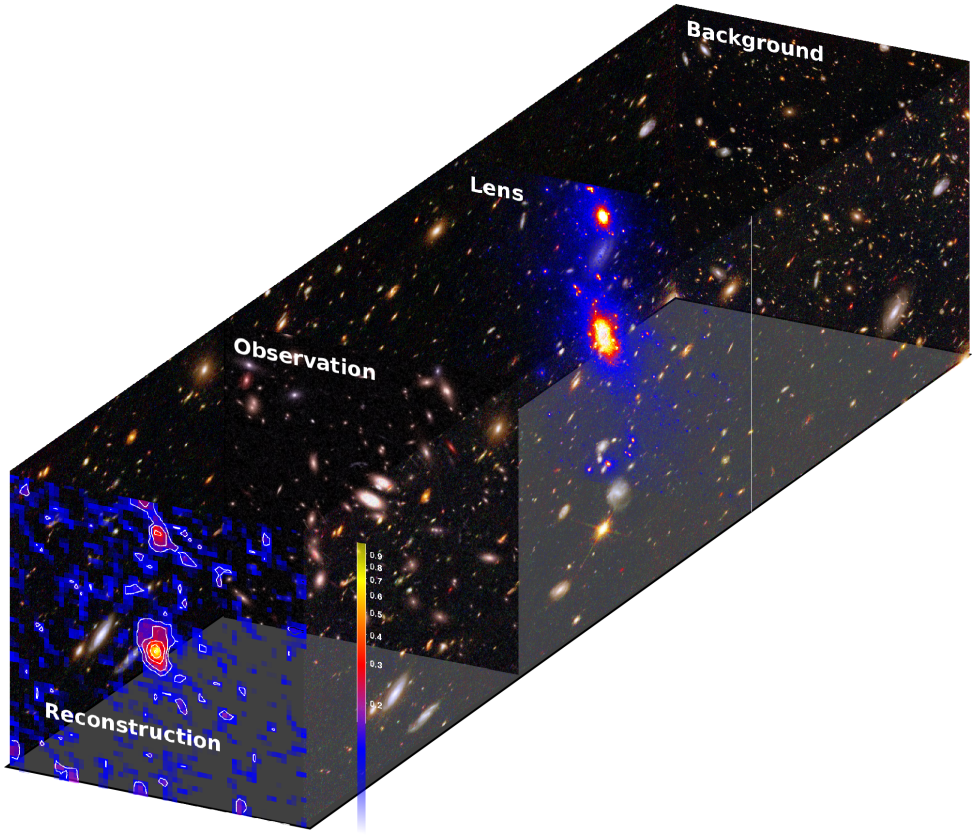
Ruprecht-Karls-Universität Heidelberg
Fakultät für Physik und Astronomie

-Gravitational Lensing-

An advanced method to recover the mass distribution of galaxy clusters



Julian Merten

Advisors:

Matthias Bartelmann
Massimo Meneghetti

Referees:

Matthias Bartelmann
Hans-Walter Rix

Heidelberg & Bologna, March/April 2010

And the sky is filled with light
Can you see it?
All the black is really white
If you believe it

As your time is running out
Let me take away your doubt
You can find a better a place
In this twilight

Trent Reznor

Über eine effiziente Methode zur Massenprofilrekonstruktion von Galaxienhaufen

Die vorliegende Arbeit beschäftigt sich mit Galaxienhaufen. Diese massereichsten, gravitativ gebundenen Objekte im beobachtbaren Universum repräsentieren das obere Ende der Massenfunktion und sind von speziellem Interesse für die Kosmologie. Nicht nur lassen sich mehrere kosmologische Parameter aus der Beobachtung und vor allem aus der Massenbestimmung von Galaxienhaufen ableiten, sie stellen auch ideale kosmische Laboratorien dar, welche einen direkten Vergleich zwischen Beobachtung und numerischer Simulation erlauben.

Die vielleicht vielversprechendste Methode um die Eigenschaften von Galaxienhaufen zu ermitteln ist der Gravitationslinseneffekt. Das Licht entfernter Hintergrundgalaxien wird aufgrund der hohen Massekonzentration in einem Galaxienhaufen auf dem Weg zum Beobachter abgelenkt und trägt daher Informationen über den Deflektor. In dieser Arbeit entwickeln wir eine neue, moderne Methode welche den sogenannten starken und schwachen Gravitationslinseneffekt optimal kombiniert und daher eine nichtparametrische Rekonstruktion der Massenverteilung des Deflektors erlaubt. Diese Methode ist in einem fortschrittlichen numerischen Algorithmus implementiert, welcher effiziente numerische Verfahren und parallele Höchstleistungs-Computersysteme ausnutzt.

Mit Rekonstruktionen numerisch simulierter Galaxienhaufen zeigen wir die Leistungsfähigkeit unserer Methode, im Vergleich mit etablierten Techniken. Wir schließen unsere Arbeit mit Rekonstruktion und Analyse von MS2137.3-2353 und CL0024+1654, zweier wohlbekannter Galaxienhaufen die spektakuläre Phänomene des starken Gravitationslinseneffektes aufweisen.

An advanced method to recover the mass distribution of galaxy clusters

This work shall be on clusters of galaxies. Those most massive, gravitationally bound objects in the observable Universe represent the high-mass tail of the mass function, rendering them as objects of interest for cosmology. Not only that they allow for the derivation of several cosmological parameters, but they are also ideal cosmic laboratories. Direct comparisons between numerical simulations and observations are particularly appealing in the case of clusters, as we will show.

Maybe the most promising method to derive cluster properties from observations is gravitational lensing. Light rays of distant background sources are bent on the way to the observer, due to the high mass concentrations of clusters, and thereby carry important information about the deflector. In this work we develop an advanced, nonparametric method to recover the mass distribution of galaxy clusters by combining weak and strong gravitational lensing. The underlying numerical algorithm makes use of modern concepts of high-performance computing and is fully parallelised.

We proof the capabilities of our method, compared to established methods, while reconstructing simulated clusters of galaxies and capitalising realistic lensing scenarios. We close our work with the reconstruction of two well-known, strongly lensing clusters, namely MS2137.3-2353 and CL0024+1654.

Contents

List of Figures

List of Tables

1 Introduction: Our picture of the Universe

The last century saw the advent of stringent mathematical formulations of physical principles, not only to explain experimental observations but also to predict them. One outstanding example of such a mathematical formulation is Einstein's theory of general relativity (see e.g. Carroll, 2003, for a modern overview), maybe marking the beginning of a modern interpretation of theoretical physics, together with the discovery of quantised phenomena in physical processes, which led to the formulation of quantum mechanics and peaked in the complex mathematical framework of quantum field theory (see e.g Srednicki, 2007, as a comprehensive textbook).

