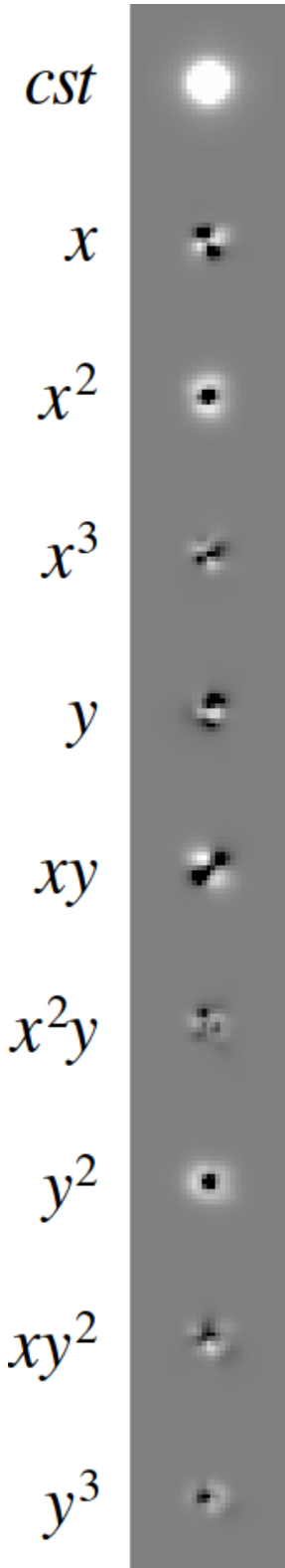
Figure 6.4: Image basis for interpolation over image locations (x, y). The PSF at each point is a sum of all images weighted by the term on the left. Here, the basis images are an example of PSF modeling for a CFHT image. *Source: Bertin [2011]*

to the shear. It is more common among different methods of
PSF modelling to use an elliptical 2D Gaussian distribution to
represent the effects, which has the advantage of being easy to
manipulate in Fourier space. However, more modern techniques
use a basis of shapes and polynomial interpolations to accurately
represent more complex optical aberrations on the PSF [Bertin,
2011]. This is crucial in space telescopes, where the blurring effect
of the atmosphere, dominant in ground-based telescopes, vanishes
and most of the contribution comes from geometric imperfections
of the optical systems.

6.5 *Adaptive Moments*

In this section, we will examine the most straightforward method
of ellipticity calculation: getting the quadrupole moments of the
image pixels. For these methods, several ideas have been published
to remove PSF distortion effects, such as Kaiser et al. [1995],
Miller et al. [2007], Hirata and Seljak [2003], Bernstein and Jarvis
[2002]. In our first practical application, we will be using HSM,
a module of GALSIM that uses re-Gaussianization and adaptive
moments, which we briefly explain below.

   The (directional) ellipticity of an object can be defined in
respect to its major and minor semi-axis as:

$$\chi = \frac{a^2 - b^2}{a^2 + b^2} e^{2\theta i} = \chi_1 + i\chi_2 \,, \qquad (6.4)$$

also represented here as $\mathbf{e} = (e_1, e_2)$, or alternatively as:

$$\epsilon = \frac{a - b}{a + b} e^{2\theta i} = \epsilon_1 + i\epsilon_2 \,, \qquad (6.5)$$

or also as $\mathbf{g} = (g_1, g_2)$, when used as an estimator for $g = \gamma/(1 - \kappa)$.[3]

   For each of these shear definitions, we can get an estimator for
the observed ϵ can calculated directly from the pixels:

[3] It is important in practical applications
to note which definition each code uses,
because both are widely used. It is also easy
to find in modules as GALSIM or NGMIX,
functions to convert one to another. NGMIX
uses (g_1, g_2), whereas LENSFIT calls
(e_1, e_2), however, using our definition of ϵ.

[4] where

$$Q_{xx} = \frac{\sum (x - \bar{x})^2 I(x,y)}{\sum I(x,y)}$$

$$Q_{yy} = \frac{\sum (y - \bar{y})^2 I(x,y)}{\sum I(x,y)}$$

$$Q_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y}) I(x,y)}{\sum I(x,y)}$$

[5] which is just to add a $w(x,y)$ term to sum terms on the moments Q_{ij}

[6] For a full description, the reader can refer to the original paper, appendix A & B, or [Bartelmann and Schneider, 2001], eqs 4.70, 4.71.

$$\epsilon = \frac{Q_{xx} - Q_{yy} + 2iQ_{xy}}{Q_{xx} + Q_{yy} + 2\sqrt{Q_{xx}Q_{yy} - Q_{xy}^2}}$$

$$\chi = \frac{Q_{xx} - Q_{yy} + 2iQ_{xy}}{Q_{xx} + Q_{yy}}$$

[4]

However, as shown in Kaiser et al. [1995], directly measuring the quadrupole moments from the pixels is very sensitive to noise. Therefore, weighing schemes are used in practice [5]. Kaiser et al. [1995] measures second moments using a circular Gaussian weight with size selected to maximize the detection significance. Bernstein and Jarvis [2002] introduced a modification, in which the moments $w(x,y)$ are iteratively *adapted* to match the shape of the object, hence the name adaptive moments. The effect of the PSF can be modelled by separating the effects into isotropic and anisotropic contributions. Kaiser et al. [1995] have shown that, using a small, highly anisotropic distribution, convolved with a large circular seeing disk that represents the isotropic contribution, the PSF effect can be then removed from objects by:

$$\epsilon^{cor} = \epsilon^{obs} - P_{i,j}^{sm} p_j \tag{6.6}$$

where p_j can be calculated as the observed ellipticity moments of the stars, that have $\epsilon^{cor} = 0$, and $P_{i,j}^{sm}$ is the responsivity of the quadrupole moments of an observed object to PSF anisotropy. The responsivity can be calculated directly from the weighted (adaptive, if it is the case) moments Q_{ij}.[6] Further refinements were introduced by Hirata and Seljak [2003] that properly accounts for the PSF anisotropies departures from elliptical Gaussians, through a process dubbed re-Gaussianization.

We used the HSM module of GALSIM to calculate ellipticities by image moments, and use re-Gaussianization to remove distortions induced by the PSFs. We take stamps from the AEGIS (All-wavelength Extended Groth strip International Survey) images cutting boxes 48×48 in size. Selecting galaxies with $(S/N) > 10, 15, 20$ in turns, from a total of ~ 2800 galaxies.
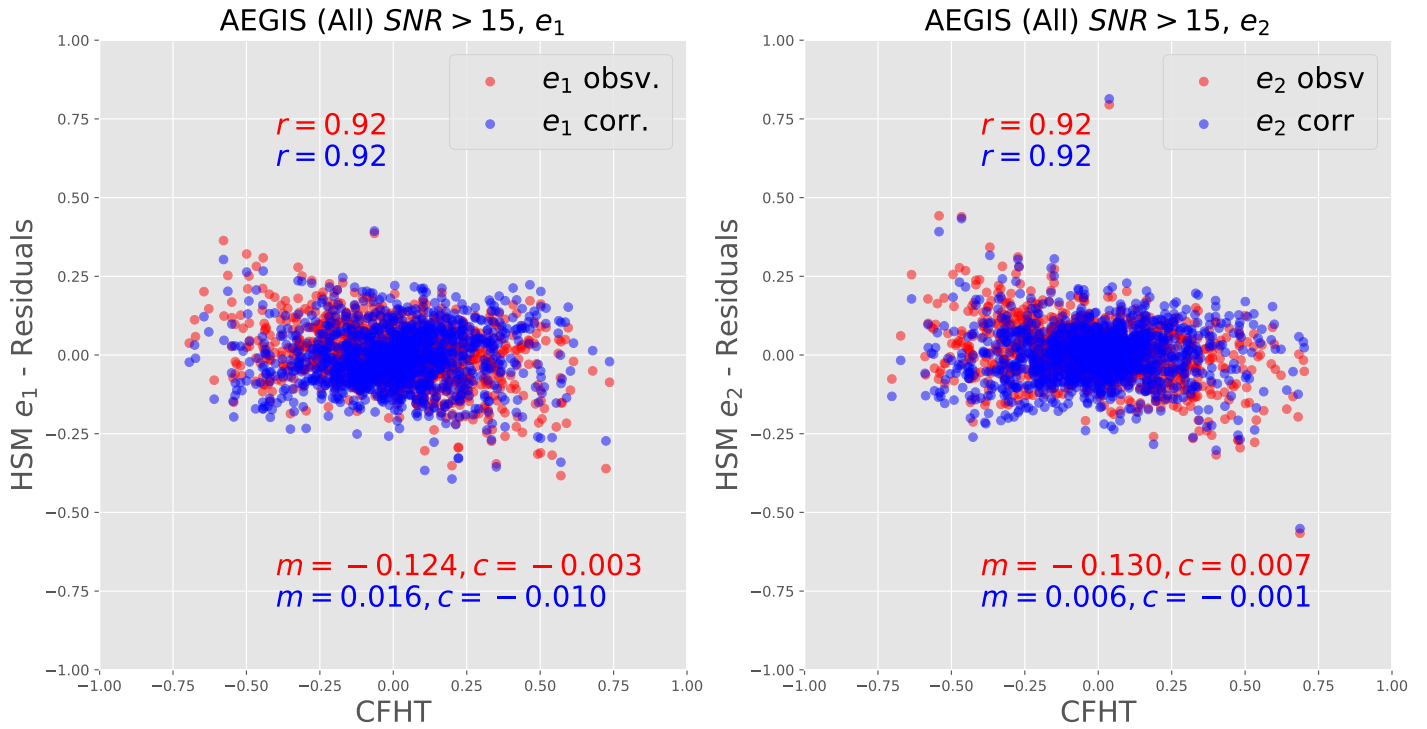
Figure 6.5: Residual results of the fitting procedure using HSM over 1287 galaxies with $(S/N) > 15$. Here we display the calculated ellipticities with and without PSF corrections, to give a magnitude of the effect. We also quote the values for the biases m and c found.

The results obtained without and with using the PSF for the $N = 1287$ galaxies with $(S/N) > 15$ is given in the figure 6.5.[7]

We have also calculated the m and c bias quantifiers and have found that the correction of the PSF significantly impacts the biases. In both cases, the multiplicative bias dropped by more than an order of magnitude.

Finally, the Pearson correlation coefficient between the CFHTLenS values and the values we calculate is $r = 0.92$, which is probably the best we can expect, considering the expected scatter in the original measurements from CFHT.

6.6 Model Fitting

The second type of method proposed to calculate galaxy shapes is fitting parametric profiles. One advantage of this method is that we can use an accurate model description of the PSF, then convolve this PSF model with the galaxy model, and then calculate the distance between this convolved model and the image data in our fitting procedure. Because of this use of data in "the model side", it is also commonly called *forward fitting*.

The description of model fitting in this section is focused on

[7]

Here we must draw an attention to a caveat: when working with the PSFs of the mini-J-PAS, there were difficulties finding accurate results. Using the data products from the survey out of the box yielded worse results than not correcting by the PSF, with much great scatter and biases - also making the fit unstable in the parametric model in the next section. As this is an ongoing work, the reasons for this are still being investigated, but is suspected to be problems with the pixel scale of the image and the PSF model. An adjustment of scale, downsampling the PSF by a factor of 2.5, provided all the results given. This is most likely due to problems with the pixel This is obviously unjustifiable as such and is intended, at this point, to be a proof-of-concept until the pipeline is fully ready.

the NGMIX code, that we will offer as a part of the pipeline. A future expansion of the same idea is a re-implementation of the code in GPUs, using parallel computing to calculate the difference between the model and all the pixels of the data during the fitting procedure at once.

In general, it is known that most regular galaxy profiles can be accurately described by the Sérsic Law [Huff and Mandelbaum, 2017]:

$$I(R) = I_e exp \left\{ -b_n \left[\left(\frac{R}{R_e} \right)^{1/n} - 1 \right] \right\} \qquad (6.7)$$

where R_e is the half-light radius, that is, the radius that encloses half of the total galaxy flux. Different n indices correspond to different types of profiles, with $n = 4$ being the de Vaucouleur's profile, typical of elliptical galaxies, and $n = 1$ the exponential profile, typical of galaxy discs. When doing parametric fits of galaxy luminosity profiles, however, especially when we are mostly interested in the overall shape, it is not only easier, but actually better to approximate these profiles by sums of Gaussians Hogg and Lang [2013]. The reason is twofold: Gaussians make it easier to work in Fourier space[8], and also, standard parametric luminosity profiles used in the literature are ill behaved near the centre [Hogg and Lang, 2013], which can complicate the fitting procedure, adding instabilities to the fit.

From a circular two-dimensional Gaussian, written as

$$f(r|\sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp(-r^2/2\sigma^2) \,, \qquad (6.8)$$

where $r = \sqrt{x^2 + y^2}$, an interpolation of Gaussians can be constructed as

$$F(r) = \sum_i \alpha_i \, f(r|\sigma_i) \,, \qquad (6.9)$$

with the α_i being coefficients that better interpolate the desired profile, and are given in Hogg and Lang [2013] for several common profiles, as the de Vaucouleur's and the Exponential[9]. The representation of profiles as a sum of Gaussians is called a *Gaussian mixture*. We use 6-Gaussian mixtures in our tests.

[8] The use of Fourier space in shear measurements is widespread among codes because convolutions become products of functions, greatly simplifying PSF treatment. Furthermore, the Fourier transform of a Gaussian function is also a Gaussian function in frequency space.

[9] The gaussian parameters are fixed and interpolate a normalised profile. All fits are performed on transformations of this interpolated curve, not the gaussians.

This radial symmetric distribution can then be transformed by the inverse of a shear matrix, as we can recall given by

$$\mathcal{A} = \begin{bmatrix} 1 - e_1 & -e_2 \\ -e_2 & 1 + e_1 \end{bmatrix} \qquad (6.10)$$

to give a galaxy shape as

$$F'(x, y) = \mathcal{A}^{-1} L(r = \sqrt{x^2 + y^2})\,. \qquad (6.11)$$

This, in turn, can be easily convolved with a PSFEx given model of the PSF, and then have the chi squared calculated from the data.

Figure 6.6: Example of fitting a single simulated galaxy image with a parametric model. The model is a sum of 6 Gaussians that best approximate an exponential profile, and the galaxy image was created in GALSIM, using routines designed for the GREAT08 [Bridle et al., 2010] shear measurement challenge. Here, the fitting code used is not NGMIX, but our own implementation of this method, running on a GPU. The data and the model are passed to the GPU, where the difference between each galaxy pixel and the model at that point is calculated in parallel. The fitting loop (a simple Nelder-Mead) is executed by the CPU, in an example of heterogeneous computation. Although this code is far from mature to be used, the preliminary evaluation is that it can calculate galaxy shapes up to an order of magnitude faster.

In our test case, we used a simple proof-of-concept procedure using NGMIX, from which we can improve adding complexity to the pipeline. The full DES Science Verification [Jarvis et al., 2016] procedure for calculating galaxy shapes uses several steps, exploring the whole likelihood space and finding maximum expectation points for the ellipticity. It also calculates the results out of the full posterior, marginalising nuisance parameters to obtain the ellipticity the same way LENSFIT does.

We use exponential profiles for galaxies and a single elliptical Gaussian to describe the PSF. The model has a total of 6 parameters for the galaxy and 6 for the PSF, which are the centre

coordinates (x, y), the ellipticity (g_1, g_2), the "size", given by the sum of the first moments: $T = Q_{xx} + Q_{yy}$, and the total flux F (which is the sum of all pixel values).

We use a standard Levenberg-Marquardt fitter, taking as initial guesses the ellipticities from the HSM, and random sorts out of our priors. For the PSF, we use a prior that is 1 magnitude smaller in shear than that of the galaxies.

To have stable fitting procedures, it is important to adopt priors, even if they are uninformative or not motivated - they just need to have sensible boundaries for the total flux, galaxy centre and the size T. We used flat priors for the T and total flux F. For the ellipticity we use the prior from Bernstein and Armstrong [2014] with average absolute ellipticity of $\sigma_i = 0.26$. Finally, we use narrow Gaussian priors for the centre around the centre of the image, with $\sigma = 1$ pixel.

We report the results together with metacalibration in the next section in figures 6.8 and 6.9.

6.7 Metacalibration

As previously mentioned, imperfections on galaxy shape measurements can lead to several type of biases. The traditional way to deal with these biases is to create realistic simulations of galaxy images, taking into account all details of the upcoming survey, and testing the method over these.

This, obviously, opens the question: how do we know a galaxy simulated image is an accurate representation of the real images? Furthermore, are the galaxies generated representative of the distribution of galaxy images we will encounter in the real experiment?

The outright impossibility of properly answering these questions draws the attention to the idea proposed in Huff and Mandelbaum [2017], to calculate the shear responsivity of a method directly on real galaxy images, and then offer an inverse transformation that remove the biases found.

In weak gravitational lensing, we use the measurement of ellipticities $\mathbf{e} = (e_1, e_2)$ to observe a shear $\gamma = (\gamma_1, \gamma_2)$ by making local averages $\langle e \rangle \approx \langle \gamma \rangle$.

Metacalibration

$$e_1, e_2 + \delta e_2$$

$$e_1 - \delta e_1, e_2 \qquad e_1, e_2 \qquad e_1 + \delta e_1, e_2$$

$$e_1, e_2 - \delta e_2$$
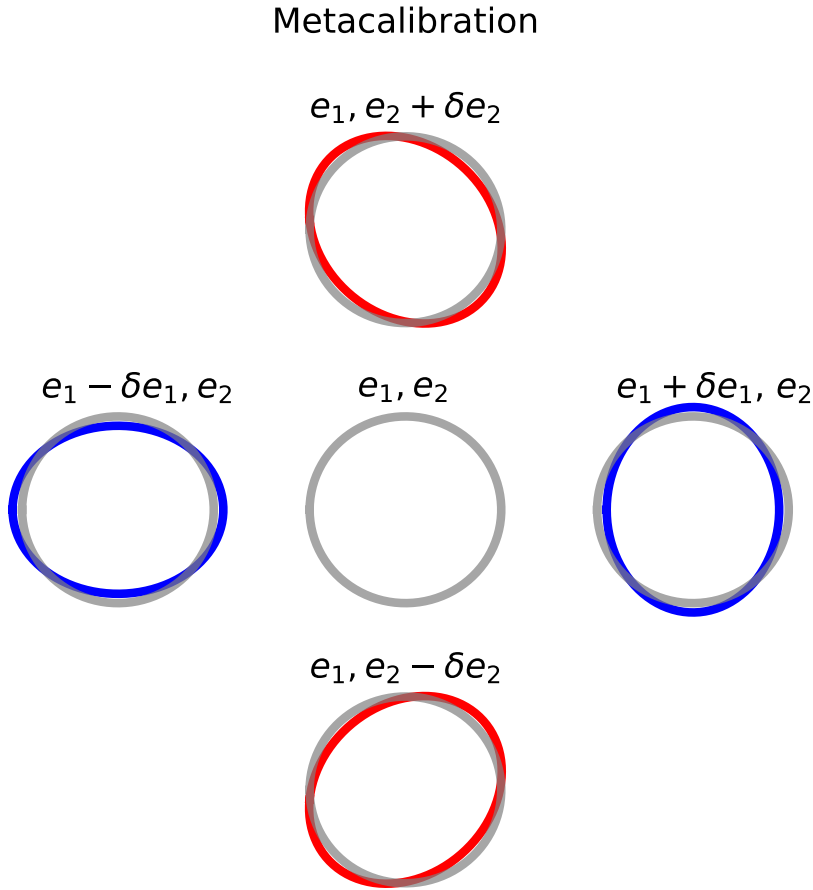
Figure 6.7: An schematic explanation of the metacalibration: We get the galaxy image, with observed ellipticity $(e_1, e2)$ (represented here as the center circle - we use a circle for convenience, but this is a galaxy of any shape). Then, we apply small shears ($+0.01$ and -0.01, in our case) this image in the e_1 (left and right) and e_2 (up and down) directions independently, and measure shears over these sheared images. The ratio between the difference in observed shape and the applied shear between them is an estimator for the responsivity **R** of the method to shear (that is, for example, the difference between e_2 measured on the right sheared image and e_2 measured on the left sheared image, divided by a $2\delta e_1$ we chose to shear those images is an estimator to $\partial e_2 / \partial \gamma_1$, and so on). By inverting this matrix, we can remove, in theory, the effect of model biases.