Today, we are in the fortunate position to be able to describe the four discovered fundamental interaction in nature. But unfortunately, this is not possible with a single, closed theory. It needs two distinct, so-called standard models to describe the Universe ranging from microscopic to macroscopic scales. The standard model of particle physics provides the microscopic part, describing three of the four fundamental interactions, being the electroweak and the strong force. This theory is a quantum field theory, or to be more precise a non-abelian gauge theory, which describes leptons, hadrons and scalar fields by mathematical objects, so-called Weyl-fields[1] in the case of leptons and hadrons and a complex scalar field. Mappings between those fields must obey symmetry under certain gauge transformations, given by the gauge transformation group $SU(2)^2 \times U(1)^3$ in the case of the electroweak force and the $SU(3)^4$ in the case of the strong force. In a quantised theory, those gauge transformation groups manifest themselves as bosonic gauge particles, namely as photons, gluons, $W^{\pm}$ and Z^0 bosons. The gauge transformation groups define the "rules" for all physical processes within the theory through their symmetry properties and the associated particles are therefore often called force carriers. The complete particle content of the standard model is shown in Fig. 1.1. This complex mathematical structure is mainly inspired by experimental results (see e.g. Perkins, 1987, for an overview), but also predicts the existence of the scalar Higgs particle, whose

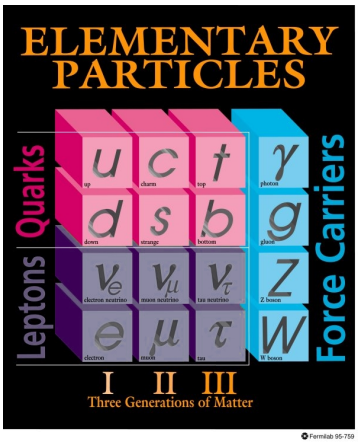


Figure 1.1: The standard model of particle physics as it presents itself today. Hadrons and leptons appear in three families, the fundamental forces are carried by bosonic gauge particles. A scalar Higgs particle is predicted, but not found experimentally, yet. Gravity is completely neglected. (Courtesy of Fermilab)

[1] Weyl fields, sometimes also called Weyl spinors, obey the right transformation properties within the Lorentz group. They are named after Hermann Klaus Hugo Weyl (1885-1955) a German mathematician.
[2] The set of unitary 2×2 matrices with determinant 1.
[3] The set of unitary numbers.
[4] The set of unitary 3×3 matrices with determinant 1.

detection was one of the main drivers for the construction of the Large Hadron Collider (LHC). In total, the standard model is described by 20 independent parameters, including e.g. the different gauge-coupling constants and the mass of the Higgs boson. This large number of parameters renders the standard model already somewhat unappealing, but despite its great success in predicting the results of a larger number of high-precision experiments, another pretty obvious conceptual problem of the standard model of particle physics is the complete disregard of gravity, being the forth fundamental force, dominating the Universe on large scales.

The physical theory closing the gap is general relativity. Einstein's field equations connect the metric field, describing the geometry of spacetime, to the energy content in the Universe. Relativity is a field theory, but not a quantum field theory. It can be shown, that a theory describing a spin-2 field with respect to the Lorentz group cannot be renormalisable (e.g. Srednicki, 2007). This renders it impossible to include Einstein's gravity in the quantisation machinery of quantum field theory. Even when not matching into a unified mathematical framework yet, only the combination of quantum field theory and general relativity describes our Universe completely.

The foundation of the standard model of cosmology is general relativity. While assuming an isotropic and homogeneous metric, the left-hand side of the field equations largely simplifies. The right-hand side describes the content of the Universe. The actual parameters, describing the cosmological standard model, are motivated by observations (see e.g. Bartelmann, 2010, and references therein). The isotropy of the Cosmic Microwave Background (CMB) (see e.g. Durrer, 2008, for an extensive overview) indicates, that all regions of the Universe were in causal contact at an early stage of the cosmic evolution. This, and the fact that the Universe seems to be spatially flat, again indicated by CMB observations, suggests an era of extremely rapid expansion in the very beginning of the Universe. This predicted epoch, called inflation marks the starting point of our ability to describe observationally and theoretically the global evolution of the Universe, by providing a mechanism to produce the initial perturbations in the cosmic fluid (see e.g. Liddle and Lyth, 2000), which partly manifest themselves today as e.g. stars and galaxies. The physics before inflation basically is a mystery to us, because existing theories will presumably break down while reaching the Planck scale[5]. Since the end of inflation, the Universe expanded with a rate which can be determined by observations. The current expansion rate is given by the Hubble[6] constant, found to be $H_0 = 72 \pm 8$ km s^{-1}Mpc^{-1} (Freedman et al., 2001)[7]. A value which is confirmed by observations of distant Type Ia supernovae, finding $H_0 = 74.2 \pm 3.6$ km s^{-1}Mpc^{-1} (Riess et al., 2009) and furthermore indicating that the Universe entered a phase of rapid expansion, again. One explanation for this accelerated expansion is given by one particular ingredient in the cosmic fluid, called dark energy or sometimes also simply denoted as cosmological constant. This mysterious energy-density component seems to be necessary also for other reasons: Measurements of element abundances in the Universe (see e.g. Steigman, 2007, for a review) indicate that ordinary matter, as it is described by the standard model of particle physics,

[5] The Planck scale, e.g. expressed as a mass, is the fundamental constant which can be constructed out of the fundamental constants in general relativity, Newton's constant and the speed of light and out of the fundamental constant in quantum mechanics, Planck's constant: $m_{\text{pl}} = \sqrt{\hbar c/G}$. Physically the Planck scale describes the presumed regime where the spacetime metric cannot be protected any more from quantum corrections. ($\hbar = 6.58211889(26) \times 10^{-16}$eV sec; $G = 6.673(10) \times 10^{-8}$cm^{3}g^{-1}sec^{-2}; $c = 2.99792458 \times 10^{10}$cm sec^{-1})

[6] The cosmic expansion is sometimes also called Hubble flow, in reference to Edwin Hubble who was one of the first to discover an increase in the recession velocity of galaxies as a function of distance (Hubble, 1929). The first discovery of this phenomenon was achieved by Vesto Sliper in the 1910's.

[7] Our prefered unit of length in this work is the Mpc. 1 Mpc $= 3.09 \times 10^{22}$m $= 3.26 \times 10^6$ light years.

contributes only a small fraction to the total matter density of the Universe. A finding which is supported by the shape of rotation curves in spiral galaxies (Sofue and Rubin, 2001). The missing matter component is thought to consist of an unknown form of matter, called dark matter since it is interacting only gravitationally. Careful mass determinations of galaxy clusters show, that matter contributes $\sim 25\%$ to the energy density of the Universe. While also considering the fact that the Universe is spatially flat, as the CMB tells us and the fact that the contribution of radiation and neutrinos is negligible today, dark energy provides the missing ingredient. A visual summary of our current understanding of the cosmic timeline is given in Fig. 1.2.