We can expand **e** about zero shear as

$$\mathbf{e} = \mathbf{e}|_{\gamma=0} + \frac{\partial \mathbf{e}}{\partial \gamma}\gamma + \mathcal{O}(\gamma^2)\dots . \qquad (6.12)$$

We call[10]

[10] more explicitly,

$$R = \begin{pmatrix} \frac{\partial e_1}{\partial \gamma_1} & \frac{\partial e_1}{\partial \gamma_2} \\ \frac{\partial e_2}{\partial \gamma_1} & \frac{\partial e_2}{\partial \gamma_2} \end{pmatrix} \qquad (6.13)$$

$$\mathbf{R} = \frac{\partial \mathbf{e}}{\partial \gamma} \qquad (6.14)$$

the (linear) shear responsivity. Ignoring terms of order two or higher, and taking the average we have

$$\langle \mathbf{e} \rangle = \langle \mathbf{e} \rangle|_{\gamma=0} + \langle \mathbf{R}\gamma \rangle . \qquad (6.15)$$

But $\langle \mathbf{e} \rangle|_{\gamma=0} = 0$, because we assume all deviation from isotropy is due to local shear effects and then

$$\langle \mathbf{e} \rangle = \langle \mathbf{R}\gamma \rangle . \qquad (6.16)$$

If e is an unbiased estimator of γ, R is the unit matrix, and we recover the usual relation of $\langle e \rangle \approx \langle \gamma \rangle$. When the model is biased, R deviates from unity, and we can use its inverse as a method to remove biases from **e** as

$$\langle \gamma \rangle \approx \langle \mathbf{R} \rangle^{-1} \langle \mathbf{e} \rangle . \tag{6.17}$$

We can estimate R directly from the image if we apply a known shear

$$R_{i,j} = \frac{e_i^+ - e_i^-}{\Delta \gamma_j} , \tag{6.18}$$

where e_i+ and e_i^- are the measured e_i ellipticity when the image is sheared by $+\Delta\gamma_j/2$ and $-\Delta\gamma_j/2$ (see Fig. 6.7 for an schematic explanation).

The process of applying shear to the image must be taken with care, as it represents the true shape of the object in the sky convolved with a PSF, as we have seen before. The response to shear of the true shape and the PSF is not necessarily the same, and we want to measure the effect over the actual shape. Therefore, we must deconvolve the image, apply the shear, and reconvolve it again - taking care of suppress noise amplified by the deconvolution [Sheldon and Huff, 2017]. In this, the Gaussian representation of light profiles and PSFs is of great aid, as these processes can be worked out in Fourier space. Metacalibration, however, can be applied to any method, in principle, and has been shown to be effective to all of those that were tested in [Sheldon and Huff, 2017].

As we are not here measuring local shears (for now), and just want to see the effect of metacalibration over the biases m and c, we apply an average R to all e measured to compare with non calibrated results. In practice, *this will not work*, as has been shown by [Sheldon and Huff, 2017], and more sophisticated methods of averaging must be used.

By running NGMIX over the same $(S/N) > 15$ galaxies as before we have again found that the correlation coefficient is $\rho = 0.92$ for both components. For the biases m and c, we find significant reduction after the metacalibration procedure, as shown in figure 6.8. However, in some cases the biases are still

higher than the HSM without metacalibration.

Here we have also investigated the different fields, to check if there were different levels of precision of the measurements or biases. Our results indicate that the measurement of all fields is homogeneous, finding no problems. We display these field-separated results in figure 6.9.

The performance is vital in a pipeline being created to be used over up to 10^9 galaxies. However, it is known that fitting procedures are computationally expensive. The average time NGMIX takes to perform the shear calculation using one filter, with PSF and metacalibration (with 4 counterfactual shears) is $\Delta t\, 0.15 \sim$ s per galaxy, on a Intel(R) Core(TM) i7-6900K CPU @ 3.20GHz, in a single thread. That is 108 days of computing using all 16 threads for 10^9 galaxies. By adding other filters, this could easily take up a whole year of computing in a desktop machine. However, by using a $\sim$ 1000 core cluster[11] a 5 filter metacalibration job is expected to take less than 9 days, which is not out of the realm of the achievable.

At this time, many settings have been done *ad hoc* on the code, and need to be addressed properly. More tests, both with available data and simulations are still needed to assess the robustness and

Figure 6.8: Residual results of the fitting procedure using NGMIX over 1287 galaxies with $(S/N) > 15$. Here we display the calculated ellipticities with and without metacalibration, to give a magnitude of the effect. We also quote the values for the biases m and c found.

[11] ...such as AlphaCrucis at IAG/USP.

reliability of the results.
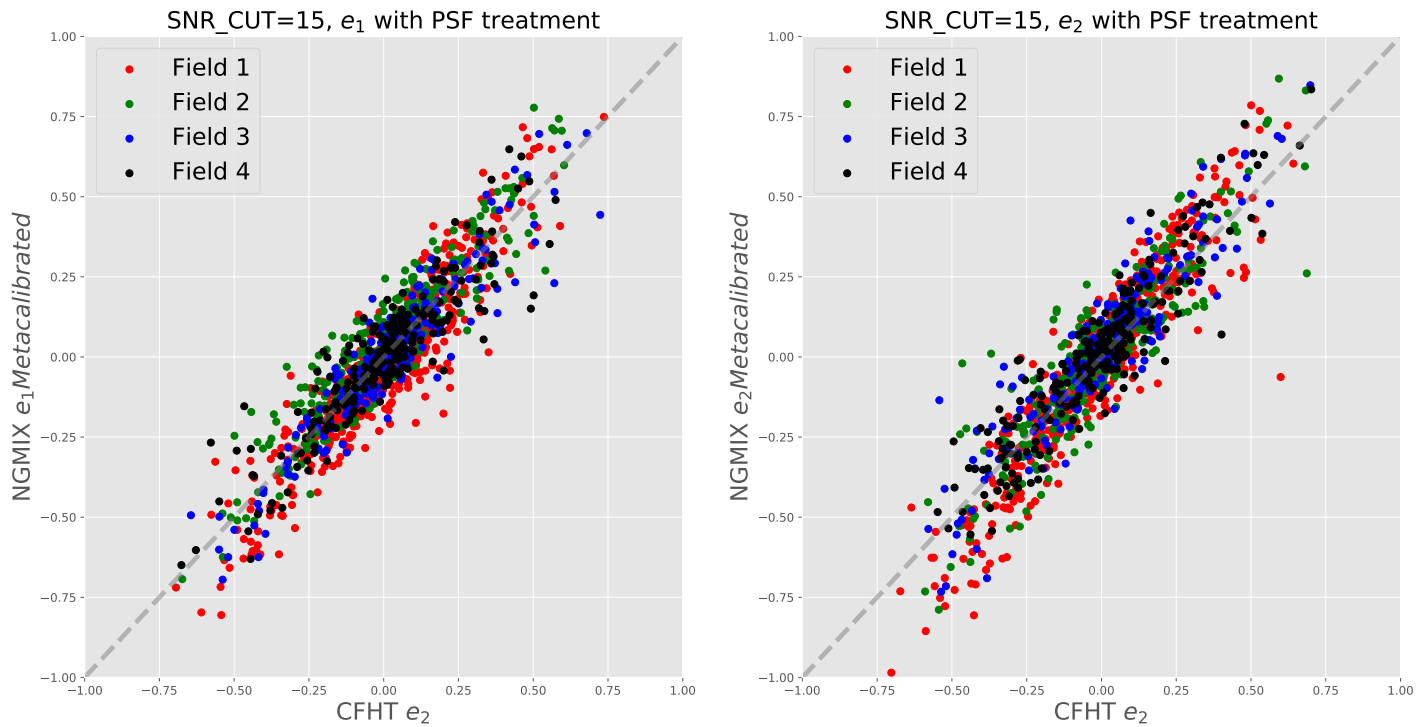
Figure 6.9: Results of the fitting procedure using NGMIX separated by fields, both with metacalibration. We find no significant difference among the four fields.

6.8 *Convolutional Neural Networks*

The use of machine learning methods, and more specifically, neural networks to astronomy problems is not new [Odewahn et al., 1992, Weir et al., 1995, Bell et al., 2006]. However, the last two years have seen an explosion in implementations of deep neural network solutions to a variety of tasks [Lanusse et al., 2018, Soo et al., 2017, Ribli et al., 2019, Ishida et al., 2019]. The main driver for the appeal is the recent development and publication of easy to use, open source libraries such as TENSORFLOW [Abadi et al., 2015] and KERAS [Chollet et al., 2015]. Also, the continuing trend of improvement in parallel computing, by means of graphic cards (GPUs) have contributed in making formerly expensive computational tasks - such as backpropagating huge networks, with millions of parameters - accessible in time and costs for the casual user.

Machine learning ML differentiates itself from other methods of computation by performing an specific task without using explicit instructions. By relying on sufficiently general models, machine

learning algorithms can learn (that is, be fitted as parametric functions) to perform complex tasks. It tends to be particularly apt to perform pattern recognition type of tasks, which are not easily describable in mathematical models, but can be done if a function with enough degrees of freedom is fitted to a large corpus of data.

The main advantage of using ML is that the model can be built very agnostic to the problem, without having to conceive a specific functional form, motivated by plausible assumptions[12]. Instead, you can have a function so generic that it can be fitted to calculate the probability that an array contains an image of a car, or with the same topology, calculate parameters of a cosmological model from a convergence map.

Care must be taken, therefore, to avoid the von Neumann aphorism that "*with four parameters, I can fit an elephant, and with five I can make him wiggle his trunk*"[13]. In ML models, specially deep neural networks, the function can be generic enough to fit any input to any results very precisely. That the results are representative of the inputs [14] is the role of the researcher to check.

Furthermore, ML strategies normally have *millions* of parameters. It is important that the quantity of data points available to perform the fit is large enough. That helps ML methods to better explore the normally large space of parameters, and avoid being trapped in minima.

The job of the training is to change parameters in the model so that for any possible input $\vec{x}$, the output $\vec{y}$ is as close as possible to "real" values $\vec{\lambda}$, on average - just as any fitting procedure. The final product of the training procedure is a model M that takes an input $\vec{x}$ and returns a *prediction* $\vec{x}$ on previously unclassified data correctly. Therefore, to get a ML model to *learn* a task is equivalent of fitting a model to data. The model M has parameters $\vec{\theta}$, and receives as input data an array $\vec{x}$.

Suppose that we have a vector $\vec{\theta}$ in the model parameter space. Then, having some input data $\vec{x}$ we can calculate

$$L = d(M(\vec{\theta}, \vec{x}), \vec{\lambda})\,, \qquad (6.19)$$

where d is a metric, a function that calculates a real positive value that represents the distance between the model M calculated with

[12] Contrast, for example, with our model in chapter 4, where in equation 4.21, each term has a very definite explanation.

[13] For an example of such an "elephant", see [Mayer et al., 2010]

[14] in other words, if they are *meaningful* in some way

the input $\vec{x}$ with parameters $\vec{\theta}$ and the value of the label $\vec{\lambda}$.

For all available data, that is, a set of $\{\vec{x}\}$ and another of $\left\{\vec{\lambda}\right\}$, we want to minimise the overall loss $\mathcal{L}$ that we have by adopting parameters $\vec{\theta}$, given by:

$$\mathcal{L} = \frac{1}{N} \sum_{\{\vec{x}\},\{\vec{\lambda}\}} d(M(\vec{\theta}, \vec{x}), \vec{\lambda})\,. \tag{6.20}$$

where N is the number of elements in $\{\vec{x}\}$. To do so, we can proceed with an iterative process: starting from some initial choice of $\vec{\theta}$ we calculate $\mathcal{L}$, and then figure out the direction in the parameter space $\vec{\theta}$ we need to move to achieve lower $\mathcal{L}$. An optimal strategy to do so, is to follow the gradient of $\mathcal{L}$ to arrive at a new $\vec{\theta}$,

$$\vec{\theta}_{t+1} = \vec{\theta}_t - \alpha \nabla \mathcal{L}\,, \tag{6.21}$$

where α is called the learning rate. This optimisation technique is called *gradient descent* and is widely used in ML.
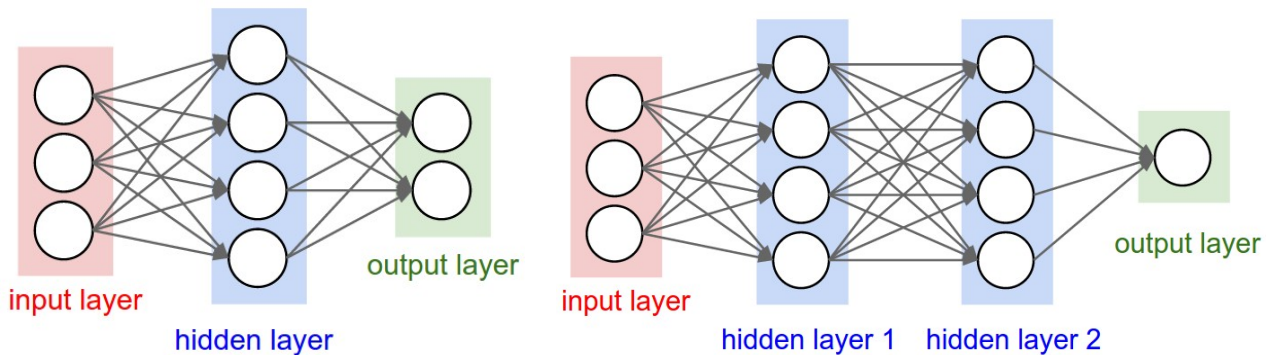
Figure 6.10: Schematic representation of two neural networks. The input layer, in red, receives an array of numbers (the data we want to classify, or in a regression, to calculate some value over). These numbers are then sent as arguments to the neurons on one or more hidden layers in blue. The neurons sum the contributions from each arrow with weights, and pass them to an activation function. The result of this activation function is then passed on to the next layer, until we reach the output (in green). On the right, a *deep* neural network, containing two hidden layers. *Source: http://cs231n.stanford.edu/*.

The choice of an appropriate learning rate is important, because too large learning rates will prevent the model from converging and too small learning rates risk trapping the method in a local minimum. A neural network (see figure 6.10) is a machine learning strategy that uses a set of compositions of functions, linked in successive layers that connect an N_i-dimensional input to a N_o-dimensional output. The inputs can be anything from images, time series, discrete values, etc., but are usually large number arrays. The outputs are usually called *labels*, owning to the task of pattern recognition these networks are often used, and can be probabilities of the input array containing a particular

feature (in a classification task), or some continuous value derived from the input (a regression task).

Neural networks are named for the inspiration from biological neurons [Kleene, 1956], that receive and propagate signals. In an artificial neural network, each neuron is a function, sitting on a layer, that receives inputs from the previous layer, and returns values for the next one. An artificial neuron is a mathematical model made up by a weighted sum of its inputs and a transfer function that are combined to give an output:

$$x_j^l = f\left(\sum w_j^k x_k^{l-1} + b_j^l\right) . \qquad (6.22)$$

[15]

In each layer (l), each j neuron receives x_k^{l-1} from all k neurons in the previous layer ($l-1$). They sum each x_k^j signal with a weight w_k, pass the result to an activation function f, and then send x_j^l to all the neurons in the next ($l+1$). In this next layer, each neuron get y_j from each neuron in the previous (l) layer, and perform the same procedure. The input layer, of course, receives its x_i inputs from the data, and the output layer values y_i are compared to a set of true values λ_i. In essence, this is the same as getting all elements of layer $l-1$ as a vector, multiply it by matrix of weights (as matrix multiplication), sum a vector of constants, and then apply f to each of the results. The resulting vector is then passed onto the next layer (in this way, this is called a feed-forward network).

Without non-linear steps, a neural network would be equivalent to a linear regression with lots of parameters. To be more flexible, neural networks incorporate activation functions f to each neuron. Most commonly used activation functions are the step function, the sigmoid function, and ReLu.[16] To calculate the gradient $\nabla \mathcal{L}$ we need to successively apply the chain rule to every layer of the network. This scheme is called *backpropagation* [LeCun et al., 1989] . To explain how backpropagation works, we will first consider a simple neural network, with a single value y as output and just one neuron per layer, with l layers (l is the layer number of the output layer).

The loss associated with just one input (considering square error

[15] The reader might remember I could have used Einstein notation as in section 1.1 with $w_j^k x_k^{l-1}$, implying a sum over k. This is where TENSORFLOW gets its name from.

[16] The most commonly used activators are:
- Rectifier Linear Unit (ReLU): $f(x) = max(0, x)$
- Sigmoid function: $f(x) = e^x/(e^x + 1)$
- Step function: $f(x) = 1$ if $x \geq 0$ or $f(x) = 0$ otherwise.

The shape of each of these activators is seen below. The argument is the weighted sum of inputs. The activator determines how much a neuron respond to the input.
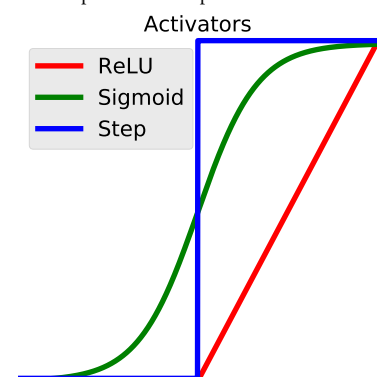
Activators

as our metric) is

$$L = (y - \lambda)^2 \tag{6.23}$$

Now, from equation 6.22, the value of y is given by

$$y = x^l = f(w^l x^{l-1} + b^l) = f(z^l) \tag{6.24}$$

where we call the argument of the activator as z^l for simplicity.

Then, by the chain rule, the derivative of the loss L to the weight w is:

$$\frac{\partial L}{\partial w^l} = \frac{\partial L}{\partial y} \frac{\partial y}{\partial z^l} \frac{\partial z^l}{\partial w^l} \,. \tag{6.25}$$

[17]

Now, if we want to know how much L changes by a change in the weight of the previous neuron, we know that

$$x^{l-1} = f(w^{l-1} x^{l-2} + b^l) = f(z^{l-1}) \,,$$

then,

$$\frac{\partial L}{\partial w^l} = \frac{\partial L}{\partial y} \frac{\partial x^l}{\partial z^l} \frac{\partial z^l}{\partial x^{l-1}} \frac{\partial x^{l-1}}{\partial z^{l-1}} \frac{\partial z^{l-1}}{\partial w^{l-1}} \,. \tag{6.26}$$

In other words, we just have to follow the links backwards adding more terms in the chain rule derivative until we arrive at the desired weight we are derivating in respect to. Moreover, all of the terms are either already there, or are easy to calculate.

When the network is not one dimensional, not much changes. We now need to sum all paths that connect the link where the weight w_j^k is to the loss L (that is, all paths from w_j^k to all nodes in the output layer. To do so, note that each node at layer l contributes with a factor $f'(z^{l-1})$, and each link between l and $l-1$ represented by a weight w_j^k, j being a neuron index in the l layer, k in the $l-1$, contributes with a factor x_k^{l-1}.

Finally, since the total loss function $\mathcal{L}$ is just the average of all L for each of the inputs in $\{\vec{x}\}$, so the derivative of $\mathcal{L}$ in respect to a weight w_j^k will also be an average of all derivatives of L to w_j^k. The vector made with all derivatives of $\mathcal{L}$ in respect of each of the weights of M is the gradient we need for gradient descent.

---

[17] Explicitly

$$\frac{\partial L}{\partial y} = 2(y - \lambda)$$

$$\frac{\partial y}{\partial z^l} = \frac{\partial x^l}{\partial z^l} = f'(z^l)$$

$$\frac{\partial z^l}{\partial w^l} = x^{l-1}$$

or

$$\frac{\partial L}{\partial w^l} = 2(y - \lambda) f'(z^l) x^{l-1}$$

[18] The two first terms we already have, the ones missing are:

$$\frac{\partial z^l}{\partial x^{l-1}} = w^{l-1}$$

$$\frac{\partial x^{l-1}}{\partial z^{l-1}} = f'(z^{l-1})$$

$$\frac{\partial z^{l-1}}{\partial w^{l-1}} = x^{l-2}$$

or

$$\frac{\partial L}{\partial w^{l-1}} = 2(y - \lambda) f'(z^l) w^{l-1} f'(z^{l-1}) x^{l-2} \,,$$

and so on.

One specific type of neural network, used in many applications related to imaging, is the convolutional neural network (CNN). CNNs are inspired by research on how the brain interprets images [Fukushima, 1980] and uses convolutional filters as layers of neurons, in which the filters are the parameters to be fitted by the backpropagation.

$$(-1 \times 3) + (0 \times 0) + (1 \times 1) +$$
$$(-2 \times 2) + (0 \times 6) + (2 \times 2) +$$
$$(-1 \times 2) + (0 \times 4) + (1 \times 1) = -3$$

Figure 6.11: How a convolution filter works: the value -3 is obtained by multiplying each term in the convolution kernel by the corresponding pixel in the source image. Then the convolution moves according to the stride to cover the image. *Source:* Towards Data Science
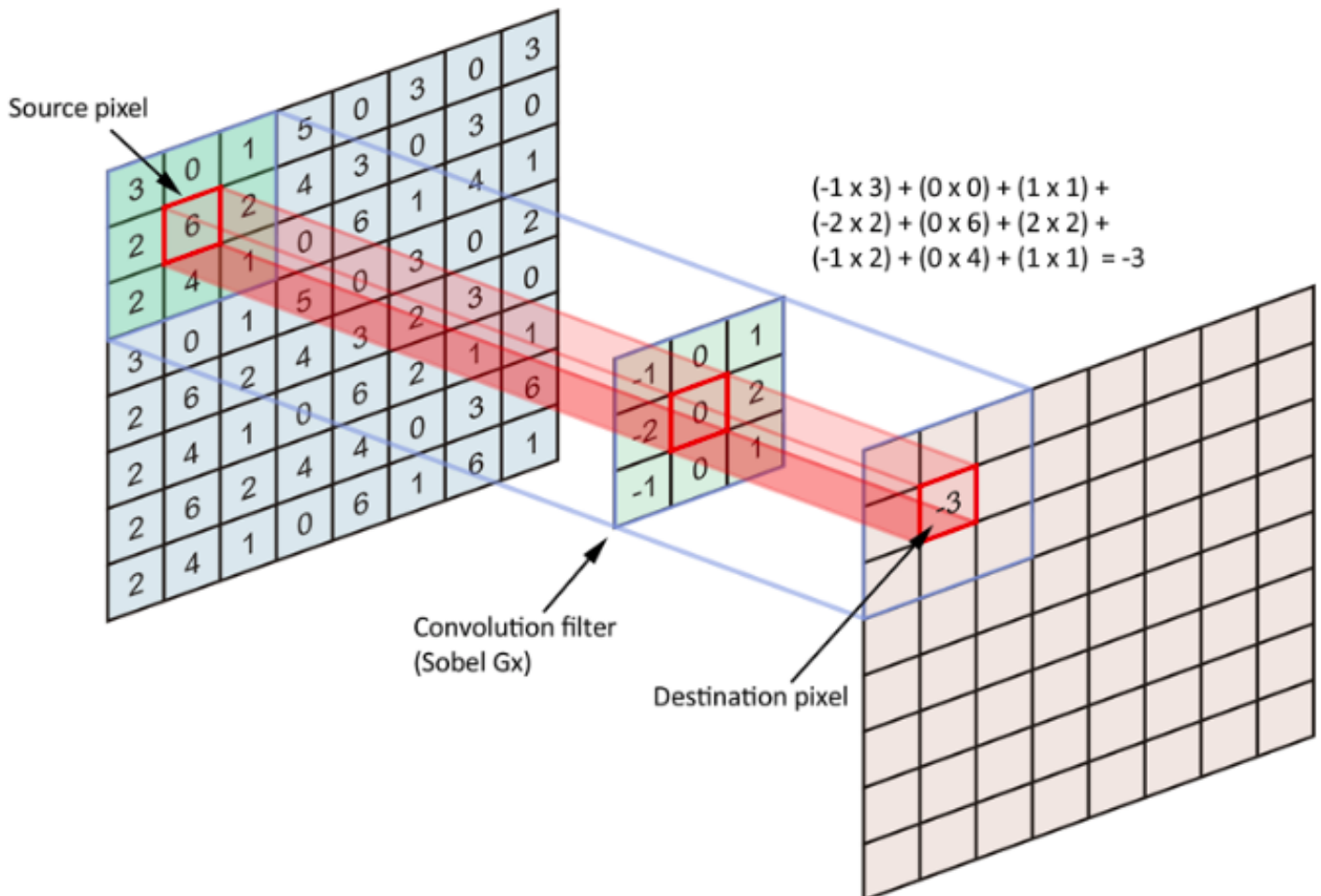
A convolution filter is the discrete version of the convolution integral, and acts as the following: take an input I that is a $n \times m$ sized matrix (an image), then cut a rectangular subset of shape $a \times b$ of this. We then do a dot product[19] between this and another $a \times b$ shaped matrix (called the convolution filter) and then store this at a place $I'_{k,l}$ of a new matrix I'. Now, move s_x (which is called the *stride*) columns to the right, and do this again, to calculate $I'_{k,l+1}$. When the k row is over, move s_y rows below, and restart, see figure 6.11. Beyond that, if you want to reproduce the size of the original input, you must define a padding that

[19] that is $p = \sum_{all} R_{i,j} C_{i,j}$, where R is a rectangular subset of I, and C is the convolution filter.

will accommodate your filter when it convolves near the borders. This is usually additional rows and columns of zeros. The idea behind convolution filters (also commonly called *convolution kernels*) is that they represent features found in images. For example, if we want to train a convolution network that identifies squares, some of the convolution filters will eventually look like corners. The convolution being applied in figure 6.11 is a classical example of an edge detection kernel [Sobel and Feldman, 1968]. Another type of filter that is important in CNNs is the pooling filter [Yamaguchi et al., 1990]. Pooling filters are used to reduce dimensionality of layers by creating a smaller layer out of the previous one, usually by selecting either the maximum value or an average out of a rectangular subset.

Figure 6.12: An example of a maxpooling filter.*Source: Wikipedia*
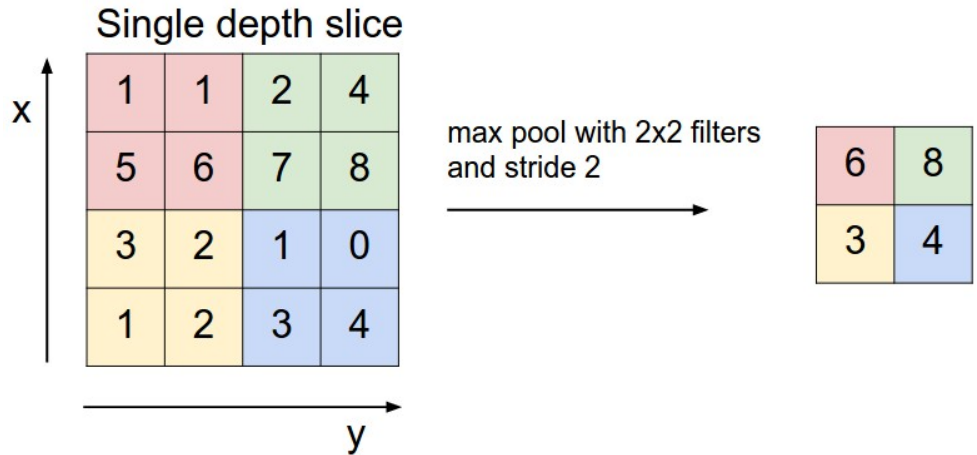
Following Ribli et al. [2019], we adopt a convolutional neural network design made of 13 successive convolution filters applied to the image, with 3 MaxPooling and 1 average pooling filters after sets of 2 to 3 convolutions. This *topology* is presented better on table 6.1. The width of the convolution filters is 3×3. Each convolutional layer is followed by batch normalization and each neuron use ReLU activators, so that the network approximate non-linear functions. The pooling filters downsample the image and aggregate information for the next layers. The final layer, represent the ellipticity outputs e_1 and e_2 that have to be trained against CFHTLenS "label" values.

The input at the beginning of the network are 48×48 "stamps" centered on individual galaxies - though given the size of the

stamp, there are several with more than one object. Again we split the tests in all galaxies with $(S/N) > 10$, 15, and 20, but focus our results in $(S/N) > 15$, which have been shown to be more stable.

#	Layer	Layer Output
1	Convolution (3×3)	$50 \times 50 \times 64$
2	Convolution (3×3)	$50 \times 50 \times 64$
-	MaxPooling (2×2)	$25 \times 25 \times 64$
3	Convolution (3×3)	$25 \times 25 \times 128$
4	Convolution (3×3)	$25 \times 25 \times 128$
-	MaxPooling (2×2)	$12 \times 12 \times 128$
5	Convolution (3×3)	$12 \times 12 \times 256$
6	Convolution (1×1)	$12 \times 12 \times 128$
7	Convolution (3×3)	$12 \times 12 \times 256$
-	MaxPooling (2×2)	$6 \times 6 \times 256$
8	Convolution (3×3)	$6 \times 6 \times 512$
9	Convolution (1×1)	$6 \times 6 \times 256$
10	Convolution (3×3)	$6 \times 6 \times 512$
-	MaxPooling (2×2)	$3 \times 3 \times 512$
11	Convolution (3×3)	$3 \times 3 \times 1024$
12	Convolution (1×1)	$3 \times 3 \times 512$
13	Convolution (3×3)	$3 \times 3 \times 1024$
-	Average Pooling (3×3)	$1 \times 1 \times 1024$
14	Dense	2

Table 6.1: This is the full architecture of the network used. The total number of fitting parameters of the model is: 13.352.002.

Out of the selected galaxies, we choose $1/5$ to serve as a test sample, to check the result of the trained network. As the network has millions of parameters, all galaxies used as a part of the training will obviously have very well "predicted" ellipticities by the trained network. Thus, we have to test the quality of the ellipticity predictor (the trained network) on galaxies that were *not* used as a part of the training.

We have also tested fits using r and i, and g, r and i image filters at the same time, using multiple channels. However, the r-only has provided much better results as of this date.

Additionally, due to the low number of input galaxies, we use data augmentation. By flipping galaxy images in all 3 possible ways (LR mirror, UD mirror, and transposition) and transforming the shears appropriately, we create an input catalogue 8 times larger than the original. Using this, the network can also learn symmetries that are intrinsic to the problem.

We train the neural network not by calculating the full gradient $\mathcal{L}$ using all data at once, due to memory restrictions of the GPU,
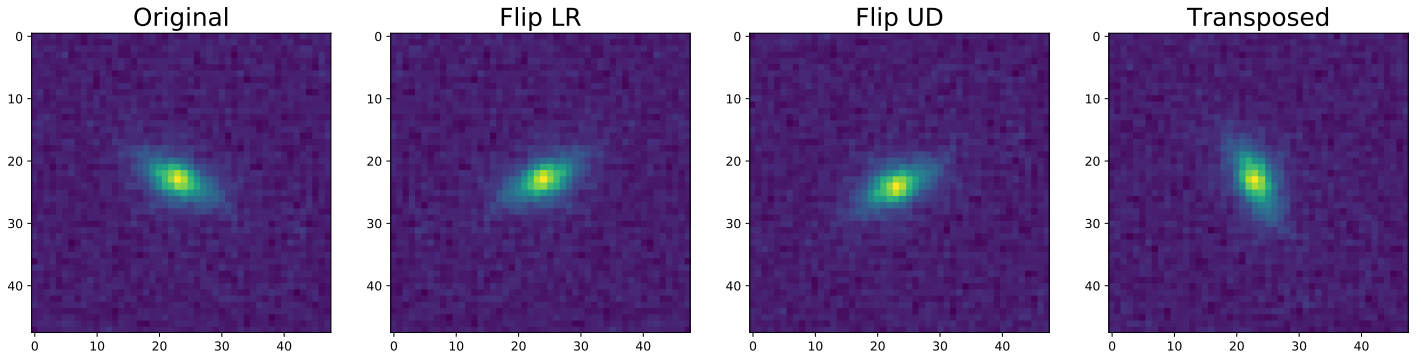
Figure 6.13: Using a high (S/N) galaxy to showcase the transformations used during data augmentation. The original ellipticities are $e_1 = 0.34$ and $e_2 = 0.38$. Both LR and UD flips reverse the sign of e_2, transposition reverses e_1.

but divide them into randomly chosen batches of 512 galaxies at a time, reshuffling them at each step. This in practice transforms the gradient descent into *stochastic* gradient descent, and although this tends to slow down the fit (in terms of numbers of steps), it does not cause problems to the procedure.
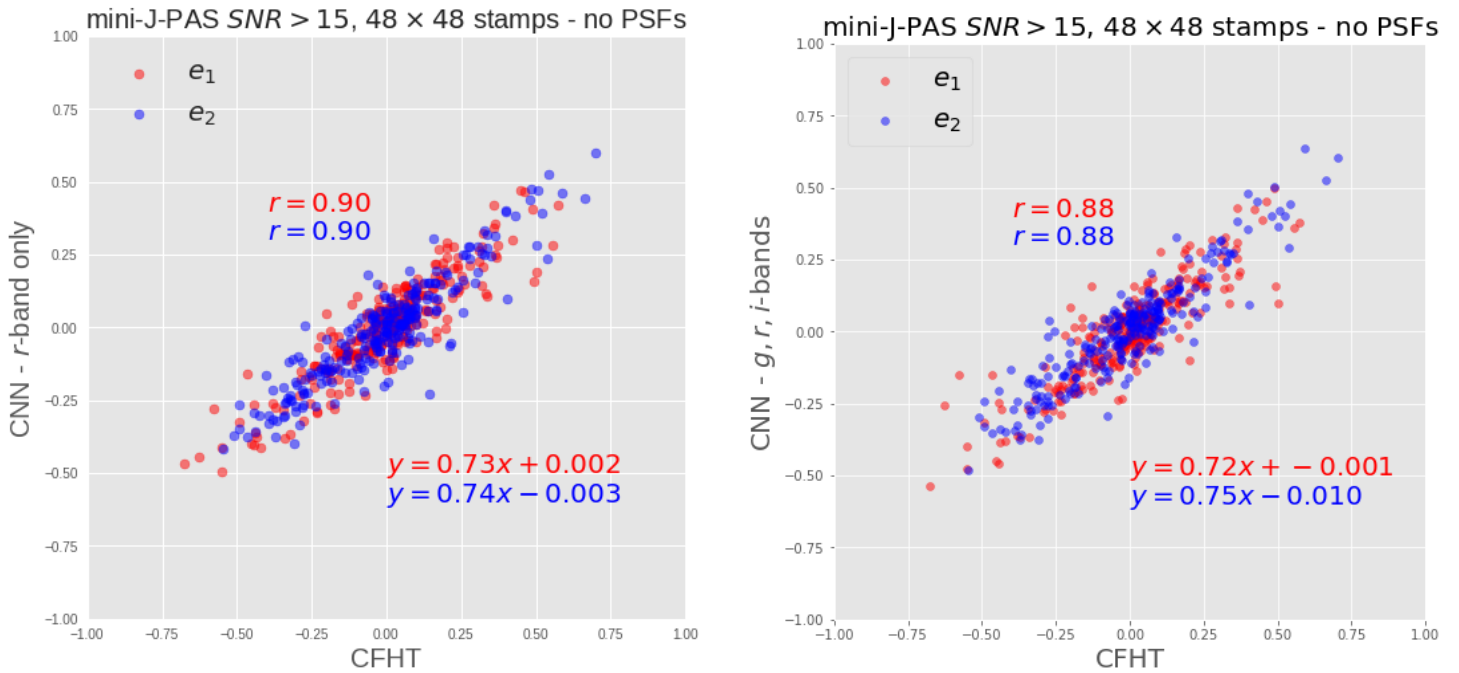
Figure 6.14: Results of CNN predictions over the test sample. We see that adding more filters did not result in an improved precision.

We have tested many schemes for controlling the learning rate, either by gradual decay each step or by reducing the learning rate after sets of 40 epochs (one epoch is the complete iteration of mini-batches, that cover all galaxies). We use a total of 120 epochs, with a learning rate of 0.005 in the first two, and 0.0005 in the final set of 40. We use a mean squared error loss function, instead of mean absolute error as in Ribli et al. [2019], as it has shown better loss decay in our tests.

As of this moment, we have not used PSF information together with the galaxies, and do not attempt to correct for that. As such, we are more interested in the Pearson correlation coefficient obtained, instead of biases. In the next iteration, however, PSFs will be introduced as new channels, running parallel with the images as additional information.

Some other suggestions being explored are: using stars together with galaxies and their PSFs, using $e_1 = e_2 = 0$, using counterfactual shears as in metacalibration in data augmentation.

We have also analysed the behaviour of the cost function along the training procedure using tensorboard. What we have concluded is that the number of epochs is adequate, and the model seems to be resilient to overfitting. The loss function stabilises before the number of epochs chosen and the validation set loss does not increase.
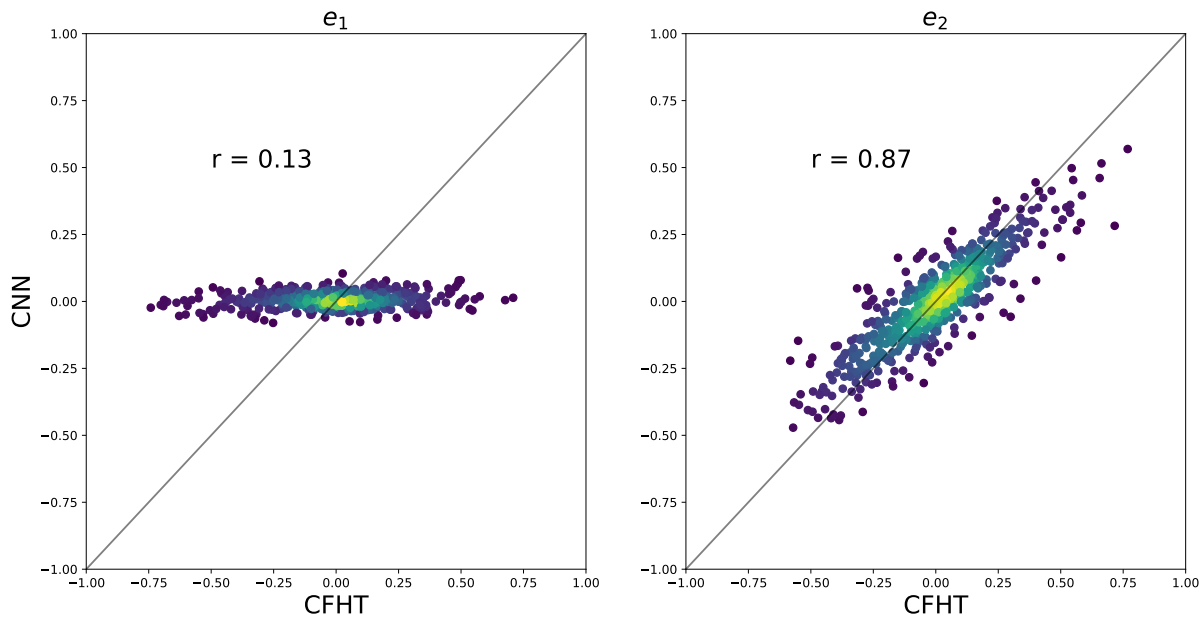
Figure 6.15: Example of fit with galaxies with $(S/N) > 10$ - at times, the training method would ignore one of the components and just fit the other.

However, there have been moments of fit instabilities. At times, the network would fit one of the ellipticity components and not the other. this is displayed in figure 6.15. This is possibly the loss function being trapped in a local minimum.

Finally, one problem to be addressed in the future is error determination. It is not clear how to properly quantify the error.

One suggestion is to get the dispersion between the test labels and the predictions, and then add this to the error bars in the predictions. This, however, gives very large error bars that might be overestimating the error.

6.9   *Discussion*

As we have seen, there are multiple options available to carry shape measurements that can be implemented, profiting from past surveys and the vast literature available. All of them require attention to detail, input from past surveys and knowledge of simulating galaxy images. However, the open source availability of these codes makes it possible to offer multiple solutions so that we can cross check the validity of the results.

We are now approaching the final phase to validate the pipeline, which will take some time, but can definitely be available before the JP-CAM becomes ready to conduct the full J-PAS. The full realisation of the J-PAS lensing capabilities will depend on the stability of the PSF of the camera, which cannot be evaluated by the pathfinder survey.

At optimum conditions, all methods performed about the same in regard to correlations between the CFHT data and our obtained shape measurements. NGMIX has been displayed more resilience to low (S/N), as shown in the table 6.2

Table 6.2: Summary of all shape measurement biases results. It must be noted that the neural network has uncomparable larger ms due to absence of correction of the PSF.

Method	S/N	r	(m, c)
	10	0.89	$(0.003, 0.02)$
HSM	15	0.92	$(0.01, 0.01)$
	20	0.92	$(0.04, 0.01)$
	10	0.87	$(0.09, 0.01)$
NGMIX	15	0.92	$(0.08, 0.01)$
	20	0.92	$(0.07, 0.01)$
	10	0.57	$(0.61, 0.03)$
CNN*	15	0.90	$(0.16, 0.003)$
(r-only)	20	0.88	$(0.15, 0.06)$

The huge volume of data of upcoming surveys makes it mandatory for us to evaluate the scalability and performance of the codes tested. HSM performed each galaxy measurement in $\Delta t = 0.005$ s, NGMIX performed in $\Delta t = 0.02$ s when not doing metacalibration and, $\Delta t = 0.15$ s with metacalibration. The CNN takes about 10 minutes to train with a single nVidia

GeForce GTX 1080 Ti, and predicts ellipticities at a rate of $\Delta t = 0.0003$ s per galaxies. The CNN is by far the easiest to scale. A total 1 million galaxy training set could take less then a day with several GPUs, and then the performance per galaxy is an order of magnitude better than HSM, the second best.