The main task of cosmology (see Dodelson, 2003; Weinberg, 2008, as textbook suggestions) is to precisely determine the parameters, fixing the cosmological standard model (Lahav and Liddle, 2010), as we show the most important of them in Tab. 1.1. The high-precision measurement of the CMB temperature seems to provide the most accurate way to achieve this task (Spergel et al., 2007; Komatsu et al., 2009; Dunkley et al., 2009), which will be even improved by the analysis of the data from the `Planck` satellite. Together with the constraints coming from observations from distant supernovae, the clustering of structure in the Universe, as e.g. observed by cosmological lensing (Bartelmann and Schneider, 2001) or baryonic acoustic oscillations (BAO) (Eisenstein, 2005; Percival et al., 2007), the cosmological parameters can be fixed within surprisingly small bounds. But also the standard model of cosmology suffers from conceptual problems. Naturally, it considers only gravity, but also the nature of the two main ingredients of the Universe today, dark energy and dark matter, though they appear as parameters, is left open.

parameter	value
Hubble parameter	$h = 0.72 \pm 0.03$
Total matter density	$\Omega_\mathrm{m} h^2 = 0.133 \pm 0.006$
Baryon density	$\Omega_\mathrm{b} h^2 = 0.0227 \pm 0.0006$
Cosmological constant	$\Omega_\Lambda = 0.742 \pm 0.03$
Power spectrum normalisation	$\sigma_8 = 0.796 \pm 0.036$
Spectral index	$n_s = 0.963^{+0.014}_{-0.015}$
Reionisation optical depth	$\tau = 0.087 \pm 0.017$

Table 1.1: The main parameters of the cosmological standard model today, as they are obtained from `WMAP5` (Komatsu et al., 2009).

The point where the cosmological standard model and the one of particle physics meet again is actually connected to their most severe problems. Both models suffer from a pretty severe, so-called hierarchy problem. In the standard model of particle physics, the smallest intrinsic length scale (or largest energy scale) is the scale of the electroweak phase transition, so the energy below which the symmetry $SU(2) \times U(1)$ is broken, to the symmetry under $U(1)$ gauge transformations, known as electromagnetism. This energy scale turns out to be 17 orders of magnitude smaller than the natural scale of an effective field theory neglecting gravity, being the Planck scale. This suggests the introduction of a new symmetry, broken at a much higher energy scale. This extension of the standard model is known under the concept of supersymmetry, implying that every fermion in the standard model should have a bosonic partner and vice versa. Supersymmetry also provides the most promising dark-matter candidate, being the so-called neutralino. But a severe problem remains, namely the hierarchy problem in the cosmological standard model: The value of the dark-energy parameter today is 120 orders of magnitude lower than the one would expect from a simple, but conceptually appealing, cosmological constant in Einstein's field equations. It is not unlikely, that the hierarchy problem of dark energy is related to the hierarchy problem in quantum

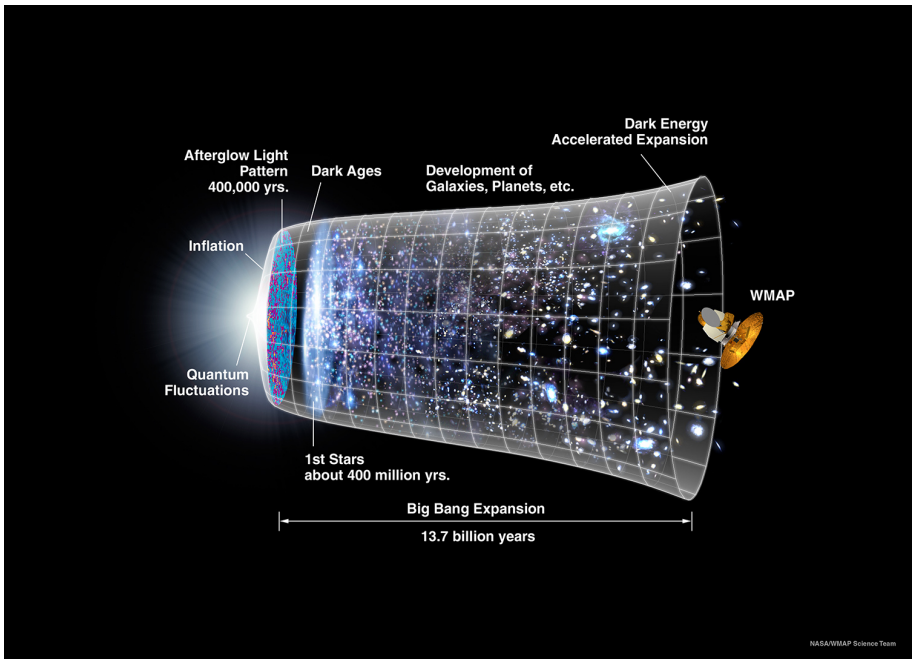


Figure 1.2: The history of the Universe as it is seen by the cosmological standard model. (Courtesy of the NASA/WMAP Science team)

field theory, regarding the fact that gravity is completely left out of the game of quantised field theories. A natural explanation for both hierarchy problems might arise from a “theory of everything”, finally unifying all four fundamental interactions. Recent developments in the field of supersymmetry and string theory show, that theories going beyond both standard models (see e.g. Dine, 2007, for an overview) may indeed be able to unify the two.

For the moment it is important to understand better the nature of dark matter and dark energy. Therefore, we will focus again on the cosmological standard model, also often called the ΛCDM model and especially on its dark-matter component. In the following, we will introduce a method, based on gravitational lensing, which is able to map accurately the mass distribution of galaxy clusters. We describe the concept of gravitational lensing in chapter 2 and highlight the special significance of clusters of galaxies within a cosmological model in chapter 3. Chapters 4 and 5 focus on the concrete implementation of our method, by also incorporating advanced numerical techniques and modern ideas of high-performance computing. The last three chapters 6, 7 and 8 are dedicated to the results on simulations and on real observations, obtained with our method and we explain there, how our method can contribute to the task of shedding light on the dark components of our current cosmological picture.

2 Gravitational Lensing

Gravitational lensing, or just lensing as we will often call it, has become a very important phenomenon in astrophysics during the last decades and created a new, independent field of research within the community. A very peculiar property of lensing is that it is observed on virtually every angular length scale and in several wave bands.

So called microlensing techniques can be used for the search of extrasolar planets, where the presence of a transiting planet characteristically changes the light curve of its host star. Using this technique, also relatively cold planets with low mass, have been found (Beaulieu et al., 2006) and large telescope networks constantly monitor the light curves of candidate systems (Albrow et al., 1995).

Moving up from the size of planets and stars to the size of galaxies, lensing is used to enhance the observations of distant quasi-stellar objects (QSOs), often called quasars. In fact, Walsh et al. (1979) discovered the first gravitational lens system, which consisted of the double image of `QSO 0957+561` seen through a foreground galaxy acting as gravitational lens.

On the scale of galaxy clusters, which deflect the light of distant background galaxies, the effects of lensing are presumably most spectacular. This work is dedicated to this special field of lensing and we shall discuss it in more detail later on.