7

Conclusion

WITH THE BEGINNING OF THE MOST ADVANCED BATCH OF NEW DARK ENERGY EXPERIMENTS THE NEXT YEARS a huge amount of new data will be available to study the most prominent questions in extragalactic astrophysics. At the same time, this data may provide new paths to study particle physics - from the masses of neutrinos to the nature of dark matter.

7.1 *Developments in Cluster Astrophysics & Cosmology*

Regarding the question of dating the formation of galaxy systems, several indicators have been found that correlate with concentration, the current prime indicator for early mass accretion histories. However, the complex dynamical environment of galaxy groups and clusters makes a definitive classification elusive. New studies, with a greater number of galaxy systems will aid in further investigating these relations.

The dynamical state and, specially, the evolution of galaxy clusters from collapse to relaxation is also poorly understood. It is not known how to properly reconcile the time irreversibility of the fast relaxation of the collapse with the reversible nature of freefalling galaxies in clusters [Beraldo e Silva et al., 2019].

Furthermore, the shape of the dark matter halos have been extensively studied, both in simulations and observations. However, there is still no derivation of the shape of the profile from physical principles. As the mass distribution of clusters is largely self-similar, it would be reasonable to expect that it follows from fundamental physics [Beraldo e Silva et al., 2013].

In cosmology, for a long time, estimates for the Hubble-Lemaitre parameter (H_0) varied between 50 and 100 km $\cdot$ s$^{-1} \cdot$ Mpc^{-1}[Overbye, 1993]. This question was among the main missions of the Hubble Space Telescope, and precise measurements were achieved [Freedman et al., 2001]. However, more recent measurements from Planck Collaboration [2018], together with new measurements by the HST [Pietrzyński et al., 2019] are at a tension of 4.4σ. Consistently, measurements made by examining the large scale structure of the universe, instead of the CMB, tend to give higher values of H_0.

On the relationship between cluster astrophysics and particle physics, many of the open questions lie on how good galaxy cluster measurements can get. To push that to the limit, deeper surveys must be carried, to fetch more data from gravitational lensing. Not only that, wider areas will be needed to enhance the numbers to drive down statistical uncertainties.

With less trouble with random scatter, however, techniques will have to be perfected as to not introduce biases from systematic effects [Mandelbaum et al., 2018].

7.2   *Upcoming Surveys*

The success of the Sloan Digital Sky Survey paved the way for a great number of new galaxy surveys in the optical spectrum, and spawning several spin offs that continue to be a driving force in extragalactic astrophysics. So many new surveys are being conducted, proposed, or being prepared for the near future that any attempt to make an exaustive list would be impossible. Here are some of those that will play a major role in cosmology and extragalactic astrophysics.

On the weak gravitational lensing, Euclid [Laureijs et al., 2011] and LSST [LSST Dark Energy Science Collaboration, 2012] are the major players. Euclid is intended to cover a total of 15000 deg^2 of the sky, with 30 galaxies per arcmin^2. This is about the same that what is expected from LSST [Chang et al., 2013].

WEAVE (WHT Enhanced Area Velocity Explorer) is a next-generation wide-field spectroscopic survey [Dalton et al., 2014].

The J-PAS [Benítez et al., 2015] provides a middle ground

between spectroscopic and photometric surveys, with a 60-filter camera that will cover 8000 deg^2 of the sky. This will provide a large number of galaxy clusters, and galaxies with precise redshifts to use in weak gravitational lensing studies.

Outside the optical spectrum, but crucial to cluster astrophysics, eROSITA [Merloni et al., 2012], a space-based X-ray observatory, will image the full sky in X-ray ($2-10$keV) over a 7-year period. It will provide invaluable data to understand galaxy cluster dynamics such as mergers.

With these and many other surveys, the interdisciplinary relationship between Statistics, Astronomy, and Computer Science will deepen to a new level. Future studies will not depend *only* on programming skills or knowledge of basic statistical analysis, but having access and familiarity with cutting-edge modern techniques and tools.

Furthermore, a new level o collaboration will be required, as new instruments are expensive, and the work force needed to perform all steps increases accordingly. These new frontiers demand new paradigms of making science, while transforming our own paradigms of understanding nature.

Part III

References and Index

Bibliography

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo,
Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis,
Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian
Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard,
Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath
Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga,
Sherry Moore, Derek Murray, Chris Olah, Mike Schuster,
Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar,
Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda
Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin
Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-
scale machine learning on heterogeneous systems, 2015. URL
https://www.tensorflow.org/. Software available from
tensorflow.org.

J. A. L. Aguerri, M. Girardi, W. Boschin, R. Barrena, J. Méndez-
Abreu, R. Sánchez-Janssen, S. Borgani, N. Castro-Rodriguez,
E. M. Corsini, C. Del Burgo, E. D'Onghia, J. Iglesias-Páramo,
N. Napolitano, and J. M. Vilchez. Fossil groups origins.
I. RX J105453.3+552102 a very massive and relaxed system
at z ~ 0.5. *Astronomy & Astrophysics*, 527:A143, March
2011. DOI: 10.1051/0004-6361/201015364. URL http:
//adsabs.harvard.edu/abs/2011A%26A...527A.143A.

The WFIRST Project and Formulation Science Investigation
Teams Akeson, Rachel L. The Wide Field Infrared Survey
Telescope: 100 Hubbles for the 2020s. *arXiv e-prints*, art.
arXiv:1902.05569, Feb 2019. URL https://ui.adsabs.
harvard.edu/abs/2019arXiv190205569A.

Andreas Albrecht, Gary Bernstein, Robert Cahn, Wendy L.
Freedman, Jacqueline Hewitt, Wayne Hu, John Huth, Marc

Kamionkowski, Edward W. Kolb, Lloyd Knox, John C. Mather, Suzanne Staggs, and Nicholas B. Suntzeff. Report of the Dark Energy Task Force. *arXiv e-prints*, art. astro-ph/0609591, Sep 2006. URL https://ui.adsabs.harvard.edu/abs/2006astro.ph..9591A.

Steven W. Allen, August E. Evrard, and Adam B. Mantz. Cosmological Parameters from Observations of Galaxy Clusters. *Annual Review of Astronomy and Astrophysics*, 49(1):409–470, Sep 2011. DOI: 10.1146/annurev-astro-081710-102514. URL https://ui.adsabs.harvard.edu/abs/2011ARA&A..49..409A.

R. A. Alpher, H. Bethe, and G. Gamow. The Origin of Chemical Elements. *Physical Review*, 73:803–804, April 1948. DOI: 10.1103/PhysRev.73.803. URL http://adsabs.harvard.edu/abs/1948PhRv...73..803A.

J. Annis, M. Soares-Santos, M. A. Strauss, A. C. Becker, S. Dodelson, X. Fan, J. E. Gunn, J. Hao, Ž. Ivezić, S. Jester, L. Jiang, D. E. Johnston, J. M. Kubo, H. Lampeitl, H. Lin, R. H. Lupton, G. Miknaitis, H.-J. Seo, M. Simet, and B. Yanny. The Sloan Digital Sky Survey Coadd: 275 deg^2 of Deep Sloan Digital Sky Survey Imaging on Stripe 82. *Astrophysical Journal*, 794:120, October 2014. DOI: 10.1088/0004-637X/794/2/120. URL http://adsabs.harvard.edu/abs/2014ApJ...794..120A.

E. A. Baltz, P. Marshall, and M. Oguri. Analytic models of plausible gravitational lens potentials. *Journal of Cosmology and Astroparticle Physics*, 1:015, January 2009. DOI: 10.1088/1475-7516/2009/01/015. URL http://adsabs.harvard.edu/abs/2009JCAP...01..015B.

J. M. Bardeen. Gauge-invariant cosmological perturbations. *Physical Review D*, 22:1882–1905, October 1980. DOI: 10.1103/PhysRevD.22.1882. URL http://adsabs.harvard.edu/abs/1980PhRvD..22.1882B.

M. Bartelmann and P. Schneider. Weak gravitational lensing. *Physics Report*, 340:291–472, January 2001. DOI:

10.1016/S0370-1573(00)00082-X. URL http://adsabs.
harvard.edu/abs/2001PhR...340..291B.

W. A. Baum. Population Inferences from Star Counts, Surface
Brightness and Colors. *Publications of the Astronomical
Society of the Pacific*, 71(419):106–117, Apr 1959. DOI:
10.1086/127346. URL https://ui.adsabs.harvard.edu/
abs/1959PASP...71..106B.

D. Baumann. TASI Lectures on Inflation. *ArXiv e-prints*,
July 2009. URL http://adsabs.harvard.edu/abs/
2009arXiv0907.5424B.

Gordon Bell, Jim Gray, and Alex Szalay. Petascale computational
systems. *Computer*, 39(1):110–112, January 2006. ISSN 0018-
9162. DOI: 10.1109/MC.2006.29. URL https://doi.org/
10.1109/MC.2006.29.

N. Benítez. Bayesian Photometric Redshift Estimation.
Astrophysical Journal, 536:571–583, June 2000. DOI:
10.1086/308947. URL http://adsabs.harvard.edu/abs/
2000ApJ...536..571B.

N. Benítez, R. Dupke, M. Moles, L. Sodré, A. J. Cenarro,
A. Marín Franch, K. Taylor, D. Cristóbal, A. Fernández-
Soto, C. Mendes de Oliveira, J. Cepa-Nogué, L. R. Abramo,
J. S. Alcaniz, R. Overzier, C. Hernández-Monteagudo,
E. J. Alfaro, A. Kanaan, M. Carvano, R. R. R. Reis, and J-
PAS collaboration. J-PAS: The Javalambre-Physics of the
Accelerated Universe Astrophysical Survey. In A. J. Cenarro,
F. Figueras, C. Hernández-Monteagudo, J. Trujillo Bueno, and
L. Valdivielso, editors, *Highlights of Spanish Astrophysics VIII*,
pages 148–153, May 2015. URL http://adsabs.harvard.
edu/abs/2015hsa8.conf..148B.

Leandro Beraldo e Silva, Walter de Siqueira Pedra, Monica Valluri,
Laerte Sodré, and Jean-Bernard Bru. The Discreteness-driven
Relaxation of Collisionless Gravitating Systems: Entropy
Evolution in External Potentials, N-dependence, and the Role
of Chaos. *Astrophysical Journal*, 870(2):128, Jan 2019. DOI:
10.3847/1538-4357/aaf397.

Leandro J. Beraldo e Silva, Marcos Lima, and Laerte Sodré. Testing phenomenological and theoretical models of dark matter density profiles with galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 436(3):2616–2624, Dec 2013. DOI: 10.1093/mnras/stt1761.

G. M. Bernstein and M. Jarvis. Shapes and Shears, Stars and Smears: Optimal Measurements for Weak Lensing. *Astrophysical Journal*, 123(2):583–618, Feb 2002. DOI: 10.1086/338085. URL https://ui.adsabs.harvard.edu/abs/2002AJ....123..583B.

Gary M. Bernstein and Robert Armstrong. Bayesian lensing shear measurement. *Monthly Notices of the Royal Astronomical Society*, 438(2):1880–1893, Feb 2014. DOI: 10.1093/mnras/stt2326. URL https://ui.adsabs.harvard.edu/abs/2014MNRAS.438.1880B.

E. Bertin. Automated Morphometry with SExtractor and PSFEx. In I. N. Evans, A. Accomazzi, D. J. Mink, and A. H. Rots, editors, *Astronomical Data Analysis Software and Systems XX*, volume 442 of *Astronomical Society of the Pacific Conference Series*, page 435, Jul 2011. URL https://ui.adsabs.harvard.edu/abs/2011ASPC..442..435B.

Richard G. Bower, J. R. Lucey, and Richard S. Ellis. Precision photometry of early-type galaxies in the Coma and Virgo clusters : a test of the universality of the colour-magnitude relation - I. The data. *Monthly Notices of the Royal Astronomical Society*, 254:589–600, Feb 1992. DOI: 10.1093/mnras/254.4.589. URL https://ui.adsabs.harvard.edu/abs/1992MNRAS.254..589B.

Sarah Bridle, Sreekumar T. Balan, Matthias Bethge, Marc Gentile, Stefan Harmeling, Catherine Heymans, Michael Hirsch, Reshad Hosseini, Mike Jarvis, Donnacha Kirk, Thomas Kitching, Konrad Kuijken, Antony Lewis, Stephane Paulin-Henriksson, Bernhard Schölkopf, Malin Velander, Lisa Voigt, Dugan Witherick, Adam Amara, Gary Bernstein, Frédéric Courbin, Mandeep Gill, Alan Heavens, Rachel Mandelbaum, Richard Massey, Baback Moghaddam, Anais Rassat, Alexandre Réfrégier,

Jason Rhodes, Tim Schrabback, John Shawe-Taylor, Marina Shmakova, Ludovic van Waerbeke, and David Wittman. Results of the GREAT08 Challenge: an image analysis competition for cosmological lensing. *Monthly Notices of the Royal Astronomical Society*, 405(3):2044–2061, Jul 2010. DOI: 10.1111/j.1365-2966.2010.16598.x. URL https://ui.adsabs.harvard.edu/abs/2010MNRAS.405.2044B.

T. Buchert, M. Carfora, G. F. R. Ellis, E. W. Kolb, M. A. H. MacCallum, J. J. Ostrowski, S. Räsänen, B. F. Roukema, L. Andersson, A. A. Coley, and D. L. Wiltshire. Is there proof that backreaction of inhomogeneities is irrelevant in cosmology? *Classical and Quantum Gravity*, 32(21):215021, Nov 2015. DOI: 10.1088/0264-9381/32/21/215021. URL https://ui.adsabs.harvard.edu/abs/2015CQGra..32u5021B.

J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. *Monthly Notices of the Royal Astronomical Society*, 321(3):559–575, Mar 2001. DOI: 10.1046/j.1365-8711.2001.04068.x. URL https://ui.adsabs.harvard.edu/abs/2001MNRAS.321..559B.

K. Bundy, A. Leauthaud, S. Saito, A. Bolton, Y.-T. Lin, C. Maraston, R. C. Nichol, D. P. Schneider, D. Thomas, and D. A. Wake. The Stripe 82 Massive Galaxy Project. I. Catalog Construction. *Astrophysical Journal*, 221:15, November 2015. DOI: 10.1088/0067-0049/221/1/15. URL http://adsabs.harvard.edu/abs/2015ApJS..221...15B.

H. Butcher and Jr. Oemler, A. The evolution of galaxies in clusters. I. ISIT photometry of Cl 0024+1654 and 3C 295. *Astrophysical Journal*, 219:18–30, Jan 1978. DOI: 10.1086/155751. URL https://ui.adsabs.harvard.edu/abs/1978ApJ...219...18B.

D. Carter and N. Metcalfe. The morphology of clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 191:325–337, May 1980. URL http://adsabs.harvard.edu/abs/1980MNRAS.191..325C.

S. Chandrasekhar. Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction. *Astrophysical Journal*, 97:255, March 1943. DOI: 10.1086/144517. URL http://adsabs.harvard.edu/abs/1943ApJ....97..255C.

C. Chang, M. Jarvis, B. Jain, S. M. Kahn, D. Kirkby, A. Connolly, S. Krughoff, E. H. Peng, and J. R. Peterson. The effective number density of galaxies for weak lensing measurements in the LSST project. *Monthly Notices of the Royal Astronomical Society*, 434(3):2121–2135, Sep 2013. DOI: 10.1093/mnras/stt1156.

François Chollet et al. Keras. https://keras.io, 2015.

O. Chwolson. Über eine mögliche Form fiktiver Doppelsterne. *Astronomische Nachrichten*, 221:329, June 1924. URL http://adsabs.harvard.edu/abs/1924AN....221..329C.

D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky. A Direct Empirical Proof of the Existence of Dark Matter. *Astrophysical Journal Letters*, 648:L109–L113, September 2006. DOI: 10.1086/508162. URL http://adsabs.harvard.edu/abs/2006ApJ...648L.109C.

S. Cole and C. Lacey. The structure of dark matter haloes in hierarchical clustering models. *Monthly Notices of the Royal Astronomical Society*, 281:716, July 1996. DOI: 10.1093/mnras/281.2.716. URL http://adsabs.harvard.edu/abs/1996MNRAS.281..716C.

M. Colless, G. Dalton, S. Maddox, W. Sutherland, P. Norberg, S. Cole, J. Bland-Hawthorn, T. Bridges, R. Cannon, C. Collins, W. Couch, N. Cross, K. Deeley, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, D. Madgwick, J. A. Peacock, B. A. Peterson, I. Price, M. Seaborne, and K. Taylor. The 2dF Galaxy Redshift Survey: spectra and redshifts. *Monthly Notices of the Royal Astronomical Society*, 328:1039–1063, December 2001. DOI: 10.1046/j.1365-8711.2001.04902.x. URL http://adsabs.harvard.edu/abs/2001MNRAS.328.1039C.

Julia M. Comerford, Leonidas A. Moustakas, and Priyamvada Natarajan. Observed Scaling Relations for Strong Lensing Clusters: Consequences for Cosmology and Cluster Assembly. *Astrophysical Journal*, 715(1):162–171, May 2010. DOI: 10.1088/0004-637X/715/1/162. URL https://ui.adsabs.harvard.edu/abs/2010ApJ...715..162C.

A. Cooray and R. Sheth. Halo models of large scale structure. *Physics Reports*, 372:1–129, December 2002. DOI: 10.1016/S0370-1573(02)00276-4. URL http://adsabs.harvard.edu/abs/2002PhR...372....1C.

D. J. Croton, V. Springel, S. D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida. The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *Monthly Notices of the Royal Astronomical Society*, 365:11–28, January 2006. DOI: 10.1111/j.1365-2966.2005.09675.x. URL http://adsabs.harvard.edu/abs/2006MNRAS.365..11C.

E. S. Cypriano, C. L. Mendes de Oliveira, and L. Sodré, Jr. Velocity Dispersion, Mass, and the Luminosity Function of the Fossil Cluster RX J1416.4+2315. *Astronomical Journal*, 132:514–520, August 2006. DOI: 10.1086/505044. URL http://adsabs.harvard.edu/abs/2006AJ....132..514C.

Gavin Dalton, Scott Trager, Don Carlos Abrams, Piercarlo Bonifacio, J. Alfonso López Aguerri, Kevin Middleton, Chris Benn, Kevin Dee, Frédéric Sayède, Ian Lewis, Johan Pragt, Sergio Pico, Nic Walton, Juerg Rey, Carlos Allende Prieto, José Peñate, Emilie Lhome, Tibor Agócs, José Alonso, David Terrett, Matthew Brock, James Gilbert, Andy Ridings, Isabelle Guinouard, Marc Verheijen, Ian Tosh, Kevin Rogers, Iain Steele, Remko Stuik, Neils Tromp, Attila Jasko, Jan Kragt, Dirk Lesman, Chris Mottram, Stuart Bates, Frank Gribbin, Luis Fernand o Rodriguez, José M. Delgado, Carlos Martin, Diego Cano, Ramón Navarro, Mike Irwin, Jim Lewis, Eduardo Gonzalez Solares, Neil O'Mahony, Andrea Bianco, Christina Zurita, Rik ter Horst, Emilio Molinari, Marcello Lodi, José Guerra, Antonella Vallenari, and Andrea Baruffolo. Project

overview and update on WEAVE: the next generation wide-field spectroscopy facility for the William Herschel Telescope. In *Proceedings of the SPIE*, volume 9147 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 91470L, Jul 2014. DOI: 10.1117/12.2055132. URL https://ui.adsabs.harvard.edu/abs/2014SPIE.9147E..0LD.

A. Dariush, H. G. Khosroshahi, T. J. Ponman, F. Pearce, S. Raychaudhury, and W. Hartley. The mass assembly of fossil groups of galaxies in the Millennium simulation. *Monthly Notices of the Royal Astronomical Society*, 382:433–442, November 2007. DOI: 10.1111/j.1365-2966.2007.12385.x. URL http://adsabs.harvard.edu/abs/2007MNRAS.382..433D.

A. A. Dariush, S. Raychaudhury, T. J. Ponman, H. G. Khosroshahi, A. J. Benson, R. G. Bower, and F. Pearce. The mass assembly of galaxy groups and the evolution of the magnitude gap. *Monthly Notices of the Royal Astronomical Society*, 405:1873–1887, July 2010. DOI: 10.1111/j.1365-2966.2010.16569.x. URL http://adsabs.harvard.edu/abs/2010MNRAS.405.1873D.

Dark Energy Survey Collaboration. Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing. *Physical Review D*, 98(4):043526, Aug 2018. DOI: 10.1103/PhysRevD.98.043526. URL https://ui.adsabs.harvard.edu/abs/2018PhRvD..98d3526A.