Reaching the largest possible scales, the whole Universe can act as a lens and its energy-density content distorts the shapes of astronomical objects all over the sky. These effects are actually observable and produce a significant signal, if treated statistically, in observations with a large field-of-view. This so-called cosmic-shear signal can be used to constrain the growth of structure in the Universe in a very powerful and direct way. The European Space Agency (ESA) satellite project `Euclid` (Refregier et al., 2010; Cimatti et al., 2009), which recently passed the assessment phase of ESA's Cosmic Visions program, is dedicated to shed light on the nature of dark energy through the impact of this mysterious component of the Universe on the growth of structure. Euclid is supposed to conduct an enormous galaxy redshift survey and thereby reveal the shape and the time evolution of the spatial correlation functions of structures, on a wide range of scales. Lensing will be a key player in this survey.

This chapter is organised as follows: We will give a basic derivation of the effects of lensing from general relativity in Sec. 2.1 and focus on more specific applications afterwards. Weak lensing is covered in Sec. 2.2 and strong lensing in Sec. 2.3. For a comprehensive textbook, which covers the complete context of lensing, the reader is referred to Schneider et al. (2006).

2.1 Basic theory of lensing

As we have seen, lensing is used in a variety of applications on different length scales, but the basic effect is always the same and very simple. Due to the curvature of spacetime, caused by e.g. massive objects, light rays, which would have followed a straight line in a Euclidean space, are bent. This effect of light deflection is called gravitational lensing. As can be derived in the stationary weak-field approximation of general relativity (Carroll, 2003), the deflection angle $\hat{\alpha}$, which describes the difference in the direction of a light ray passing

a massive point mass and the direction of the same light ray in the absence of this mass, is given by

$$\hat{\boldsymbol{\alpha}} = \frac{4GM}{c^2}\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|^2}, \tag{2.1}$$

with M being the mass of the point mass, G Newton's constant, c the speed of light and $\boldsymbol{\xi}$ the impact vector, orthogonal to the incoming direction of the light ray and describing its distance from the deflector (compare Fig. 2.1). In the following, we will describe extended deflectors of mass density $\rho(\boldsymbol{r})$ as continuous ensembles of point masses, each of which contains an infinitesimal mass element dm in an infinitesimal volume dV

$$dm = \rho(\boldsymbol{r})dV. \tag{2.2}$$

2.1.1 Thin screen approximation

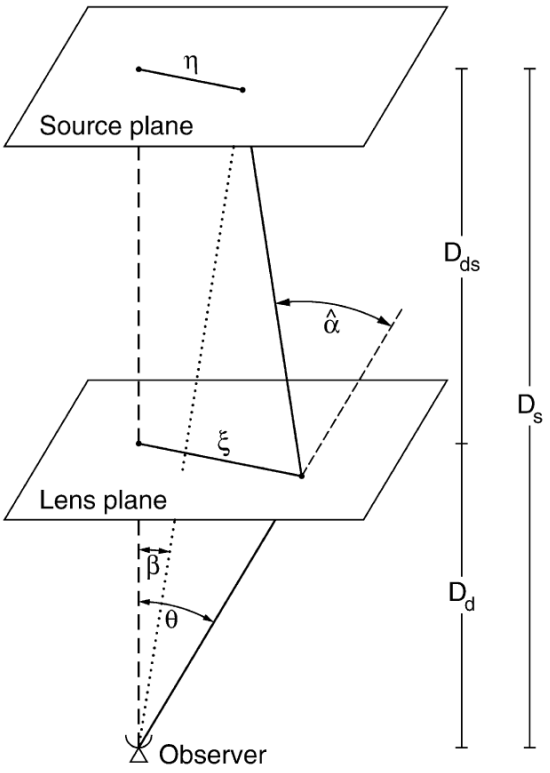


Figure 2.1: From Bartelmann and Schneider (2001). Sketch of a typical gravitational lens system.

We will make an assumption on the geometry of the lensing situation, which is generally fulfilled within the context of lensing by galaxy clusters. We assume, that the lensing happens within a distance along the line-of-sight, which is very small compared to the typical distances between the deflector in the lens plane and the light-emitting source in the source plane D_{ds}, the deflector and the observer D_{d} and between the observer and the source D_{s}. This typical lensing situation is sketched in Fig. 2.1. In other words, we approximate the trajectory of the light ray as a straight line, which is abruptly deflected via Eq. 2.1 within an infinitesimally thin screen, perpendicular to the line-of-sight. This assumption is actually equivalent to the Born approximation in scattering theory. Thus, we project the mass density of the lens along the line-of-sight, which direction we denote as the z coordinate and obtain the surface-mass density Σ as

$$\Sigma(\boldsymbol{\xi}) = \int \rho(\boldsymbol{\xi}, z)dz. \tag{2.3}$$

We derive the total deflection angle by integrating over the total mass distribution

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2}\int d^2\xi' \Sigma(\boldsymbol{\xi}')\frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{|\boldsymbol{\xi}-\boldsymbol{\xi}'|^2}. \tag{2.4}$$

2.1.2 The lens equation

By having a close look at Fig. 2.1, we can read off the mapping, which relates the 2D coordinate of a point source in the source plane $\boldsymbol{\eta}$ to its 2D coordinate in the lens plane $\boldsymbol{\xi}$, after

deflection by the angle $\hat{\boldsymbol{\alpha}}$

$$\boldsymbol{\eta} = \frac{D_\mathrm{s}}{D_\mathrm{d}}\boldsymbol{\xi} - D_\mathrm{ds}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{2.5}$$

It should be noted, that the distances $D_\mathrm{s}, D_\mathrm{d}$ and D_ds have to be defined such, that Eq. 2.5 holds for any spatial geometry of the Universe. Therefore, we will use the angular diameter distance (can e.g. be found in Weinberg, 2008), which satisfies this condition by construction. With this in mind, we introduce the angular source plane $\boldsymbol{\beta} = \boldsymbol{\eta}/D_\mathrm{s}$ and lens plane $\boldsymbol{\theta} = \boldsymbol{\xi}/D_\mathrm{d}$ coordinates and use them to transform Eq. 2.5 into

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\boldsymbol{\alpha}}(D_\mathrm{d}\boldsymbol{\theta}), \tag{2.6}$$

which motivates the definition of the reduced deflection angle $\boldsymbol{\alpha}$, to obtain the final, convenient form of the lens equation

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{2.7}$$

This central equation in gravitational lensing describes the lens mapping from the source plane to the lens plane and we will make constant use of it.