A. J. Deason, C. Conroy, A. R. Wetzel, and J. L. Tinker. Stellar Mass-gap as a Probe of Halo Assembly History and Concentration: Youth Hidden among Old Fossils. *Astrophysical Journal*, 777:154, November 2013. DOI: 10.1088/0004-637X/777/2/154. URL http://adsabs.harvard.edu/abs/2013ApJ...777..154D.

J. P. Dietrich, A. Böhnert, M. Lombardi, S. Hilbert, and J. Hartlap. The origin of peak-offsets in weak-lensing maps. *Monthly Notices of the Royal Astronomical Society*, 419(4):3547–3552, Feb 2012. DOI: 10.1111/j.1365-

2966.2011.19995.x. URL https://ui.adsabs.harvard.edu/abs/2012MNRAS.419.3547D.

S. Dodelson and G. Efstathiou. Modern Cosmology. *Physics Today*, 57:60–61, July 2004. DOI: 10.1063/1.1784308. URL http://adsabs.harvard.edu/abs/2004PhT....57R..60D.

E. D'Onghia and G. Lake. Cold Dark Matter's Small-Scale Crisis Grows Up. *Astrophysical Journal*, 612:628–632, September 2004. DOI: 10.1086/422794. URL http://adsabs.harvard.edu/abs/2004ApJ...612..628D.

E. D'Onghia, J. Sommer-Larsen, A. D. Romeo, A. Burkert, K. Pedersen, L. Portinari, and J. Rasmussen. The Formation of Fossil Galaxy Groups in the Hierarchical Universe. *Astrophysical Journal Letters*, 630:L109–L112, September 2005. DOI: 10.1086/491651. URL http://adsabs.harvard.edu/abs/2005ApJ...630L.109D.

A. Dressler. The Evolution of Galaxies in Clusters. *Annual Review of Astronomy and Astrophysics*, 22:185–222, 1984. DOI: 10.1146/annurev.aa.22.090184.001153. URL http://adsabs.harvard.edu/abs/1984ARA%26A..22..185D.

J. Dubinski. The Origin of the Brightest Cluster Galaxies. *Astrophysical Journal*, 502:141–149, July 1998. DOI: 10.1086/305901. URL http://adsabs.harvard.edu/abs/1998ApJ...502..141D.

J. Dubinski and R. G. Carlberg. The structure of cold dark matter halos. *Astrophysical Journal*, 378:496–503, September 1991. DOI: 10.1086/170451. URL http://adsabs.harvard.edu/abs/1991ApJ...378..496D.

A. R. Duffy, J. Schaye, S. T. Kay, and C. Dalla Vecchia. Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology. *Monthly Notices of the Royal Astronomical Society*, 390:L64–L68, October 2008. DOI: 10.1111/j.1745-3933.2008.00537.x. URL http://adsabs.harvard.edu/abs/2008MNRAS.390L..64D.

F. Durret, C. Adami, E. Bertin, J. Hao, I. Márquez, N. Martinet, S. Maurogordato, T. Sauvaget, N. Scepi, and M. P. Ulmer. Galaxy clusters in the SDSS Stripe 82. In J. Ballet, F. Martins, F. Bournaud, R. Monier, and C. Reylé, editors, *SF2A-2014: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, pages 315–319, December 2014. URL http://adsabs.harvard.edu/abs/2014sf2a.conf. .315D.

Cora Dvorkin, Martina Gerbino, David Alonso, Nicholas Battaglia, Simeon Bird, Ana Diaz Rivero, Andreu Font-Ribera, George Fuller, Massimiliano Lattanzi, Marilena Loverde, Julian B. Muñoz, Blake Sherwin, Anze Slosar, and Francisco Villaescusa-Navarro. Neutrino Mass from Cosmology: Probing Physics Beyond the Standard Model. *Bulletin of the American Astronomical Society*, 51(3):64, May 2019. URL https://ui.adsabs.harvard.edu/abs/2019BAAS...51c..64D.

F. W. Dyson, A. S. Eddington, and C. Davidson. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Royal Society of London Philosophical Transactions Series A*, 220:291–333, 1920. DOI: 10.1098/rsta.1920.0009. URL http://adsabs.harvard.edu/abs/1920RSPTA.220..291D.

A. S. Eddington. *Space, time and gravitation. an outline of the general relativity theory.* 1920. URL http://adsabs.harvard.edu/abs/1920stga.book.....E.

J. Ehlers, P. Geren, and R. K. Sachs. Isotropic solutions of the Einstein-Liouville equations. *Journal of Mathematical Physics*, 9:1344–1349, 1968. DOI: 10.1063/1.1664720. URL http://adsabs.harvard.edu/abs/1968JMP.....9.1344E.

Jürgen Ehlers. Contributions to the relativistic mechanics of continuous media (transl. orig. 1961). *General Relativity and Gravitation*, 25(12):1225–1266, Dec 1993. ISSN 1572-9532. DOI: 10.1007/BF00759031. URL https://doi.org/10.1007/BF00759031.

A. Einstein. Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field. *Science*, 84:506–507, December 1936. DOI: 10.1126/science.84.2188.506. URL http://adsabs.harvard.edu/abs/1936Sci....84..506E.

G. F. R. Ellis. Relativistic Cosmology. In R. K. Sachs, editor, *General Relativity and Cosmology*, pages 104–182, 1971. URL http://adsabs.harvard.edu/abs/1971grc..conf..104E.

G. F. R. Ellis. On the Raychaudhuri equation. *Pramana*, 69:15, July 2007. DOI: 10.1007/s12043-007-0107-4. URL http://adsabs.harvard.edu/abs/2007Prama..69...15E.

G. F. R. Ellis, R. Maartens, and M. A. H. MacCallum. *Relativistic Cosmology*. March 2012. URL http://adsabs.harvard.edu/abs/2012reco.book.....E.

T. Erben, L. Van Waerbeke, E. Bertin, Y. Mellier, and P. Schneider. How accurately can we measure weak gravitational shear? *Astronomy and Astrophysics*, 366:717–735, Feb 2001. DOI: 10.1051/0004-6361:20010013. URL https://ui.adsabs.harvard.edu/abs/2001A&A...366..717E.

T. Erben, H. Hildebrandt, L. Miller, L. van Waerbeke, C. Heymans, H. Hoekstra, T. D. Kitching, Y. Mellier, J. Benjamin, C. Blake, C. Bonnett, O. Cordes, J. Coupon, L. Fu, R. Gavazzi, B. Gillis, E. Grocutt, S. D. J. Gwyn, K. Holhjem, M. J. Hudson, M. Kilbinger, K. Kuijken, M. Milkeraitis, B. T. P. Rowe, T. Schrabback, E. Semboloni, P. Simon, M. Smit, O. Toader, S. Vafaei, E. van Uitert, and M. Velander. CFHTLenS: the Canada-France-Hawaii Telescope Lensing Survey - imaging data and catalogue products. *Monthly Notices of the Royal Astronomical Society*, 433:2545–2563, August 2013. DOI: 10.1093/mnras/stt928. URL http://adsabs.harvard.edu/abs/2013MNRAS.433.2545E.

C. W. F. Everitt, D. B. Debra, B. W. Parkinson, J. P. Turneaure, J. W. Conklin, M. I. Heifetz, G. M. Keiser, A. S. Silbergleit, T. Holmes, J. Kolodziejczak, M. Al-Meshari, J. C. Mester, B. Muhlfelder, V. G. Solomonik, K. Stahl, P. W. Worden, Jr., W. Bencze, S. Buchman, B. Clarke, A. Al-Jadaan, H. Al-Jibreen,

J. Li, J. A. Lipa, J. M. Lockhart, B. Al-Suwaidan, M. Taber, and S. Wang. Gravity Probe B: Final Results of a Space Experiment to Test General Relativity. *Physical Review Letters*, 106(22): 221101, June 2011. DOI: 10.1103/PhysRevLett.106.221101. URL http://adsabs.harvard.edu/abs/2011PhRvL. 106v1101E.

A. C. Fabian. Cooling Flows in Clusters of Galaxies. *Annual Review of Astronomy and Astrophysics*, 32:277–318, 1994. DOI: 10.1146/annurev.aa.32.090194.001425. URL http: //adsabs.harvard.edu/abs/1994ARA%26A..32..277F.

V. Fock. *The Theory of Space, Time and Gravitation*. Elsevier Science, 1955/2015. ISBN 9781483184906. URL https: //books.google.com.br/books?id=X7A3BQAAQBAJ.

J. Ford, L. Van Waerbeke, M. Milkeraitis, C. Laigle, H. Hildebrandt, T. Erben, C. Heymans, H. Hoekstra, T. Kitching, Y. Mellier, L. Miller, A. Choi, J. Coupon, L. Fu, M. J. Hudson, K. Kuijken, N. Robertson, B. Rowe, T. Schrabback, and M. Velander. CFHTLenS: a weak lensing shear analysis of the 3D-Matched-Filter galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 447:1304–1318, February 2015. DOI: 10.1093/mnras/stu2545. URL http://adsabs.harvard.edu/abs/2015MNRAS.447. 1304F.

D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125:306–312, March 2013. DOI: 10.1086/670067. URL http://adsabs.harvard.edu/ abs/2013PASP..125..306F.

James Foster and J. David Nightingale. *A Short Course in General Relativity*. 2006. DOI: 10.1007/978-0-387-27583-3. URL https://ui.adsabs.harvard.edu/abs/2006asci.book. ....F.

Wendy L. Freedman, Barry F. Madore, Brad K. Gibson, Laura Ferrarese, Daniel D. Kelson, Shoko Sakai, Jeremy R. Mould, Jr. Kennicutt, Robert C., Holland C. Ford, John A. Graham, John P. Huchra, Shaun M. G. Hughes, Garth D. Illingworth,

Lucas M. Macri, and Peter B. Stetson. Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant. *Astrophysical Journal*, 553(1):47–72, May 2001. DOI: 10.1086/320638.

A. Friedmann. Über die Krümmung des Raumes. *Zeitschrift fur Physik*, 10:377–386, Jan 1922. DOI: 10.1007/BF01332580. URL https://ui.adsabs.harvard.edu/abs/1922ZPhy...10..377F.

Kunihiko Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202, Apr 1980. ISSN 1432-0770. DOI: 10.1007/BF00344251. URL https://doi.org/10.1007/BF00344251.

M. Girardi, E. Escalera, D. Fadda, G. Giuricin, F. Mardirossian, and M. Mezzetti. Optical Substructures in 48 Galaxy Clusters: New Insights from a Multiscale Analysis. *Astrophysical Journal*, 482:41–62, June 1997. URL http://adsabs.harvard.edu/abs/1997ApJ...482...41G.

M. Girardi, J. A. L. Aguerri, S. De Grandi, E. D'Onghia, R. Barrena, W. Boschin, J. Méndez-Abreu, R. Sánchez-Janssen, S. Zarattini, A. Biviano, N. Castro-Rodriguez, E. M. Corsini, C. del Burgo, J. Iglesias-Páramo, and J. M. Vilchez. Fossil group origins. III. The relation between optical and X-ray luminosities. *Astronomy & Astrophysics*, 565:A115, May 2014. DOI: 10.1051/0004-6361/201323311. URL http://adsabs.harvard.edu/abs/2014A%26A...565A.115G.

Michael D. Gladders, Omar López-Cruz, H. K. C. Yee, and Tadayuki Kodama. The Slope of the Cluster Elliptical Red Sequence: A Probe of Cluster Evolution. *Astrophysical Journal*, 501(2):571–577, Jul 1998. DOI: 10.1086/305858. URL https://ui.adsabs.harvard.edu/abs/1998ApJ...501..571G.

III Gott, J. R. Recent theories of galaxy formation. *Annual Review of Astronomy and Astrophysics*, 15:235–266, Jan 1977.

DOI: 10.1146/annurev.aa.15.090177.001315. URL https://ui.adsabs.harvard.edu/abs/1977ARA&A..15..235G.

G. Gozaliasl, A. Finoguenov, H. G. Khosroshahi, M. Mirkazemi, M. Salvato, D. M. Z. Jassur, G. Erfanianfar, P. Popesso, M. Tanaka, M. Lerchster, J. P. Kneib, H. J. McCracken, Y. Mellier, E. Egami, M. J. Pereira, F. Brimioulle, T. Erben, and S. Seitz. Mining the gap: evolution of the magnitude gap in X-ray galaxy groups from the 3-square-degree XMM coverage of CFHTLS. *Astronomy and Astrophysics*, 566:A140, June 2014. DOI: 10.1051/0004-6361/201322459. URL http://adsabs.harvard.edu/abs/2014A%26A...566A.140G.

Stephen R. Green and Robert M. Wald. Examples of backreaction of small-scale inhomogeneities in cosmology. *Physical Review D*, 87(12):124037, Jun 2013. DOI: 10.1103/PhysRevD.87.124037. URL https://ui.adsabs.harvard.edu/abs/2013PhRvD..8714037G.

J. E. Gunn and J. R. Gott, III. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *Astrophysical Journal*, 176:1, August 1972. DOI: 10.1086/151605. URL http://adsabs.harvard.edu/abs/1972ApJ...176....1G.

N. Hand, A. Leauthaud, S. Das, B. D. Sherwin, G. E. Addison, J. R. Bond, E. Calabrese, A. Charbonnier, M. J. Devlin, J. Dunkley, T. Erben, A. Hajian, M. Halpern, J. Harnois-Déraps, C. Heymans, H. Hildebrandt, A. D. Hincks, J.-P. Kneib, A. Kosowsky, M. Makler, L. Miller, K. Moodley, B. Moraes, M. D. Niemack, L. A. Page, B. Partridge, N. Sehgal, H. Shan, J. L. Sievers, D. N. Spergel, S. T. Staggs, E. R. Switzer, J. E. Taylor, L. Van Waerbeke, C. Welker, and E. J. Wollack. First measurement of the cross-correlation of CMB lensing and galaxy lensing. *Physical Review D*, 91(6):062001, March 2015. DOI: 10.1103/PhysRevD.91.062001. URL http://adsabs.harvard.edu/abs/2015PhRvD..91f2001H.

Jiangang Hao, Timothy A. McKay, Benjamin P. Koester, Eli S. Rykoff, Eduardo Rozo, James Annis, Risa H. Wechsler, August Evrard, Seth R. Siegel, Matthew Becker, Michael Busha, David

Gerdes, David E. Johnston, and Erin Sheldon. A GMBCG Galaxy Cluster Catalog of 55,424 Rich Clusters from SDSS DR7. *Astrophysical Journal Suplement*, 191(2):254–274, Dec 2010. DOI: 10.1088/0067-0049/191/2/254. URL https://ui.adsabs.harvard.edu/abs/2010ApJS..191..254H.

John W. Hardy. *Adaptive Optics for Astronomical Telescopes*. 1998. URL https://ui.adsabs.harvard.edu/abs/1998aoat.book.....H.

C. D. Harrison, C. J. Miller, J. W. Richards, E. J. Lloyd-Davies, B. Hoyle, A. K. Romer, N. Mehrtens, M. Hilton, J. P. Stott, D. Capozzi, C. A. Collins, P.-J. Deadman, A. R. Liddle, M. Sahlén, S. A. Stanford, and P. T. P. Viana. The XMM Cluster Survey: The Stellar Mass Assembly of Fossil Galaxies. *Astrophysical Journal*, 752:12, June 2012. DOI: 10.1088/0004-637X/752/1/12. URL http://adsabs.harvard.edu/abs/2012ApJ...752...12H.

J. Hartlap, P. Simon, and P. Schneider. Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix. *Astronomy & Astrophysics*, 464:399–404, March 2007. DOI: 10.1051/0004-6361:20066170. URL http://adsabs.harvard.edu/abs/2007A%26A...464..399H.

D. Harvey, R. Massey, T. Kitching, A. Taylor, and E. Tittley. The nongravitational interactions of dark matter in colliding galaxy clusters. *Science*, 347:1462–1465, March 2015. DOI: 10.1126/science.1261381. URL http://adsabs.harvard.edu/abs/2015Sci...347.1462H.

S. W. Hawking and G. F. R. Ellis. *The Large-scale Structure of Space-time*. 1973. URL http://adsabs.harvard.edu/abs/1973lsss.book.....H.

S. W. Hawking and R. Penrose. The Singularities of Gravitational Collapse and Cosmology. *Royal Society of London Proceedings Series A*, 314:529–548, January 1970. DOI: 10.1098/rspa.1970.0021. URL http://adsabs.harvard.edu/abs/1970RSPSA.314..529H.

A. P. Hearin, A. R. Zentner, J. A. Newman, and A. A. Berlind. Mind the gap: tightening the mass-richness relation with magnitude gaps. *Monthly Notices of the Royal Astronomical Society*, 430:1238–1246, April 2013. DOI: 10.1093/mnras/sts699. URL http://adsabs.harvard.edu/abs/2013MNRAS.430.1238H.

Alan Heavens, Alexandre Refregier, and Catherine Heymans. Intrinsic correlation of galaxy shapes: implications for weak lensing measurements. *Monthly Notices of the Royal Astronomical Society*, 319(2):649–656, Dec 2000. DOI: 10.1046/j.1365-8711.2000.03907.x. URL https://ui.adsabs.harvard.edu/abs/2000MNRAS.319..649H.

Catherine Heymans, Ludovic Van Waerbeke, David Bacon, Joel Berge, Gary Bernstein, Emmanuel Bertin, Sarah Bridle, Michael L. Brown, Douglas Clowe, Håkon Dahle, Thomas Erben, Meghan Gray, Marco Hetterscheidt, Henk Hoekstra, Patrick Hudelot, Mike Jarvis, Konrad Kuijken, Vera Margoniner, Richard Massey, Yannick Mellier, Reiko Nakajima, Alexandre Refregier, Jason Rhodes, Tim Schrabback, and David Wittman. The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations. *Monthly Notices of the Royal Astronomical Society*, 368(3):1323–1339, May 2006. DOI: 10.1111/j.1365-2966.2006.10198.x. URL https://ui.adsabs.harvard.edu/abs/2006MNRAS.368.1323H.

H. Hildebrandt, M. Viola, C. Heymans, S. Joudaki, K. Kuijken, C. Blake, T. Erben, B. Joachimi, D. Klaes, L. Miller, C. B. Morrison, R. Nakajima, G. Verdoes Kleijn, A. Amon, A. Choi, G. Covone, J. T. A. de Jong, A. Dvornik, I. Fenech Conti, A. Grado, J. Harnois-Déraps, R. Herbonnet, H. Hoekstra, F. Köhlinger, J. McFarland, A. Mead, J. Merten, N. Napolitano, J. A. Peacock, M. Radovich, P. Schneider, P. Simon, E. A. Valentijn, J. L. van den Busch, E. van Uitert, and L. Van Waerbeke. KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 465(2):1454–1498,

Feb 2017. DOI: 10.1093/mnras/stw2805. URL https://ui.adsabs.harvard.edu/abs/2017MNRAS.465.1454H.

G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *Astrophysical Journal Supplements*, 208: 19, October 2013. DOI: 10.1088/0067-0049/208/2/19. URL http://adsabs.harvard.edu/abs/2013ApJS..208...19H.

Christopher Hirata and Uroš Seljak. Shear calibration biases in weak-lensing surveys. *Monthly Notices of the Royal Astronomical Society*, 343(2):459–480, Aug 2003. DOI: 10.1046/j.1365-8711.2003.06683.x. URL https://ui.adsabs.harvard.edu/abs/2003MNRAS.343..459H.

H. Hoekstra, H. K. C. Yee, and M. D. Gladders. Current status of weak gravitational lensing. *New Astronomy Review*, 46:767–781, November 2002. DOI: 10.1016/S1387-6473(02)00245-2. URL http://adsabs.harvard.edu/abs/2002NewAR..46..767H.

H. Hoekstra, M. Bartelmann, H. Dahle, H. Israel, M. Limousin, and M. Meneghetti. Masses of Galaxy Clusters from Gravitational Lensing. *Space Science Reviews*, 177:75–118, August 2013. DOI: 10.1007/s11214-013-9978-5. URL http://adsabs.harvard.edu/abs/2013SSRv..177...75H.

David W. Hogg and Dustin Lang. Replacing Standard Galaxy Profiles with Mixtures of Gaussians. *Publications of the Astronomical Society of the Pacific*, 125(928):719, Jun 2013. DOI: 10.1086/671228. URL https://ui.adsabs.harvard.edu/abs/2013PASP..125..719H.

Hanako Hoshino, Alexie Leauthaud, Claire Lackner, Chiaki Hikage, Eduardo Rozo, Eli Rykoff, Rachel Mand elbaum, Surhud More, Anupreeta More, Shun Saito, and Benedetta Vulcani. Luminous red galaxies in clusters: central occupation,

spatial distributions and miscentring. *Monthly Notices of the Royal Astronomical Society*, 452(1):998–1013, Sep 2015. DOI: 10.1093/mnras/stv1271. URL https://ui.adsabs.harvard.edu/abs/2015MNRAS.452..998H.