2.1.3 The lensing potential

In adaptation to the practical considerations of observations, we describe lensing with the help of angular coordinates. Thus, it appears useful to define a dimensionless equivalent to the surface-mass density, which we will call the convergence κ. Due to the relation between the reduced deflection angle $\boldsymbol{\alpha}$ and the deflection angle $\hat{\boldsymbol{\alpha}}$, we define it as

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_\mathrm{d}\boldsymbol{\theta})}{\Sigma_\mathrm{cr}}, \tag{2.8}$$

with the critical surface density

$$\Sigma_\mathrm{cr} = \frac{c^2}{4\pi G}\frac{D_\mathrm{s}}{D_\mathrm{d}D_\mathrm{ds}}. \tag{2.9}$$

The reduced deflection angle now reads

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\theta' \kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta}-\boldsymbol{\theta}'}{|\boldsymbol{\theta}-\boldsymbol{\theta}'|^2}, \tag{2.10}$$

which suggests introducing a 2D gravitational potential ψ, which is adapted to the thin-screen approximation and for which

$$\nabla^2\psi = 2\kappa \tag{2.11}$$

holds. It describes a rescaled and line-of-sight integrated analogue of the Newtonian potential. We shall call it lensing potential and formally define it as

$$\psi(\boldsymbol{\theta}) := \frac{1}{\pi}\int d^2\theta' \kappa(\boldsymbol{\theta}')\ln|\boldsymbol{\theta}-\boldsymbol{\theta}'|. \tag{2.12}$$

2.2 Weak lensing

We will now begin to describe the effects of lensing on the observations of astronomical objects, which are in our case distant galaxies with redshifts typically ranging from ~ 0.5 to ~ 3. One aspect of Liouville's[1] theorem, which is interesting to us, is that lensing neither creates nor destroys photons. Translated into the relation between the unlensed surface-brightness distribution of an object $I^{(s)}$ and the observed surface-brightness distribution I and while using the lens equation Eq. 2.7 it reads

$$I(\boldsymbol{\theta}) = I^{(s)}[\boldsymbol{\beta}(\boldsymbol{\theta})]. \tag{2.13}$$

This equation looks rather harmless, but depending on the lensing potential of the deflector it can become highly non-linear. We therefore distinguish different lensing regimes. In the case of so-called weak lensing, we assume that the deflection angles are small and we will locally expand Eq. 2.13 in terms of the angular coordinate $\boldsymbol{\theta}$. Before we do so, we need to conveniently describe the observed surface-brightness distribution, which is why we decompose it into its multipole moments of different order

$$M = \int d^2\theta\ I(\boldsymbol{\theta}) \tag{2.14}$$

$$D_i = \int d^2\theta\ I(\boldsymbol{\theta})\theta_i \qquad i \in [1,2] \tag{2.15}$$

$$Q_{ij} = \int d^2\theta\ I(\boldsymbol{\theta})\theta_i\theta_j \qquad i,j \in [1,2] \tag{2.16}$$

$$Q_{ij...n} = \int d^2\theta\ I(\boldsymbol{\theta})\theta_i...\theta_n \qquad i,j,...n \in [1,2]. \tag{2.17}$$

2.2.1 Complex spin fields

We have already mentioned, that the lens mapping given by Eqs. 2.7, 2.13 can become quite complicated because of its nonlinearity and we will now describe a formulation of the problem, that tremendously simplifies the equations and exploits the symmetries in the structure of the problem.

A special advantage of the decomposition of the observed images into their multipole components is the fact, that the individual elements of the expansion encode symmetries under rotations in the plane, so under transformations represented by the group $SO(2)$[2]. We say, that a quantity has spin s if it is invariant under a rotation of the Cartesian coordinate frame by a rotation angle $\varphi = 2\pi/s$ and $s \in [1,2,...]$. A scalar quantity, which is invariant under any rotation, we say to have spin 0. Since the groups $SO(2)$ and $SU(1)$[3] are at least locally isomorphic[4] (see e.g. Tung, 1985, as a textbook on group theory) we can furthermore introduce the complex mapping

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \rightarrow v_1 + \mathrm{i}v_2 = \left(v_1^2 + v_2^2\right)^{1/2} e^{\mathrm{i}\varphi} \quad \text{with} \quad \varphi = arctan\frac{v_2}{v_1} \quad \text{and} \quad \varphi \in [0, 2\pi)\,^{5}, \tag{2.18}$$

[1] Joseph Liouville (1809-1882) was a French mathematician.
[2] The set of orthogonal 2×2 matrices with a determinant of one
[3] The set of unitary numbers with a determinant of one
[4] Meaning that there exists a bijective map $f : SO(2) \rightarrow U(1)$, such that for f and f^{-1} it holds that $f(u * v) = f(u) \cdot f(v)$, where u, v are elements of $SO(2)$, $*$ the group operation in SO(2) and $\cdot$ the group operation in $U(1)$.
[5] Under this restriction the mapping is isomorphic

for every vector quantity, especially $\boldsymbol{\theta}$ and simplify the notation.

In this complex notation, a quantity is defined to have spin s if it gets multiplied by a phase factor of $exp(-\mathrm{i}s\varphi)$ under rotations of the original Cartesian coordinate frame by an angle of φ. In analogy to algebraic solutions of the angular-momentum problem (Varshalovich and Moskalev, 1988), we define the spin-raising and spin-lowering operators ∂ and ∂^* with respect to the angular directions θ_1 and θ_2 by

$$\partial := (\frac{\partial}{\partial\theta_1} + \mathrm{i}\frac{\partial}{\partial\theta_2}) \qquad \partial^* := (\frac{\partial}{\partial\theta_1} - \mathrm{i}\frac{\partial}{\partial\theta_2}), \tag{2.19}$$

which increase or decrease the spin of a complex quantity by one, respectively. Following the ideas above, we decompose the symmetric[6] Cartesian tensors of rank 2, 3 and 4 into complex quantities, which we will call fields, of spin 0, 1, 2, 3 and 4 defined as

$$Q_0 := Q_{11} + Q_{22} \tag{2.20}$$
$$Q_2 := Q_{11} - Q_{22} + 2\mathrm{i}Q_{12} \tag{2.21}$$

$$T_1 := Q_{111} + Q_{122} + \mathrm{i}(Q_{122} + Q_{222}) \tag{2.22}$$
$$T_3 := Q_{111} - 3Q_{122} + \mathrm{i}(3Q_{122} - Q_{222}) \tag{2.23}$$

$$\tilde{F}_0 := Q_{1111} + 2Q_{1122} + Q_{2222} \tag{2.24}$$
$$\tilde{F}_2 := Q_{1111} - Q_{2222} + 2\mathrm{i}(Q_{1112} + Q_{1222}) \tag{2.25}$$
$$\tilde{F}_4 := Q_{1111} - 6Q_{1122} + Q_{2222} + 4\mathrm{i}(Q_{1112} - Q_{1222}). \tag{2.26}$$

Finally, some simple arithmetic rules for calculations with spin fields as given in Schneider and Er (2008):

1. Terms in a valid equation must have the same spin.
2. The product of a spin-m with a spin-n quantity has spin $m + n$.
3. Complex conjugation changes the sign of the spin.