E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929. DOI: 10.1073/pnas.15.3.168. URL http://adsabs.harvard.edu/abs/1929PNAS...15..168H.

Eric Huff and Rachel Mandelbaum. Metacalibration: Direct Self-Calibration of Biases in Shear Measurement. *arXiv e-prints*, art. arXiv:1702.02600, Feb 2017. URL https://ui.adsabs.harvard.edu/abs/2017arXiv170202600H.

R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *Astrophysical Journal Letters*, 195:L51–L53, Jan 1975. DOI: 10.1086/181708. URL https://ui.adsabs.harvard.edu/abs/1975ApJ...195L..51H.

A. Huss, B. Jain, and M. Steinmetz. How Universal Are the Density Profiles of Dark Halos? *Astrophysical Journal*, 517: 64–69, May 1999. DOI: 10.1086/307161. URL http://adsabs.harvard.edu/abs/1999ApJ...517...64H.

E. E. O. Ishida, R. Beck, S. González-Gaitán, R. S. de Souza, A. Krone-Martins, J. W. Barrett, N. Kennamer, R. Vilalta, J. M. Burgess, B. Quint, A. Z. Vitorelli, A. Mahabal, E. Gangler, and COIN Collaboration. Optimizing spectroscopic follow-up strategies for supernova photometric classification with active learning. *Monthly Notices of the Royal Astronomical Society*, 483(1):2–18, Feb 2019. DOI: 10.1093/mnras/sty3015. URL https://ui.adsabs.harvard.edu/abs/2019MNRAS.483...2I.

M. Jarvis, E. Sheldon, J. Zuntz, T. Kacprzak, S. L. Bridle, A. Amara, R. Armstrong, M. R. Becker, G. M. Bernstein, C. Bonnett, C. Chang, R. Das, J. P. Dietrich, A. Drlica-Wagner, T. F. Eifler, C. Gangkofner, D. Gruen, M. Hirsch, E. M. Huff, B. Jain, S. Kent, D. Kirk, N. MacCrann, P. Melchior, A. A. Plazas, A. Refregier, B. Rowe, E. S. Rykoff, S. Samuroff,

C. Sánchez, E. Suchyta, M. A. Troxel, V. Vikram, T. Abbott, F. B. Abdalla, S. Allam, J. Annis, A. Benoit-Lévy, E. Bertin, D. Brooks, E. Buckley-Geer, D. L. Burke, D. Capozzi, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, J. Clampitt, M. Crocce, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, D. L. DePoy, S. Desai, H. T. Diehl, P. Doel, A. Fausti Neto, B. Flaugher, P. Fosalba, J. Frieman, E. Gaztanaga, D. W. Gerdes, R. A. Gruendl, G. Gutierrez, K. Honscheid, D. J. James, K. Kuehn, N. Kuropatkin, O. Lahav, T. S. Li, M. Lima, M. March, P. Martini, R. Miquel, J. J. Mohr, E. Neilsen, B. Nord, R. Ogando, K. Reil, A. K. Romer, A. Roodman, M. Sako, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, M. E. C. Swanson, G. Tarle, J. Thaler, D. Thomas, A. R. Walker, and R. H. Wechsler. The DES Science Verification weak lensing shear catalogues. *Monthly Notices of the Royal Astronomical Society*, 460(2):2245–2281, Aug 2016. DOI: 10.1093/mnras/stw990. URL https://ui.adsabs.harvard.edu/abs/2016MNRAS.460.2245J.

Y. P. Jing. The Density Profile of Equilibrium and Nonequilibrium Dark Matter Halos. *Astrophysical Journal*, 535(1):30–36, May 2000. DOI: 10.1086/308809. URL https://ui.adsabs.harvard.edu/abs/2000ApJ...535...30J.

D. E. Johnston, E. S. Sheldon, R. H. Wechsler, E. Rozo, B. P. Koester, J. A. Frieman, T. A. McKay, A. E. Evrard, M. R. Becker, and J. Annis. Cross-correlation Weak Lensing of SDSS galaxy Clusters II: Cluster Density Profiles and the Mass–Richness Relation. *ArXiv e-prints*, September 2007. URL http://adsabs.harvard.edu/abs/2007arXiv0709.1159J.

L. R. Jones, T. J. Ponman, and D. A. Forbes. Multiwavelength observations of an evolved galaxy group: an end-point of galaxy merging? *Monthly Notices of the Royal Astronomical Society*, 312:139–150, February 2000. DOI: 10.1046/j.1365-8711.2000.03118.x. URL http://adsabs.harvard.edu/abs/2000MNRAS.312..139J.

L. R. Jones, T. J. Ponman, A. Horton, A. Babul, H. Ebeling, and

D. J. Burke. The nature and space density of fossil groups of galaxies. *Monthly Notices of the Royal Astronomical Society*, 343:627–638, August 2003. DOI: 10.1046/j.1365-8711.2003.06702.x. URL http://adsabs.harvard.edu/abs/2003MNRAS.343..627J.

F. Kahlhoefer, K. Schmidt-Hoberg, M. T. Frandsen, and S. Sarkar. Colliding clusters and dark matter self-interactions. *Monthly Notices of the Royal Astronomical Society*, 437:2865–2881, January 2014. DOI: 10.1093/mnras/stt2097. URL http://adsabs.harvard.edu/abs/2014MNRAS.437.2865K.

Nick Kaiser, Gordon Squires, and Tom Broadhurst. A Method for Weak Lensing Observations. *Astrophysical Journal*, 449:460, Aug 1995. DOI: 10.1086/176071. URL https://ui.adsabs.harvard.edu/abs/1995ApJ...449..460K.

N. Katz. Dissipationless collapse in an expanding universe. *Astrophysical Journal*, 368:325–336, February 1991. DOI: 10.1086/169696. URL http://adsabs.harvard.edu/abs/1991ApJ...368..325K.

H. G. Khosroshahi, L. R. Jones, and T. J. Ponman. An old galaxy group: Chandra X-ray observations of the nearby fossil group NGC 6482. *Monthly Notices of the Royal Astronomical Society*, 349:1240–1250, April 2004. DOI: 10.1111/j.1365-2966.2004.07575.x. URL http://adsabs.harvard.edu/abs/2004MNRAS.349.1240K.

H. G. Khosroshahi, T. J. Ponman, and L. R. Jones. The central elliptical galaxy in fossil groups and formation of brightest cluster galaxies. *Monthly Notices of the Royal Astronomical Society*, 372:L68–L72, October 2006. DOI: 10.1111/j.1745-3933.2006.00228.x. URL http://adsabs.harvard.edu/abs/2006MNRAS.372L..68K.

H. G. Khosroshahi, T. J. Ponman, and L. R. Jones. Scaling relations in fossil galaxy groups. *Monthly Notices of the Royal Astronomical Society*, 377:595–606, May 2007. DOI: 10.1111/j.1365-2966.2007.11591.x. URL http://adsabs.harvard.edu/abs/2007MNRAS.377..595K.

T. D. Kitching, L. Miller, C. E. Heymans, L. van Waerbeke, and A. F. Heavens. Bayesian galaxy shape measurement for weak lensing surveys - II. Application to simulations. *Monthly Notices of the Royal Astronomical Society*, 390:149–167, October 2008. DOI: 10.1111/j.1365-2966.2008.13628.x. URL http://adsabs.harvard.edu/abs/2008MNRAS.390..149K.

T. D. Kitching, S. T. Balan, S. Bridle, N. Cantale, F. Courbin, T. Eifler, M. Gentile, M. S. S. Gill, S. Harmeling, C. Heymans, M. Hirsch, K. Honscheid, T. Kacprzak, D. Kirkby, D. Margala, R. J. Massey, P. Melchior, G. Nurbaeva, K. Patton, J. Rhodes, B. T. P. Rowe, A. N. Taylor, M. Tewes, M. Viola, D. Witherick, L. Voigt, J. Young, and J. Zuntz. Image analysis for cosmology: results from the GREAT10 Galaxy Challenge. *Monthly Notices of the Royal Astronomical Society*, 423(4):3163–3208, Jul 2012. DOI: 10.1111/j.1365-2966.2012.21095.x. URL https://ui.adsabs.harvard.edu/abs/2012MNRAS.423.3163K.

S. C. Kleene. Representation of events in nerve nets and finite automata. In Claude Shannon and John McCarthy, editors, *Automata Studies*, pages 3–41. Princeton University Press, Princeton, NJ, 1956. URL https://www.bibsonomy.org/bibtex/2cc51c86a4093c3f989db80fb8d7831be/nlp.

A. A. Klypin, S. Trujillo-Gomez, and J. Primack. Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation. *Astrophysical Journal*, 740:102, October 2011. DOI: 10.1088/0004-637X/740/2/102. URL http://adsabs.harvard.edu/abs/2011ApJ...740..102K.

Benjamin P. Koester, Timothy A. McKay, James Annis, Risa H. Wechsler, August E. Evrard, Eduardo Rozo, Lindsey Bleem, Erin S. Sheldon, and David Johnston. MaxBCG: A Red-Sequence Galaxy Cluster Finder. *Astrophysical Journal*, 660 (1):221–238, May 2007. DOI: 10.1086/512092. URL https://ui.adsabs.harvard.edu/abs/2007ApJ...660..221K.

A. V. Kravtsov and S. Borgani. Formation of Galaxy Clusters. *Annual Review of Astronomy and Astrophysics*, 50:353–409, September 2012. DOI: 10.1146/annurev-astro-081811-125502.

URL http://adsabs.harvard.edu/abs/2012ARA%26A.
.50..353K.

J. Krempec-Krygier and B. Krygier. Interaction of Abell Cluster
2063 and the Group of Galaxies MKW3s. *Acta Astronomica*,
49:403–420, September 1999. URL http://adsabs.
harvard.edu/abs/1999AcA....49..403K.

A. Kundert, F. Gastaldello, E. D'Onghia, M. Girardi, J. A. L.
Aguerri, R. Barrena, E. M. Corsini, S. De Grandi, E. Jiménez-
Bailón, M. Lozada-Muñoz, J. Méndez-Abreu, R. Sánchez-
Janssen, E. Wilcots, and S. Zarattini. Fossil group origins - VI.
Global X-ray scaling relations of fossil galaxy clusters. *Monthly
Notices of the Royal Astronomical Society*, 454:161–176,
November 2015. DOI: 10.1093/mnras/stv1879. URL http:
//adsabs.harvard.edu/abs/2015MNRAS.454..161K.

A. Kundert, E. D'Onghia, and J. A. L. Aguerri. Are Fossil Groups
Early-forming Galaxy Systems? *Astrophysical Journal*, 845:45,
August 2017. DOI: 10.3847/1538-4357/aa7b88. URL http:
//adsabs.harvard.edu/abs/2017ApJ...845...45K.

M. Lachieze-Rey and J. Luminet. Cosmic topology. *Physics
Reports*, 254:135–214, March 1995. DOI: 10.1016/0370-
1573(94)00085-H. URL http://adsabs.harvard.edu/
abs/1995PhR...254..135L.

S. M. LaMassa, C. M. Urry, E. Glikman, N. Cappelluti,
A. Comastri, H. Boehringer, G. T. Richards, and S. S. Murray.
Stripe 82 X: X-ray Survey of SDSS Stripe 82. In *AAS/High
Energy Astrophysics Division*, volume 13 of *AAS/High Energy
Astrophysics Division*, page 109.05, April 2013. URL http:
//adsabs.harvard.edu/abs/2013HEAD...1310905L.

François Lanusse, Quanbin Ma, Nan Li, Thomas E. Collett,
Chun-Liang Li, Siamak Ravanbakhsh, Rachel Mandelbaum,
and Barnabás Póczos. CMU DeepLens: deep learning for
automatic image-based galaxy-galaxy strong lens finding.
Monthly Notices of the Royal Astronomical Society, 473(3):
3895–3906, Jan 2018. DOI: 10.1093/mnras/stx1665. URL
https://ui.adsabs.harvard.edu/abs/2018MNRAS.473.
3895L.

R. Laureijs, J. Amiaux, S. Arduini, J. . Auguères, J. Brinchmann,
R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, and
et al. Euclid Definition Study Report. *ArXiv e-prints*,
October 2011. URL http://adsabs.harvard.edu/abs/
2011arXiv1110.3193L.

A. Leauthaud, S. Saito, S. Hilbert, A. Barreira, S. More,
M. White, S. Alam, P. Behroozi, K. Bundy, J. Coupon,
T. Erben, C. Heymans, H. Hildebrandt, R. Mandelbaum,
L. Miller, B. Moraes, M. E. S. Pereira, S. A. Rodríguez-Torres,
F. Schmidt, H.-Y. Shan, M. Viel, and F. Villaescusa-Navarro.
Lensing is low: cosmology, galaxy formation or new physics?
Monthly Notices of the Royal Astronomical Society, 467:3024–
3047, May 2017. DOI: 10.1093/mnras/stx258. URL http:
//adsabs.harvard.edu/abs/2017MNRAS.467.3024L.

Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard,
W. Hubbard, and L. D. Jackel. Backpropagation applied
to handwritten zip code recognition. *Neural Comput.*, 1
(4):541–551, December 1989. ISSN 0899-7667. DOI:
10.1162/neco.1989.1.4.541. URL http://dx.doi.org/10.
1162/neco.1989.1.4.541.

G. Lemaître. Un Univers homogène de masse constante
et de rayon croissant rendant compte de la vitesse radiale
des nébuleuses extra-galactiques. *Annales de la
Société Scientifique de Bruxelles*, 47:49–59, Jan
1927. URL https://ui.adsabs.harvard.edu/abs/
1927ASSB...47...49L.

Julien Lesgourgues. The Cosmic Linear Anisotropy Solving
System (CLASS) I: Overview. *arXiv e-prints*, art.
arXiv:1104.2932, Apr 2011. URL https://ui.adsabs.
harvard.edu/abs/2011arXiv1104.2932L.

A. Lewis and A. Challinor. Evolution of cosmological dark matter
perturbations. *Physical Review D*, 66(2):023531, July 2002.
DOI: 10.1103/PhysRevD.66.023531. URL http://adsabs.
harvard.edu/abs/2002PhRvD..66b3531L.

R. Li, H. Shan, H. Mo, J.-P. Kneib, X. Yang, W. Luo, F. C.
van den Bosch, T. Erben, B. Moraes, and M. Makler. First

galaxy-galaxy lensing measurement of satellite halo mass in the CFHT Stripe-82 Survey. *Monthly Notices of the Royal Astronomical Society*, 438:2864–2870, March 2014. DOI: 10.1093/mnras/stt2395. URL http://adsabs.harvard.edu/abs/2014MNRAS.438.2864L.

R. Li, H. Shan, J.-P. Kneib, H. Mo, E. Rozo, A. Leauthaud, J. Moustakas, L. Xie, T. Erben, L. Van Waerbeke, M. Makler, E. Rykoff, and B. Moraes. Measuring subhalo mass in redMaPPer clusters with CFHT Stripe 82 Survey. *Monthly Notices of the Royal Astronomical Society*, 458:2573–2583, May 2016. DOI: 10.1093/mnras/stw494. URL http://adsabs.harvard.edu/abs/2016MNRAS.458.2573L.

LIGO Scientific Collaboration and Virgo Collaboration. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6):061102, Feb 2016. DOI: 10.1103/PhysRevLett.116.061102. URL https://ui.adsabs.harvard.edu/abs/2016PhRvL.116f1102A.

M. Lima and W. Hu. Self-calibration of cluster dark energy studies: Observable-mass distribution. *Physical Review D*, 72 (4):043006, August 2005. DOI: 10.1103/PhysRevD.72.043006. URL http://adsabs.harvard.edu/abs/2005PhRvD..72d3006L.

G.B. Lima Neto. *Astronomia Extragaláctica*. 2014.

Yen-Ting Lin and Joseph J. Mohr. K-band Properties of Galaxy Clusters and Groups: Brightest Cluster Galaxies and Intracluster Light. *Astrophysical Journal*, 617(2):879–895, Dec 2004. DOI: 10.1086/425412. URL https://ui.adsabs.harvard.edu/abs/2004ApJ...617..879L.

X. Liu, C. Pan, R. Li, H. Shan, Q. Wang, L. Fu, Z. Fan, J.-P. Kneib, A. Leauthaud, L. Van Waerbeke, M. Makler, B. Moraes, T. Erben, and A. Charbonnier. Cosmological constraints from weak lensing peak statistics with Canada-France-Hawaii Telescope Stripe 82 Survey. *Monthly Notices of the Royal Astronomical Society*, 450:2888–2902, July 2015. DOI: 10.1093/mnras/stv784. URL http://adsabs.harvard.edu/abs/2015MNRAS.450.2888L.

Arthur Loureiro, Andrei Cuceu, Filipe B. Abdalla, Bruno Moraes, Lorne Whiteway, Michael McLeod, Sreekumar T. Balan, Ofer Lahav, Aurélien Benoit-Lévy, Marc Manera, Richard P. Rollins, and Henrique S. Xavier. Upper Bound of Neutrino Masses from Combined Cosmological Observations and Particle Physics Experiments. *Physical Review Letters*, 123(8):081301, Aug 2019. DOI: 10.1103/PhysRevLett.123.081301. URL https://ui.adsabs.harvard.edu/abs/2019PhRvL.123h1301L.

David Lovelock. The Einstein Tensor and Its Generalizations. *Journal of Mathematical Physics*, 12(3):498–501, Mar 1971. DOI: 10.1063/1.1665613. URL https://ui.adsabs.harvard.edu/abs/1971JMP....12..498L.

LSST Dark Energy Science Collaboration. Large Synoptic Survey Telescope: Dark Energy Science Collaboration. *arXiv e-prints*, art. arXiv:1211.0310, Nov 2012. URL https://ui.adsabs.harvard.edu/abs/2012arXiv1211.0310L.

D. Lynden-Bell. Statistical mechanics of violent relaxation in stellar systems. *Monthly Notices of the Royal Astronomical Society*, 136:101, 1967. URL http://adsabs.harvard.edu/abs/1967MNRAS.136..101L.

R. Lynds and V. Petrosian. Giant Luminous Arcs in Galaxy Clusters. In *Bulletin of the American Astronomical Society*, volume 18 of *Bulletin of the American Astronomical Society*, page 1014, September 1986. URL http://adsabs.harvard.edu/abs/1986BAAS...18R1014L.

R. Mandelbaum, U. Seljak, and C. M. Hirata. A halo mass-concentration relation from weak lensing. *Journal of Cosmology and Astroparticle Physics*, 8:006, August 2008a. DOI: 10.1088/1475-7516/2008/08/006. URL http://adsabs.harvard.edu/abs/2008JCAP...08..006M.

R. Mandelbaum, U. Seljak, C. M. Hirata, S. Bardelli, M. Bolzonella, A. Bongiorno, M. Carollo, T. Contini, C. E. Cunha, B. Garilli, A. Iovino, P. Kampczyk, J.-P. Kneib, C. Knobel, D. C. Koo, F. Lamareille, O. Le Fèvre, J.-F. Le Borgne, S. J. Lilly, C. Maier, V. Mainieri, M. Mignoli, J. A. Newman, P. A. Oesch, E. Perez-Montero, E. Ricciardelli,

M. Scodeggio, J. Silverman, and L. Tasca. Precision photometric redshift calibration for galaxy-galaxy weak lensing. *Monthly Notices of the Royal Astronomical Society*, 386:781–806, May 2008b. DOI: 10.1111/j.1365-2966.2008.12947.x. URL http://adsabs.harvard.edu/abs/2008MNRAS.386..781M.

Rachel Mandelbaum, Barnaby Rowe, Robert Armstrong, Deborah Bard, Emmanuel Bertin, James Bosch, Dominique Boutigny, Frederic Courbin, William A. Dawson, Annamaria Donnarumma, Ian Fenech Conti, Raphaël Gavazzi, Marc Gentile, Mandeep S. S. Gill, David W. Hogg, Eric M. Huff, M. James Jee, Tomasz Kacprzak, Martin Kilbinger, Thibault Kuntzer, Dustin Lang, Wentao Luo, Marisa C. March, Philip J. Marshall, Joshua E. Meyers, Lance Miller, Hironao Miyatake, Reiko Nakajima, Fred Maurice Ngolé Mboula, Guldariya Nurbaeva, Yuki Okura, Stéphane Paulin-Henriksson, Jason Rhodes, Michael D. Schneider, Huanyuan Shan, Erin S. Sheldon, Melanie Simet, Jean-Luc Starck, Florent Sureau, Malte Tewes, Kristian Zarb Adami, Jun Zhang, and Joe Zuntz. GREAT3 results - I. Systematic errors in shear estimation and the impact of real galaxy morphology. *Monthly Notices of the Royal Astronomical Society*, 450(3):2963–3007, Jul 2015. DOI: 10.1093/mnras/stv781. URL https://ui.adsabs.harvard.edu/abs/2015MNRAS.450.2963M.