While locally expanding Eq. 2.7, with each additional order in θ more and more spatial complexity is subsequently added to the lens mapping. We should expect that to be reflected in the spin properties of the mapping and since the lensing potential is the underlying scalar quantity, causing the deflection in Eq. 2.7, we define the following five spin fields, by applying the spin-raising and spin-lowering operators.

$$\alpha := \partial\psi \qquad s = 1 \tag{2.27}$$
$$2\gamma := \partial\partial\psi \qquad s = 2 \tag{2.28}$$
$$2\kappa := \partial\partial^*\psi \qquad s = 0 \tag{2.29}$$
$$2F := \partial\partial\partial^*\psi \qquad s = 1 \tag{2.30}$$
$$2G := \partial\partial\partial\psi \qquad s = 3 \tag{2.31}$$

We remark that the quantities in Eqs. 2.27, 2.29 are the complex equivalents to the reduced deflection angle and to the convergence. The other quantities we call from top to bottom,

[6] This means that a tensor is invariant under index permutations

the shear, the F-flexion and the G-flexion. The second order expansion of the lens Eq. 2.7 with respect to the position angle and with the use of our complex notation then reads

$$\beta = (1-\kappa)\theta - \gamma\theta^* - \frac{1}{4}F^*\theta^2 - \frac{1}{2}F\theta\theta^* - \frac{1}{4}G\,(\theta^*)^2 + \mathcal{O}(\theta^3). \tag{2.32}$$

This equation is the central result of this section. The locally expanded lens equation decomposes the lens mapping into quantities with simple spin properties, derived from the lensing potential of the deflector. If we define observable quantities through Eqs. 2.20 - 2.26, we might constrain the deflector through its spin fields (Eq. 2.27 - 2.31), by analysing the lens mapping given by Eq. 2.32.

2.2.2 Lens mapping of extended sources

In the following derivation, we consider a single background source with a given intrinsic brightness distribution in the source plane $I^s(\beta)$ and an observed brightness distribution in the image or lens plane $I(\theta)$. We chose our coordinate frame such, that the centre-of-light in the image plane coincides with its origin $\theta_0 = 0$

$$\int d^2\theta\ \theta\ I(\theta) = 0. \tag{2.33}$$

Following the derivations given in Schneider and Er (2008); Okura et al. (2007) or Goldberg and Leonard (2007), we start from the linearised Jacobian determinant of Eq. 2.32

$$\begin{aligned}\det\mathcal{A}(\theta) = \left|\frac{\partial\beta}{\partial\theta}\right| = (1-\kappa)^2 - \gamma\gamma^* - \theta\left[(1-\kappa)F^* + \frac{\gamma^*F + \gamma G^*}{2}\right] \\ - \theta^*\left[(1-\kappa)F + \frac{\gamma^*G + \gamma F^*}{2}\right] + \mathcal{O}(\theta^2),\end{aligned} \tag{2.34}$$

and use it to calculate the relations between the brightness distribution in the source and in the lens plane. We separate the calculations by multipole order and follow the notation of Schneider and Er (2008).

Total fluxes (monopole)

While defining the intrinsic, unlensed flux of the source as S_0, one finds the following relation to the observed flux S, with the help of Eq. 2.33

$$\begin{aligned}S_0 &= \int d^2\beta\ I^s(\beta) = \int d^2\theta\ \det\mathcal{A}(\theta)\ I(\theta) = \left[(1-\kappa)^2 - \gamma\gamma^*\right]\int d^2\theta\ I(\theta) \\ &= \det\mathcal{A}(\theta_0)S.\end{aligned} \tag{2.35}$$

Centroid shift (dipole)

Another interesting effect of gravitational lensing is the shift of the centre-of-light, which is defined in the source plane as

$$\bar{\beta} = \frac{1}{S_0}\int d^2\beta\ \beta\ I^s(\beta). \tag{2.36}$$

Again, we make use of the lens Eq. 2.32 and its Jacobian determinant given by Eq. 2.34 to obtain

$$\bar{\beta} = \frac{1}{S\left[(1-\kappa)^2 - \gamma\gamma^*\right]} \int d^2\theta \left[(1-\kappa)\theta - \gamma\theta^* - \frac{1}{4}F^*\theta^2 - \frac{1}{2}F\theta\theta^* - \frac{1}{4}G\left(\theta^*\right)^2\right] \times \left\{(1-\kappa)^2 - \gamma\gamma^* - \theta\left[(1-\kappa)F^* + \frac{\gamma^*F + \gamma G^*}{2}\right] - \theta^*\left[(1-\kappa)F + \frac{\gamma^*G + \gamma F^*}{2}\right]\right\} I(\theta) \tag{2.37}$$

Obviously, already for the calculation of this first order moment, the formulas become quite cumbersome, but by using the complex second-order brightness moments

$$Q_0 \equiv \frac{1}{S}\int d^2\theta\ \theta\ \theta^*\ I(\theta) \qquad \text{and} \qquad Q_2 \equiv \frac{1}{S}\int d^2\theta\ \theta^2\ I(\theta), \tag{2.38}$$

which have the spin properties that we described in the last section, one can show that Eq. 2.37 becomes a simple matrix multiplication

$$\bar{\beta} = \mathrm{BG}. \tag{2.39}$$

In the last equation the matrix B depends on Q_0 and Q_2 and the matrix G depends on the convergence, the shear and the two flexion fields.

The mass-sheet degeneracy

Before we can investigate the more interesting behaviour of higher-order brightness moments under the lens mapping, we have to deal with a peculiar ambiguity in gravitational lensing . Gorenstein et al. (1988) were the first to point out, that the effects of lensing are unchanged while adding a sheet of constant surface-mass density. Therefore, the convergence of the deflector can be transformed like

$$\kappa(\theta) \rightarrow \kappa'(\theta) = \lambda\kappa(\theta) + (1-\lambda). \tag{2.40}$$

Eq. 2.40 is called the mass-sheet degeneracy . Because of this degeneracy, the shear alone is not a weak-lensing observable but only the reduced shear defined as

$$g := \frac{\gamma}{(1-\kappa)}, \tag{2.41}$$

which is invariant under the mass-sheet transformation. We also introduce the reduced flexion as derivatives of the reduced shear

$$G_1 := \partial^* g = \frac{F + gF^*}{(1-\kappa)}, \qquad G_3 := \partial g = \frac{G + gF}{(1-\kappa)}, \tag{2.42}$$

and isotropically rescale the lens equation to obtain

$$\hat{\beta} = \frac{\beta}{(1-\kappa)} = \theta - g\theta^* - \frac{1}{4}\frac{F^*}{(1-\kappa)}\theta^2 - \frac{1}{2}\frac{F}{(1-\kappa)}\theta\theta^* - \frac{1}{4}\frac{G}{(1-\kappa)}\left(\theta^*\right)^2. \tag{2.43}$$

The last step ensures invariance of the lens mapping under the transformation given by Eq. 2.40. The Jacobian determinant of the lens mapping is then given by

$$\det\hat{\mathcal{A}} = \frac{\det\mathcal{A}}{(1-\kappa)^2} = 1 - gg^* - \eta^*\theta - \eta\theta^*, \tag{2.44}$$

where η is related to the reduced flexion as follows

$$\eta = G_1 - \frac{gG_1^*}{2} + \frac{g^* G_3}{2}. \tag{2.45}$$

Ellipticities (quadrupole)
Now we can have a look at the transformation of the spin-2 brightness moment of the source, given by

$$\begin{aligned} Q_2^s &= \frac{1}{S_0}\int d^2\hat{\beta}\,(\hat{\beta}-\bar{\beta})^2\,I^s(\hat{\beta}) = \frac{1}{S(1-gg^*)}\int d^2\theta\,(\hat{\beta}(\theta)-\bar{\beta})^2\,\det\hat{\mathcal{A}}(\theta)\,I(\theta) \\ &= \frac{1}{S(1-gg^*)}\int d^2\theta\left(\theta - g\theta^* - \frac{1}{4}\frac{F^*}{(1-\kappa)}\theta^2 - \frac{1}{2}\frac{F}{(1-\kappa)}\theta\theta^* - \frac{1}{4}\frac{G}{(1-\kappa)}(\theta^*)^2 - \bar{\beta}\right)^2 \\ &\times (1-gg^*-\eta^*\theta-\eta\theta^*)\,I(\theta), \end{aligned} \tag{2.46}$$

which can be represented in a more clear form by using the spin-1 and 3 third-order brightness moments

$$T_1 \equiv \frac{1}{S}\int d^2\theta\,\theta^2\theta^*\,I(\theta), \qquad T_3 \equiv \frac{1}{S}\int d^2\theta\,\theta^3\,I(\theta). \tag{2.47}$$

We find

$$Q_2^s = Q_2 - 2gQ_0 + g^2Q_2^* + \mathsf{AG} - \bar{\beta}^2, \tag{2.48}$$

where A is a matrix, whose elements only depend on T_1, T_3 and g. The reader is referred to Schneider and Er (2008) for more details.

Higher-order moments
In order to obtain also an estimate for the flexion of a lens, we need to evaluate the transformation of third-order brightness moments given by

$$T_1^s = \frac{1}{S_0}\int d^2\hat{\beta}\,(\hat{\beta}-\bar{\beta})^2(\hat{\beta}^*-\bar{\beta}^*)\,I^s(\hat{\beta}), \tag{2.49}$$

and

$$T_3^s = \frac{1}{S_0}\int d^2\hat{\beta}\,(\hat{\beta}-\bar{\beta})^3\,I^s(\hat{\beta}). \tag{2.50}$$

We use the spin 0, 2 and 4 forth-order brightness moments

$$\tilde{F}_0 \equiv \frac{1}{S}\int d^2\theta\,(\theta\theta^*)^2\,I(\theta), \qquad \tilde{F}_2 \equiv \frac{1}{S}\int d^2\theta\,\theta^3\theta^*\,I(\theta), \qquad \tilde{F}_4 \equiv \frac{1}{S}\int d^2\theta\,\theta^4\,I(\theta), \tag{2.51}$$

to obtain

$$\mathcal{T}^s = \tau + \mathsf{CG} + \mathcal{O}(\bar{\beta}^3). \tag{2.52}$$

This heavily simplified equation contains the third-order brightness moments T_1^s and T_3^s in the vector of length four $\mathcal{T}^s$. The respective observed third order-moments are truncated into τ and the 4×4 matrix C contains fourth and second order brightness moments. Again, the explicit representation of those quantities can be found in Schneider and Er (2008).

Shear and flexion estimators
With Eqs. 2.48, 2.52 we have closed the system of equations to derive estimators for shear and flexion of the lens. By neglecting the higher-order terms in the centroid shift, we invert Eq. 2.52, which is linear in G, to obtain

$$\mathsf{G} = \mathsf{C}^{-1}(\mathcal{T}^s - \tau). \tag{2.53}$$

Inserting the equation above into Eq. 2.48 yields

$$Q_2^s = Q_2 - 2gQ_0 + g^2 Q_2^* + \mathsf{A}\mathsf{C}^{-1}(\mathcal{T}^s - \tau) - (\mathsf{B}\mathsf{G})^2. \tag{2.54}$$

We will now assume, that sources are randomly oriented, which means that while averaging over a number of background sources, the expectation values of Q_2^s and $\mathcal{T}^s$ vanish. Thus, we are left with

$$Q_2 - 2gQ_0 + g^2 Q_2^* = \mathsf{A}\mathsf{C}^{-1}\tau + (\mathsf{B}\mathsf{C}^{-1}\tau)^2 =: Y(g), \tag{2.55}$$

which is an equation where the reduced shear just depends on the observed brightness moments. Schneider and Er (2008) suggested to solve it in an iterative way, since the right-hand side is small compared to the left-hand side. Hence, the solutions are implicitly given by

$$g = \frac{\chi}{|\chi|^2}\left(1 \pm \sqrt{1 - |\chi|^2 + \frac{Y\chi^*}{Q_0}}\right), \tag{2.56}$$

where we have defined the complex ellipticity as

$$\chi := \frac{Q_2}{Q_0}. \tag{2.57}$$

In the case of a weak shear signal, this reduces to the well-known result as e.g. found by Schneider and Seitz (1995), which relates complex ellipticity and reduced shear

$$g \approx \frac{\chi}{2}. \tag{2.58}$$

To obtain an estimator for the flexion, one considers Eq. 2.53 and assumes e.g. $|g| \ll 1$, which simplifies the appearance of C significantly and the distinction between flexion and reduced flexion vanishes. Goldberg and Leonard (2007) and Schneider and Er (2008) find

$$G_1 \approx F \approx \frac{4}{9\tilde{F}_0 - 12Q_0^2}T_1, \qquad G_3 \approx G \approx \frac{4}{3\tilde{F}_0}T_3. \tag{2.59}$$

One last thing, that we have to take into account, is the fact that the redshift of the source enters into the convergence of the deflector, via the scaling with the critical surface-mass density in Eq. 2.9. Bartelmann and Schneider (2001) define therefore the cosmological weight function $Z(z)$ for a given lens redshift z_d by

$$Z(z) := \frac{\lim_{z\to\infty}\Sigma_{cr}(z_\mathsf{d}, z)}{\Sigma_{cr}(z_\mathsf{d}, z)} H(z - z_\mathsf{d}) = \frac{D_\infty D_{\mathsf{d}s}}{D_{\mathsf{d}\infty} D_\mathsf{s}} H(z - z_\mathsf{d}), \tag{2.60}$$

where the Heavyside step function H ensures, that sources in front of the lens are weighted with zero. If we consider sources at different redshifts, all relevant lensing quantities of the deflector, like deflection angle, convergence, shear and flexion scale with the cosmological weight factor: $\alpha(z) = Z(z)\alpha$, $\kappa(z) = Z(z)\kappa$, $\gamma(z) = Z(z)\gamma$, $F(z) = Z(z)F$, $G(z) = Z(z)G$ where α, κ, γ, F, and G, refer to a fiducial source redshift of $z = \infty$. We obtain the final estimators

$$g(z) = \frac{Z(z)\gamma}{(1 - Z(z)\kappa)} = \left\langle \frac{Q_2(z)}{2Q_0(z)} \right\rangle \tag{2.61}$$

$$G_1(z) = \frac{Z(z)F + g(z)Z(z)F^*}{(1 - Z(z)\kappa)} \approx Z(z)F = \left\langle \frac{4T_1(z)}{9F_0(z) - 12Q_0(z)^2} \right\rangle \tag{2.62}$$

$$G_3(z) = \frac{Z(z)G + g(z)Z(z)F}{(1 - Z(z)\kappa)} \approx Z(z)G = \left\langle \frac{4T_3(z)}{3F_0(z)} \right\rangle, \tag{2.63}$$

where the brackets $< \cdot >$ indicate, that we are considering expectation values after averaging over a number of images.