Rachel Mandelbaum, Hironao Miyatake, Takashi Hamana, Masamune Oguri, Melanie Simet, Robert Armstrong, James Bosch, Ryoma Murata, François Lanusse, Alexie Leauthaud, Jean Coupon, Surhud More, Masahiro Takada, Satoshi Miyazaki, Joshua S. Speagle, Masato Shirasaki, Cristóbal Sifón, Song Huang, Atsushi J. Nishizawa, Elinor Medezinski, Yuki Okura, Nobuhiro Okabe, Nicole Czakon, Ryuichi Takahashi, William R. Coulton, Chiaki Hikage, Yutaka Komiyama, Robert H. Lupton, Michael A. Strauss, Masayuki Tanaka, and Yousuke Utsumi. The first-year shear catalog of the Subaru Hyper Suprime-Cam Subaru Strategic Program Survey. *Publications of the Astronomical Society of Japan*, 70:S25, Jan 2018. DOI: 10.1093/pasj/psx130.

Andrew W. Mann and Harald Ebeling. X-ray-optical classification

of cluster mergers and the evolution of the cluster merger fraction. *Monthly Notices of the Royal Astronomical Society*, 420(3):2120–2138, Mar 2012. DOI: 10.1111/j.1365-2966.2011.20170.x. URL https://ui.adsabs.harvard.edu/abs/2012MNRAS.420.2120M.

A. B. Mantz, S. W. Allen, R. G. Morris, R. W. Schmidt, A. von der Linden, and O. Urban. Cosmology and astrophysics from relaxed galaxy clusters - I. Sample selection. *Monthly Notices of the Royal Astronomical Society*, 449:199–219, May 2015. DOI: 10.1093/mnras/stv219. URL http://adsabs.harvard.edu/abs/2015MNRAS.449..199M.

Maxim Markevitch and Alexey Vikhlinin. Shocks and cold fronts in galaxy clusters. *Physics Reports*, 443(1):1–53, May 2007. DOI: 10.1016/j.physrep.2007.01.001. URL https://ui.adsabs.harvard.edu/abs/2007PhR...443....1M.

Richard Massey, Catherine Heymans, Joel Bergé, Gary Bernstein, Sarah Bridle, Douglas Clowe, Håkon Dahle, Richard Ellis, Thomas Erben, Marco Hetterscheidt, F. William High, Christopher Hirata, Henk Hoekstra, Patrick Hudelot, Mike Jarvis, David Johnston, Konrad Kuijken, Vera Margoniner, Rachel Mandelbaum, Yannick Mellier, Reiko Nakajima, Stephane Paulin-Henriksson, Molly Peeples, Chris Roat, Alexandre Refregier, Jason Rhodes, Tim Schrabback, Mischa Schirmer, Uroš Seljak, Elisabetta Semboloni, and Ludovic van Waerbeke. The Shear Testing Programme 2: Factors affecting high-precision weak-lensing analyses. *Monthly Notices of the Royal Astronomical Society*, 376(1):13–38, Mar 2007. DOI: 10.1111/j.1365-2966.2006.11315.x. URL https://ui.adsabs.harvard.edu/abs/2007MNRAS.376...13M.

Richard Massey, Henk Hoekstra, Thomas Kitching, Jason Rhodes, Mark Cropper, Jérôme Amiaux, David Harvey, Yannick Mellier, Massimo Meneghetti, Lance Miller, Stéphane Paulin-Henriksson, Sand rine Pires, Roberto Scaramella, and Tim Schrabback. Origins of weak lensing systematics, and requirements on future instrumentation (or knowledge of instrumentation). *Monthly Notices of the Royal Astronomical Society*, 429(1):661–678, Feb 2013. DOI:

10.1093/mnras/sts371. URL https://ui.adsabs.harvard.edu/abs/2013MNRAS.429..661M.

B. J. Maughan, C. Jones, W. Forman, and L. Van Speybroeck. Images, Structural Properties, and Metal Abundances of Galaxy Clusters Observed with Chandra ACIS-I at 0.1 < z < 1.3. *Astrophysical Journal Supplement*, 174(1):117–135, Jan 2008. DOI: 10.1086/521225. URL https://ui.adsabs.harvard.edu/abs/2008ApJS..174..117M.

J. Mayer, K. Khairy, and J. Howard. Drawing an elephant with four complex parameters. *American Journal of Physics*, 78:648–649, June 2010. DOI: 10.1119/1.3254017.

C. L. Mendes de Oliveira, E. S. Cypriano, and L. Sodré, Jr. The Luminosity Function of the Fossil Group RX J1552.2+2013. *Astronomical Journal*, 131:158–167, January 2006. DOI: 10.1086/498083. URL http://adsabs.harvard.edu/abs/2006AJ....131..158M.

C. L. Mendes de Oliveira, E. S. Cypriano, R. A. Dupke, and L. Sodré, Jr. An Optical and X-Ray Study of the Fossil Group RX J1340.6+4018. *Astronomical Journal*, 138:502-509, August 2009. DOI: 10.1088/0004-6256/138/2/502. URL http://adsabs.harvard.edu/abs/2009AJ....138..502M.

J. Méndez-Abreu, J. A. L. Aguerri, R. Barrena, R. Sánchez-Janssen, W. Boschin, N. Castro-Rodriguez, E. M. Corsini, C. Del Burgo, E. D'Onghia, M. Girardi, J. Iglesias-Páramo, N. Napolitano, J. M. Vilchez, and S. Zarattini. Fossil group origins. II. Unveiling the formation of the brightest group galaxies through their scaling relations. *Astronomy & Astrophysics*, 537:A25, January 2012. DOI: 10.1051/0004-6361/201117755. URL http://adsabs.harvard.edu/abs/2012A%26A...537A..25M.

A. Merloni, P. Predehl, W. Becker, H. Böhringer, T. Boller, H. Brunner, M. Brusa, K. Dennerl, M. Freyberg, P. Friedrich, A. Georgakakis, F. Haberl, G. Hasinger, N. Meidinger, J. Mohr, K. Nandra, A. Rau, T. H. Reiprich, J. Robrade, M. Salvato, A. Santangelo, M. Sasaki, A. Schwope, J. Wilms, and the

German eROSITA Consortium. eROSITA Science Book: Mapping the Structure of the Energetic Universe. *arXiv e-prints*, art. arXiv:1209.3114, Sep 2012. URL https://ui.adsabs.harvard.edu/abs/2012arXiv1209.3114M.

Julian Merten. A Combining Weak and Strong Lensing for Galaxy-Cluster Mass Reconstructions. Master's thesis, University of Heidelberg, 2008.

Julian Merten. *An advanced method to recover the mass distribution of galaxy clusters*. PhD thesis, University of Heidelberg, 2010.

P. Meszaros. The behaviour of point masses in an expanding cosmological substratum. *Astronomy & Astrophysics*, 37:225–228, December 1974. URL http://adsabs.harvard.edu/abs/1974A%26A....37..225M.

G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, and J. Wambsganss, editors. *Gravitational Lensing: Strong, Weak and Micro*, 2006. URL http://adsabs.harvard.edu/abs/2006glsw.conf.....M.

L. Miller, T. D. Kitching, C. Heymans, A. F. Heavens, and L. van Waerbeke. Bayesian galaxy shape measurement for weak lensing surveys - I. Methodology and a fast-fitting algorithm. *Monthly Notices of the Royal Astronomical Society*, 382:315–324, November 2007. DOI: 10.1111/j.1365-2966.2007.12363.x. URL http://adsabs.harvard.edu/abs/2007MNRAS.382..315M.

L. Miller, C. Heymans, T. D. Kitching, L. van Waerbeke, T. Erben, H. Hildebrandt, H. Hoekstra, Y. Mellier, B. T. P. Rowe, J. Coupon, J. P. Dietrich, L. Fu, J. Harnois-Déraps, M. J. Hudson, M. Kilbinger, K. Kuijken, T. Schrabback, E. Semboloni, S. Vafaei, and M. Velander. Bayesian galaxy shape measurement for weak lensing surveys - III. Application to the Canada-France-Hawaii Telescope Lensing Survey. *Monthly Notices of the Royal Astronomical Society*, 429:2858–2880, March 2013. DOI: 10.1093/mnras/sts454. URL http://adsabs.harvard.edu/abs/2013MNRAS.429.2858M.

C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation.* 1973. URL http://adsabs.harvard.edu/abs/1973grav.book. ....M.

Joseph J. Mohr, August E. Evrard, Daniel G. Fabricant, and Margaret J. Geller. Cosmological Constraints from Observed Cluster X-Ray Morphologies. *Astrophysical Journal*, 447:8, Jul 1995. DOI: 10.1086/175852. URL https://ui.adsabs. harvard.edu/abs/1995ApJ...447....8M.

B. Moraes, J.-P. Kneib, A. Leauthaud, M. Makler, L. Van Waerbeke, K. Bundy, T. Erben, C. Heymans, H. Hildebrandt, L. Miller, H. Y. Shan, D. Woods, A. Charbonnier, and M. E. Pereira. The CFHT/MegaCam Stripe-82 Survey. In *Revista Mexicana de Astronomia y Astrofisica Conference Series*, volume 44 of *Revista Mexicana de Astronomia y Astrofisica, vol. 27*, pages 202–203, October 2014. URL http://adsabs. harvard.edu/abs/2014RMxAC..44..202M.

J. S. Mulchaey and A. I. Zabludoff. The Isolated Elliptical NGC 1132: Evidence for a Merged Group of Galaxies? *Astrophysical Journal*, 514:133–137, March 1999. DOI: 10.1086/306952. URL http://adsabs.harvard.edu/abs/1999ApJ...514. .133M.

R. Narayan, R. Blandford, and R. Nityananda. Multiple imaging of quasars by galaxies and clusters. *Nature*, 310:112–115, July 1984. DOI: 10.1038/310112a0. URL http://adsabs. harvard.edu/abs/1984Natur.310..112N.

J. F. Navarro, C. S. Frenk, and S. D. M. White. The assembly of galaxies in a hierarchically clustering universe. *Monthly Notices of the Royal Astronomical Society*, 275:56–66, July 1995. URL http://adsabs.harvard.edu/abs/1995MNRAS.275... 56N.

J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. *Astrophysical Journal*, 462:563, May 1996. DOI: 10.1086/177173. URL http://adsabs. harvard.edu/abs/1996ApJ...462..563N.

J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *Astrophysical*

Journal, 490:493–508, December 1997. URL http://
adsabs.harvard.edu/abs/1997ApJ...490..493N.

A. Niemiec, E. Jullo, M. Limousin, C. Giocoli, T. Erben,
H. Hildebrant, J.-P. Kneib, A. Leauthaud, M. Makler,
B. Moraes, M. E. S. Pereira, H. Shan, E. Rozo, E. Rykoff, and
L. Van Waerbeke. Stellar-to-halo mass relation of cluster
galaxies. *ArXiv e-prints*, March 2017. URL http:
//adsabs.harvard.edu/abs/2017arXiv170303348N.

E. Noether. Invariante Variationsprobleme. *Nachrichten von der
Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-
Physikalische Klasse*, 1:235–257, July 1918.

S. C. Odewahn, E. B. Stockwell, R. L. Pennington, R. M.
Humphreys, and W. A. Zumach. Automated Star/Galaxy
Discrimination With Neural Networks. *Astronomical Journal*,
103:318, Jan 1992. DOI: 10.1086/116063. URL https:
//ui.adsabs.harvard.edu/abs/1992AJ....103..318O.

Augustus Oemler. *The systematic properties of clusters of galaxies*.
PhD thesis, California Institute of Technology, Sep 1974. URL
https://ui.adsabs.harvard.edu/abs/1974PhDT.....
...370.

Masamune Oguri. A cluster finding algorithm based on the
multiband identification of red sequence galaxies. *Monthly
Notices of the Royal Astronomical Society*, 444(1):147–161,
Oct 2014. DOI: 10.1093/mnras/stu1446. URL https:
//ui.adsabs.harvard.edu/abs/2014MNRAS.444..147O.

Masamune Oguri and Philip J. Marshall. Gravitationally
lensed quasars and supernovae in future wide-field optical
imaging surveys. *Monthly Notices of the Royal Astronomical
Society*, 405(4):2579–2593, Jul 2010. DOI: 10.1111/j.1365-
2966.2010.16639.x. URL https://ui.adsabs.harvard.
edu/abs/2010MNRAS.405.2579O.

A. L. O'Mill, F. Duplancic, D. García Lambas, and L. Sodré, Jr.
Photometric redshifts and k-corrections for the Sloan Digital
Sky Survey Data Release 7. *Monthly Notices of the Royal
Astronomical Society*, 413:1395–1408, May 2011. DOI:

10.1111/j.1365-2966.2011.18222.x. URL http://adsabs.harvard.edu/abs/2011MNRAS.413.1395O.

D. Overbye. *Lonely Hearts of the Cosmos: The Story of the Scientific Quest for the Secret of the Universe*. Picador, 1993. ISBN 9780330295857. URL https://books.google.com.br/books?id=KwU6NAAACAAJ.

B. Paczynski. Giant luminous arcs discovered in two clusters of galaxies. *Nature*, 325:572–573, February 1987. DOI: 10.1038/325572a0. URL http://adsabs.harvard.edu/abs/1987Natur.325..572P.

T. Padmanabhan. *Structure Formation in the Universe*. May 1993. URL http://adsabs.harvard.edu/abs/1993sfu..book.....P.

A. Paranjape and R. K. Sheth. The luminosities of the brightest cluster galaxies and brightest satellites in SDSS groups. *Monthly Notices of the Royal Astronomical Society*, 423:1845–1855, June 2012. DOI: 10.1111/j.1365-2966.2012.21008.x. URL http://adsabs.harvard.edu/abs/2012MNRAS.423.1845P.

J. A. Peacock. Large-scale surveys and cosmic structure. *ArXiv Astrophysics e-prints*, September 2003. URL http://adsabs.harvard.edu/abs/2003astro.ph..9240P.

P. J. E. Peebles. The origin of galaxies and clusters of galaxies. *Science*, 224:1385–1391, June 1984. DOI: 10.1126/science.224.4656.1385. URL http://adsabs.harvard.edu/abs/1984Sci...224.1385P.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal*, 142:419–421, Jul 1965. DOI: 10.1086/148307. URL https://ui.adsabs.harvard.edu/abs/1965ApJ...142..419P.

Maria E. S. Pereira, Marcelle Soares-Santos, Martin Makler, James Annis, Huan Lin, Antonella Palmese, André Z. Vitorelli, Brian Welch, Gabriel B. Caminha, Thomas Erben, Bruno Moraes, and Huanyuan Shan. Weak-lensing calibration of a stellar mass-based mass proxy for redMaPPer and Voronoi

Tessellation clusters in SDSS Stripe 82. *Monthly Notices of the Royal Astronomical Society*, 474(1):1361–1372, Feb 2018. DOI: 10.1093/mnras/stx2831. URL https://ui.adsabs.harvard.edu/abs/2018MNRAS.474.1361P.

S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *Astrophysical Journal*, 517(2):565–586, Jun 1999. DOI: 10.1086/307221. URL https://ui.adsabs.harvard.edu/abs/1999ApJ...517..565P.

J. R. Peterson and A. C. Fabian. X-ray spectroscopy of cooling clusters. *Physics Reports*, 427(1):1–39, Apr 2006. DOI: 10.1016/j.physrep.2005.12.007. URL https://ui.adsabs.harvard.edu/abs/2006PhR...427....1P.

G. Pietrzyński, D. Graczyk, A. Gallenne, W. Gieren, I. B. Thompson, B. Pilecki, P. Karczmarek, M. Górski, K. Suchomska, M. Taormina, B. Zgirski, P. Wielgórski, Z. Kołaczkowski, P. Konorski, S. Villanova, N. Nardetto, P. Kervella, F. Bresolin, R. P. Kudritzki, J. Storm, R. Smolec, and W. Narloch. A distance to the Large Magellanic Cloud that is precise to one per cent. *Nature*, 567(7747):200–203, Mar 2019. DOI: 10.1038/s41586-019-0999-4. URL https://ui.adsabs.harvard.edu/abs/2019Natur.567..200P.

Planck Collaboration. Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, art. arXiv:1807.06209, Jul 2018.

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *ArXiv e-prints*, February

2015.   URL http://adsabs.harvard.edu/abs/2015arXiv150201589P.

T. J. Ponman, D. J. Allan, L. R. Jones, M. Merrifield, I. M. McHardy, H. J. Lehto, and G. A. Luppino.   A possible fossil galaxy group.   *Nature*, 369:462–464, June 1994.   DOI: 10.1038/369462a0.   URL http://adsabs.harvard.edu/abs/1994Natur.369..462P.

W. H. Press and P. Schechter.   Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *Astrophysical Journal*, 187:425–438, February 1974.   DOI: 10.1086/152650.   URL http://adsabs.harvard.edu/abs/1974ApJ...187..425P.

R. N. Proctor, C. M. de Oliveira, R. Dupke, R. L. de Oliveira, E. S. Cypriano, E. D. Miller, and E. Rykoff.   On the mass-to-light ratios of fossil groups. Are they simply dark clusters? *Monthly Notices of the Royal Astronomical Society*, 418: 2054–2073, December 2011.   DOI: 10.1111/j.1365-2966.2011.19625.x.   URL http://adsabs.harvard.edu/abs/2011MNRAS.418.2054P.

E. Puchwein and M. Bartelmann.   Probing the dynamical state of galaxy clusters.   *Astronomy & Astrophysics*, 474(3):745–754, Nov 2007.   DOI: 10.1051/0004-6361:20078020.   URL https://ui.adsabs.harvard.edu/abs/2007A&A...474..745P.

Mojtaba Raouf, Habib G. Khosroshahi, Trevor J. Ponman, Ali A. Dariush, Alireza Molaeinezhad, and Saeed Tavasoli. Ultimate age-dating method for galaxy groups; clues from the Millennium Simulations.   *Monthly Notices of the Royal Astronomical Society*, 442(2):1578–1585, Aug 2014.   DOI: 10.1093/mnras/stu963.   URL https://ui.adsabs.harvard.edu/abs/2014MNRAS.442.1578R.

Mojtaba Raouf, Habib G. Khosroshahi, and A. Dariush. Evolution of Galaxy Groups in the Illustris Simulation. *Astrophysical Journal*, 824(2):140, Jun 2016.   DOI: 10.3847/0004-637X/824/2/140.   URL https://ui.adsabs.harvard.edu/abs/2016ApJ...824..140R.

Jesper Rasmussen, John S. Mulchaey, Lei Bai, Trevor J. Ponman, Somak Raychaudhury, and Ali Dariush. The Suppression of Star Formation and the Effect of the Galaxy Environment in Low-redshift Galaxy Groups. *Astrophysical Journal*, 757(2):122, Oct 2012. DOI: 10.1088/0004-637X/757/2/122. URL https://ui.adsabs.harvard.edu/abs/2012ApJ...757..122R.

S. Refsdal. The gravitational lens effect. *Monthly Notices of the Royal Astronomical Society*, 128:295, 1964a. URL http://adsabs.harvard.edu/abs/1964MNRAS.128..295R.

S. Refsdal. On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect. *Monthly Notices of the Royal Astronomical Society*, 128:307, 1964b. URL http://adsabs.harvard.edu/abs/1964MNRAS.128..307R.

Dezső Ribli, László Dobos, and István Csabai. Galaxy shape measurement with convolutional neural networks. *Monthly Notices of the Royal Astronomical Society*, 489(4):4847–4859, Nov 2019. DOI: 10.1093/mnras/stz2374. URL https://ui.adsabs.harvard.edu/abs/2019MNRAS.489.4847R.

A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ*, 116:1009–1038, September 1998. DOI: 10.1086/300499. URL http://adsabs.harvard.edu/abs/1998AJ....116.1009R.

Adam G. Riess, Lucas M. Macri, Samantha L. Hoffmann, Dan Scolnic, Stefano Casertano, Alexei V. Filippenko, Brad E. Tucker, Mark J. Reid, David O. Jones, Jeffrey M. Silverman, Ryan Chornock, Peter Challis, Wenlong Yuan, Peter J. Brown, and Ryan J. Foley. A 2.4% Determination of the Local Value of the Hubble Constant. *Astrophysical Journal*, 826(1):56, Jul 2016. DOI: 10.3847/0004-637X/826/1/56. URL https://ui.adsabs.harvard.edu/abs/2016ApJ...826...56R.

Adam G. Riess, Stefano Casertano, Wenlong Yuan, Lucas M. Macri, and Dan Scolnic. Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM. *Astrophysical Journal*, 876(1):85, May 2019. DOI: 10.3847/1538-4357/ab1422. URL https://ui.adsabs.harvard.edu/abs/2019ApJ...876...85R.

E. S. Rykoff, B. P. Koester, E. Rozo, J. Annis, A. E. Evrard, S. M. Hansen, J. Hao, D. E. Johnston, T. A. McKay, and R. H. Wechsler. Robust Optical Richness Estimation with Reduced Scatter. *Astrophysical Journal*, 746:178, February 2012. DOI: 10.1088/0004-637X/746/2/178. URL http://adsabs.harvard.edu/abs/2012ApJ...746..178R.

E. S. Rykoff, E. Rozo, M. T. Busha, C. E. Cunha, A. Finoguenov, A. Evrard, J. Hao, B. P. Koester, A. Leauthaud, B. Nord, M. Pierre, R. Reddick, T. Sadibekova, E. S. Sheldon, and R. H. Wechsler. redMaPPer. I. Algorithm and SDSS DR8 Catalog. *Astrophysical Journal*, 785:104, April 2014. DOI: 10.1088/0004-637X/785/2/104. URL http://adsabs.harvard.edu/abs/2014ApJ...785..104R.

A. D. Sakharov. SPECIAL ISSUE: Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34:392–393, May 1967. DOI: 10.1070/PU1991v034n05ABEH002497. URL http://adsabs.harvard.edu/abs/1991SvPhU..34..392S.

A. Sandage. Current Problems in the Extragalactic Distance Scale. *Astrophysical Journal*, 127:513, May 1958. DOI: 10.1086/146483. URL http://adsabs.harvard.edu/abs/1958ApJ...127..513S.

P. Schechter. An analytic expression for the luminosity function for galaxies. *Astrophysical Journal*, 203:297–306, January 1976. DOI: 10.1086/154079. URL http://adsabs.harvard.edu/abs/1976ApJ...203..297S.

M. Schmidt. 3C 273 : A Star-Like Object with Large Red-Shift. *Nature*, 197:1040, March 1963. DOI: 10.1038/1971040a0.

URL http://adsabs.harvard.edu/abs/1963Natur.197.
1040S.

P. Schneider. *Extragalactic Astronomy and Cosmology*. 2006. URL
http://adsabs.harvard.edu/abs/2006eac..book.....
S.

P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*. 1992.
URL http://adsabs.harvard.edu/abs/1992grle.book.
....S.

SDSS Collaboration. The Eleventh and Twelfth Data Releases
of the Sloan Digital Sky Survey: Final Data from SDSS-III.
Astrophysical Journal Supplement Series, 219(1):12, Jul
2015. DOI: 10.1088/0067-0049/219/1/12. URL https:
//ui.adsabs.harvard.edu/abs/2015ApJS..219...12A.

H. Shan, J.-P. Kneib, R. Li, J. Comparat, T. Erben, M. Makler,
B. Moraes, L. Van Waerbeke, J. E. Taylor, and A. Charbonnier.
The Mass-Concentration Relation and the Stellar-to-Halo
Mass Ratio in the CFHT Stripe 82 Survey. *ArXiv e-prints*,
February 2015. URL http://adsabs.harvard.edu/abs/
2015arXiv150200313S.

H. Y. Shan, J.-P. Kneib, J. Comparat, E. Jullo, A. Charbonnier,
T. Erben, M. Makler, B. Moraes, L. Van Waerbeke, F. Courbin,
G. Meylan, C. Tao, and J. E. Taylor. Weak lensing mass map
and peak statistics in Canada-France-Hawaii Telescope Stripe 82
survey. *Monthly Notices of the Royal Astronomical Society*,
442:2534–2542, August 2014. DOI: 10.1093/mnras/stu1040.
URL http://adsabs.harvard.edu/abs/2014MNRAS.442.
2534S.

Erin Sheldon. NGMIX: Gaussian mixture models for 2D images,
Aug 2015. URL https://ui.adsabs.harvard.edu/abs/
2015ascl.soft08008S.

Erin S. Sheldon and Eric M. Huff. Practical Weak-lensing Shear
Measurement with Metacalibration. *Astrophysical Journal*, 841
(1):24, May 2017. DOI: 10.3847/1538-4357/aa704b. URL
https://ui.adsabs.harvard.edu/abs/2017ApJ...841.
..24S.

R. A. Skibba and A. V. Macciò. Properties of dark matter haloes and their correlations: the lesson from principal component analysis. *Monthly Notices of the Royal Astronomical Society*, 416:2388–2400, September 2011. DOI: 10.1111/j.1365-2966.2011.19218.x. URL http://adsabs.harvard.edu/abs/2011MNRAS.416.2388S.

V. M. Slipher. Radial velocity observations of spiral nebulae. *The Observatory*, 40:304–306, August 1917. URL http://adsabs.harvard.edu/abs/1917Obs....40..304S.

Graham P. Smith, Alastair C. Edge, Vincent R. Eke, Robert C. Nichol, Ian Smail, and Jean-Paul Kneib. Measuring σ_8 with Cluster Lensing: Biases from Unrelaxed Clusters. *Astrophysical Journal Letters*, 590(2):L79–L82, Jun 2003. DOI: 10.1086/376747. URL https://ui.adsabs.harvard.edu/abs/2003ApJ...590L..79S.

Graham P. Smith, Jean-Paul Kneib, Ian Smail, Pasquale Mazzotta, Harald Ebeling, and Oliver Czoske. A Hubble Space Telescope lensing survey of X-ray luminous galaxy clusters - IV. Mass, structure and thermodynamics of cluster cores at z= 0.2. *Monthly Notices of the Royal Astronomical Society*, 359(2): 417–446, May 2005. DOI: 10.1111/j.1365-2966.2005.08911.x. URL https://ui.adsabs.harvard.edu/abs/2005MNRAS.359..417S.

G. F. Smoot. COBE observations and results. In L. Maiani, F. Melchiorri, and N. Vittorio, editors, *3K cosmology*, volume 476 of *American Institute of Physics Conference Series*, pages 1–10, May 1999. DOI: 10.1063/1.59326. URL http://adsabs.harvard.edu/abs/1999AIPC..476....1S.

Irwin Sobel and G Feldman. An isotropic 3x3 image gradient operator for image processing. *Mach. Vis. Three–Dimens. Scenes*, (June):376–379, 1968.

J. G. v Soldner. Ueber die Ablenkung eines Lichtstrals von seiner geradlinigen Bewegung, durch die Attraktion eines Weltkörpers, an welchem er nahe vorbei geht. *Astronomisches Jahrbuch für das Jahr 1804*, pages 161–172, March 1804.

J. Y. H. Soo, B. Moraes, B. Joachimi, W. Hartley, O. Lahav, A. Charbonnier, M. Makler, M. E. S. Pereira, J. Comparat, T. Erben, A. Leauthaud, H. Shan, and L. Van Waerbeke. Morpho-z: improving photometric redshifts with galaxy morphology. *ArXiv e-prints*, July 2017. URL http://adsabs.harvard.edu/abs/2017arXiv170703169S.

G. Soucail, B. Fort, Y. Mellier, and J. P. Picat. A blue ring-like structure, in the center of the A 370 cluster of galaxies. *Astronomy & Astrophysics*, 172:L14–L16, January 1987. URL http://adsabs.harvard.edu/abs/1987A%26A...172L..14S.

G. Soucail, Y. Mellier, B. Fort, G. Mathez, and M. Cailloux. The giant arc in A 370 - Spectroscopic evidence for gravitational lensing from a source at Z = 0.724. *Astronomy & Astrophysics*, 191:L19–L21, February 1988. URL http://adsabs.harvard.edu/abs/1988A%26A...191L..19S.

V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435: 629–636, June 2005. DOI: 10.1038/nature03597. URL http://adsabs.harvard.edu/abs/2005Natur.435..629S.

G. Steigman. Primordial Nucleosynthesis in the Precision Cosmology Era. *Annual Review of Nuclear and Particle Science*, 57:463–491, November 2007. DOI: 10.1146/annurev.nucl.56.080805.140437. URL http://adsabs.harvard.edu/abs/2007ARNPS..57..463S.

W. R. Stoeger, R. Maartens, and G. F. R. Ellis. Proving almost-homogeneity of the universe: an almost Ehlers-Geren-Sachs theorem. *Astrophysical Journal*, 443:1–5, April 1995. DOI: 10.1086/175496. URL http://adsabs.harvard.edu/abs/1995ApJ...443....1S.

N. R. Tanvir, D. B. Fox, A. J. Levan, E. Berger, K. Wiersema, J. P. U. Fynbo, A. Cucchiara, T. Krühler, N. Gehrels, J. S. Bloom, J. Greiner, P. A. Evans, E. Rol, F. Olivares, J. Hjorth,

P. Jakobsson, J. Farihi, R. Willingale, R. L. C. Starling, S. B. Cenko, D. Perley, J. R. Maund, J. Duke, R. A. M. J. Wijers, A. J. Adamson, A. Allan, M. N. Bremer, D. N. Burrows, A. J. Castro-Tirado, B. Cavanagh, A. de Ugarte Postigo, M. A. Dopita, T. A. Fatkhullin, A. S. Fruchter, R. J. Foley, J. Gorosabel, J. Kennea, T. Kerr, S. Klose, H. A. Krimm, V. N. Komarova, S. R. Kulkarni, A. S. Moskvitin, C. G. Mundell, T. Naylor, K. Page, B. E. Penprase, M. Perri, P. Podsiadlowski, K. Roth, R. E. Rutledge, T. Sakamoto, P. Schady, B. P. Schmidt, A. M. Soderberg, J. Sollerman, A. W. Stephens, G. Stratta, T. N. Ukwatta, D. Watson, E. Westra, T. Wold, and C. Wolf. A γ-ray burst at a redshift of $z \sim 8.2$. *Nature*, 461:1254–1257, October 2009. DOI: 10.1038/nature08459. URL http://adsabs.harvard.edu/abs/2009Natur.461.1254T.

Argyro Tasitsiomi, Andrey V. Kravtsov, Stefan Gottlöber, and Anatoly A. Klypin. Density Profiles of ΛCDM Clusters. *Astrophysical Journal*, 607(1):125–139, May 2004. DOI: 10.1086/383219. URL https://ui.adsabs.harvard.edu/abs/2004ApJ...607..125T.

J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. *Astrophysical Journal*, 688:709–728, December 2008. DOI: 10.1086/591439. URL http://adsabs.harvard.edu/abs/2008ApJ...688..709T.

M. P. Viero, V. Asboth, I. G. Roseboom, L. Moncelsi, G. Marsden, E. Mentuch Cooper, M. Zemcov, G. Addison, A. J. Baker, A. Beelen, J. Bock, C. Bridge, A. Conley, M. J. Devlin, O. Doré, D. Farrah, S. Finkelstein, A. Font-Ribera, J. E. Geach, K. Gebhardt, A. Gill, J. Glenn, A. Hajian, M. Halpern, S. Jogee, P. Kurczynski, A. Lapi, M. Negrello, S. J. Oliver, C. Papovich, R. Quadri, N. Ross, D. Scott, B. Schulz, R. Somerville, D. N. Spergel, J. D. Vieira, L. Wang, and R. Wechsler. The Herschel Stripe 82 Survey (HerS): Maps and Early Catalog. *Astrophysical Journal, Supplement*, 210:22, February 2014. DOI: 10.1088/0067-0049/210/2/22. URL

http://adsabs.harvard.edu/abs/2014ApJS..210...
22V.

André Z. Vitorelli, Eduardo S. Cypriano, Martín Makler, Maria E. S. Pereira, Thomas Erben, and Bruno Moraes. On mass concentrations and magnitude gaps of galaxy systems in the CS82 survey. *Monthly Notices of the Royal Astronomical Society*, 474(1):866–875, Feb 2018. DOI: 10.1093/mnras/stx2791. URL https://ui.adsabs.harvard.edu/abs/2018MNRAS.474..866V.

A. Voevodkin, K. Borozdin, K. Heitmann, S. Habib, A. Vikhlinin, A. Mescheryakov, A. Hornstrup, and R. Burenin. Fossil Systems in the 400d Cluster Catalog. *Astrophysical Journal*, 708:1376–1387, January 2010. DOI: 10.1088/0004-637X/708/2/1376. URL http://adsabs.harvard.edu/abs/2010ApJ...708.1376V.

M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *Monthly Notices of the Royal Astronomical Society*, 444:1518–1547, October 2014. DOI: 10.1093/mnras/stu1536. URL http://adsabs.harvard.edu/abs/2014MNRAS.444.1518V.

A. M. von Benda-Beckmann, E. D'Onghia, S. Gottlöber, M. Hoeft, A. Khalatyan, A. Klypin, and V. Müller. The fossil phase in the life of a galaxy group. *Monthly Notices of the Royal Astronomical Society*, 386:2345–2352, June 2008. DOI: 10.1111/j.1365-2966.2008.13221.x. URL http://adsabs.harvard.edu/abs/2008MNRAS.386.2345V.

A. von der Linden, P. N. Best, G. Kauffmann, and S. D. M. White. How special are brightest group and cluster galaxies? *Monthly Notices of the Royal Astronomical Society*, 379:867–893, August 2007. DOI: 10.1111/j.1365-2966.2007.11940.x. URL http://adsabs.harvard.edu/abs/2007MNRAS.379..867V.

D. Walsh, R. F. Carswell, and R. J. Weymann. 0957 + 561 A, B - Twin quasistellar objects or gravitational lens. *Nature*, 279:

381–384, May 1979. DOI: 10.1038/279381a0. URL http://adsabs.harvard.edu/abs/1979Natur.279..381W.

Risa H. Wechsler, James S. Bullock, Joel R. Primack, Andrey V. Kravtsov, and Avishai Dekel. Concentrations of Dark Halos from Their Assembly Histories. *Astrophysical Journal*, 568 (1):52–70, Mar 2002. DOI: 10.1086/338765. URL https://ui.adsabs.harvard.edu/abs/2002ApJ...568...52W.

D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo. Observational probes of cosmic acceleration. *Physics Reports*, 530:87–255, September 2013. DOI: 10.1016/j.physrep.2013.05.001. URL http://adsabs.harvard.edu/abs/2013PhR...530...87W.

S. Weinberg. *Cosmology*. Oxford University Press, 2008. URL http://adsabs.harvard.edu/abs/2008cosm.book.....W.

Nicholas Weir, Usama M. Fayyad, and S. Djorgovski. Automated Star/Galaxy Classification for Digitized Poss-II. *Astronomical Journal*, 109:2401, Jun 1995. DOI: 10.1086/117459. URL https://ui.adsabs.harvard.edu/abs/1995AJ....109.2401W.

Z. L. Wen, J. L. Han, and F. S. Liu. A Catalog of 132,684 Clusters of Galaxies Identified from Sloan Digital Sky Survey III. *Astrophysical Journal Suplement*, 199(2):34, Apr 2012. DOI: 10.1088/0067-0049/199/2/34.

C. M. Will. The Confrontation between General Relativity and Experiment. *Living Reviews in Relativity*, 17:4, June 2014. DOI: 10.12942/lrr-2014-4. URL http://adsabs.harvard.edu/abs/2014LRR....17....4W.

C. O. Wright and T. G. Brainerd. Gravitational Lensing by NFW Halos. *Astrophysical Journal*, 534:34–40, May 2000. DOI: 10.1086/308744. URL http://adsabs.harvard.edu/abs/2000ApJ...534...34W.

Kouichi Yamaguchi, Kenji Sakamoto, Toshio Akabane, and Yoshiji Fujimoto. A neural network for speaker-independent isolated word recognition. *ICSLP-1990*, pages 1077–1080, 1990.

X. Yang, H. J. Mo, F. C. van den Bosch, Y. P. Jing, S. M.
    Weinmann, and M. Meneghetti. Weak lensing by galaxies in
    groups and clusters - I. Theoretical expectations. *Monthly
    Notices of the Royal Astronomical Society*, 373:1159–1172,
    December 2006. DOI: 10.1111/j.1365-2966.2006.11091.x.
    URL http://adsabs.harvard.edu/abs/2006MNRAS.373.
    1159Y.

S. Zarattini, R. Barrena, M. Girardi, N. Castro-Rodriguez,
    W. Boschin, J. A. L. Aguerri, J. Méndez-Abreu, R. Sánchez-
    Janssen, C. Catalán-Torrecilla, E. M. Corsini, C. del Burgo,
    E. D'Onghia, N. Herrera-Ruiz, J. Iglesias-Páramo, E. Jimenez
    Bailon, M. Lozada Muoz, N. Napolitano, and J. M. Vilchez.
    Fossil group origins. IV. Characterization of the sample and
    observational properties of fossil systems. *Astronomy &
    Astrophysics*, 565:A116, May 2014. DOI: 10.1051/0004-
    6361/201323351. URL http://adsabs.harvard.edu/
    abs/2014A%26A...565A.116Z.

S. Zarattini, J. A. L. Aguerri, R. Sánchez-Janssen, R. Barrena,
    W. Boschin, C. del Burgo, N. Castro-Rodriguez, E. M. Corsini,
    E. D'Onghia, M. Girardi, J. Iglesias-Páramo, A. Kundert,
    J. Méndez-Abreu, and J. M. Vilchez. Fossil group origins. V.
    The dependence of the luminosity function on the magnitude
    gap. *Astronomy & Astrophysics*, 581:A16, September
    2015. DOI: 10.1051/0004-6361/201425506. URL http:
    //adsabs.harvard.edu/abs/2015A%26A...581A..16Z.

S. Zarattini, M. Girardi, J. A. L. Aguerri, W. Boschin,
    R. Barrena, C. del Burgo, N. Castro-Rodriguez, E. M. Corsini,
    E. D'Onghia, A. Kundert, J. Méndez-Abreu, and R. Sánchez-
    Janssen. Fossil group origins. VII. Galaxy substructures in
    fossil systems. *Astronomy & Astrophysics*, 586:A63, February
    2016. DOI: 10.1051/0004-6361/201527175. URL http:
    //adsabs.harvard.edu/abs/2016A%26A...586A..63Z.

Y. B. Zel'dovich. Gravitational instability: An approximate theory
    for large density perturbations. *Astronomy & Astrophysics*, 5:
    84–89, March 1970. URL http://adsabs.harvard.edu/
    abs/1970A%26A.....5...84Z.

Y. Zhang, T. Jeltema, D. L. Hollowood, S. Everett, E. Rozo, A. Farahi, A. Bermeo, S. Bhargava, P. Giles, A. K. Romer, R. Wilkinson, E. S. Rykoff, A. Mantz, H. T. Diehl, A. E. Evrard, C. Stern, D. Gruen, A. von der Linden, M. Splettstoesser, X. Chen, M. Costanzi, S. Allen, C. Collins, M. Hilton, M. Klein, R. G. Mann, M. Manolopoulou, G. Morris, J. Mayers, M. Sahlen, J. Stott, C. Vergara Cervantes, P. T. P. Viana, R. H. Wechsler, S. Allam, S. Avila, K. Bechtol, E. Bertin, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, L. N. da Costa, J. De Vicente, S. Desai, J. P. Dietrich, P. Doel, B. Flaugher, P. Fosalba, J. Frieman, J. García-Bellido, E. Gaztanaga, R. A. Gruendl, J. Gschwend, G. Gutierrez, W. G. Hartley, K. Honscheid, B. Hoyle, E. Krause, K. Kuehn, N. Kuropatkin, M. Lima, M. A. G. Maia, J. L. Marshall, P. Melchior, F. Menanteau, C. J. Miller, R. Miquel, R. L. C. Ogando, A. A. Plazas, E. Sanchez, V. Scarpine, R. Schindler, S. Serrano, I. Sevilla-Noarbe, M. Smith, M. Soares-Santos, E. Suchyta, M. E. C. Swanson, G. Tarle, D. Thomas, D. L. Tucker, V. Vikram, W. Wester, and DES Collaboration. Dark Energy Surveyed Year 1 results: calibration of cluster mis-centring in the redMaPPer catalogues. *Monthly Notices of the Royal Astronomical Society*, 487(2):2578–2593, Aug 2019. DOI: 10.1093/mnras/stz1361. URL https://ui.adsabs.harvard.edu/abs/2019MNRAS.487.2578Z.

A. Zitrin, I. Labbe, S. Belli, R. Bouwens, R. S. Ellis, G. Roberts-Borsani, D. P. Stark, P. A. Oesch, and R. Smit. Lyman-alpha Emission from a Luminous z=8.68 Galaxy: Implications for Galaxies as Tracers of Cosmic Reionization. *ArXiv e-prints*, July 2015. URL http://adsabs.harvard.edu/abs/2015arXiv150702679Z.

F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933. URL http://adsabs.harvard.edu/abs/1933AcHPh...6..110Z.

F. Zwicky. Nebulae as Gravitational Lenses. *Physical Review*, 51:290–290, February 1937. DOI: 10.1103/PhysRev.51.290.

URL http://adsabs.harvard.edu/abs/1937PhRv...51.
.290Z.

Index