2.2.3 Shear and flexion in reality

The last section assumed, that images are observed under ideal conditions, so without any systematic effects introduced by the observational process. Unfortunately, this does not refer to reality as it is shown in Fig. 2.3. In real astronomical images one has to deal with several instrumental noises like read-out shot-noise and pixelisation of the detector, sky background and the blurring of images caused by the atmosphere[7]. As a result, one cannot directly observe those brightness moments of an image, which are only due to the intrinsic shape of an object and to the shape distortions caused by the deflector.

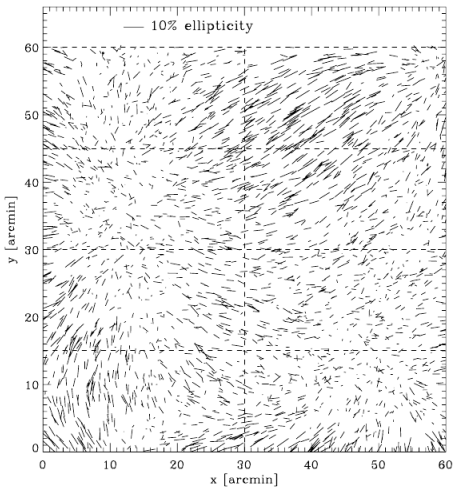


Figure 2.2: From Rowe (2010). A so-called whisker plot, showing the measured ellipticities of a simulated stellar field containing 2500 stars. The amplitude of the ellipticities is indicated by the length of the whiskers and its orientation by their direction.

It has been realised, that reliably measuring shear from real observational surveys like the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) (Fu et al., 2008; Semboloni et al., 2006) or the COSMOS field (Massey et al., 2007; Schrabback et al., 2009) is more challenging than it was expected initially. Therefore, a series of test programs have been established, to investigate the performance of existing shape-measurement pipelines. This includes the Shear TEsting Program (STEP) (Heymans et al., 2006; Massey et al., 2007) and the GRavitational lEnsing Accuracy Testing 2008 (GREAT08) (Bridle et al., 2008). With the help of those programs, the performance of the methods, that we will present further on, in measuring shapes from realistic images, improved significantly.

In order to get an estimate for the distortions, which are not induced by lensing but by the optics of the telescopes and the seeing, one uses reference stars, which can be assumed to be ideal point sources. To get an idea of the spatial variation of this effect over the whole field-of-view of the telescope, one observes stellar fields and measures the distortions for individual stars as indicated in Fig. 2.2. From these measurements the so-called point-spread function (PSF) of the telescope can be modeled (Rowe, 2010) and subtracted from the shape measurement. How this is done in practice varies within the different methods.

Weighted multipole moments

As can be seen from Fig. 2.4, the noise level contained in a realistic image of a galaxy can hamper the measurement of the multipole moments of this object in areas with low signal-to-noise ratio. For this reason, a weighting function $W(\theta)$ is introduced in the determination of the different moments

$$Q_{ij\ldots n} = \int d^2\theta \, W(\theta) \, I(\theta) \, \theta_i \ldots \theta_n \qquad i, j, \ldots n \in [1, 2]. \tag{2.64}$$

[7] In the literature often called seeing.

To account for the fact that the relations derived in Sec. 2.2.2 are not valid any more due to the weighting function and the additional convolution with the PSF, a method originally developed by Kaiser et al. (1995) and commonly called KSB corrects for those effects by assuming that they can be described as a convolution with a narrow, but highly anisotropic and a broad, but isotropic kernel. These assumptions do not necessarily hold, especially when describing complicated PSF's. Anyway, KSB and its successor called KSB+ (Luppino and Kaiser, 1997; Hoekstra et al., 1998; Heymans et al., 2005) have proven, that a simple correction to the measured ellipticity, described via the moments up to the fourth order of the observed brightness-distribution, the weighting function and its spatial derivatives up to the second order, provides reasonable estimates for the reduced shear. This approach has been generalised by Okura et al. (2007); Goldberg and Leonard (2007) and Okura et al. (2008) to provide the possibility to measure a flexion signal, where the correction terms include moments of the brightness distribution up to the sixth order and spatial derivatives of the weighting function up to the third order.

Shapelets

A completely different approach is followed when the objects, which shapes we want to measure, are not described by their multipole moments but when they are decomposed into another orthogonal set of basis functions called shapelets. This set of basis functions is identical to the Eigenfunctions of the quantum harmonical oscillator and can be represented in Cartesian Refregier (2003) or polar coordinates (Massey and Refregier, 2005). Due to the convenient mathematical properties of this basis, typical image transformations like convolution and deconvolution or shearing operations can be performed particularly fast and easily in shapelet space. Therefore, the subtraction of the convolution kernel, which also has to be modeled by shapelets, from the object of interest is possible and the shape of an object can directly be measured from the deconvolved shapelet representation. This, so-called 'passive' approach, where the measurement is directly obtained from the observed object, is followed by Refregier (2003). To perform this measurement in an 'active' way, the methods by Kuijken (2006) and Bernstein and Jarvis (2002) create an initial, unsheared model, convolve it with PSF and weighting function and subsequently apply shear to it, until the model matches best with the image in a least-squares sense. Shapelet methods are in principle also able to conveniently measure flexion signals and do not suffer from the strong assumptions of the KSB approach, but Melchior et al. (2010) pointed out, that shapelet models are not ideal to describe galaxies, whose shapes differ too strongly from the typical Gaussian character of the basis function. In those cases, measurements based on shapelets tend to underestimate the shear. Nevertheless, shapelets have several interesting astronomical applications and an overview is presented in Melchior (2010).

Model-fitting techniques

Another class of 'active' methods, are direct model-fitting techniques. A very generic model to describe the shape of a galaxy like an elliptical Sérsic model (Sérsic, 1963) is convolved with a given PSF model and then compared with the original image. After a best fit is found, the shear estimate is just given by the ellipticity of the analytic model. The biggest problem with this approach is the slow model selection process, since several model parameters have to be fit. The most developed tool available right now is called LensFit (Miller et al., 2007; Kitching et al., 2008), which uses a Bayesian technique for model selection. Parameter marginalisation can be used, since we are only interested in the ellipticity measurement and a self-calibrated prior helps to maximise the posterior probability of the model. With all these techniques, LensFit is able to find a best-fit model in less than one second per galaxy image.

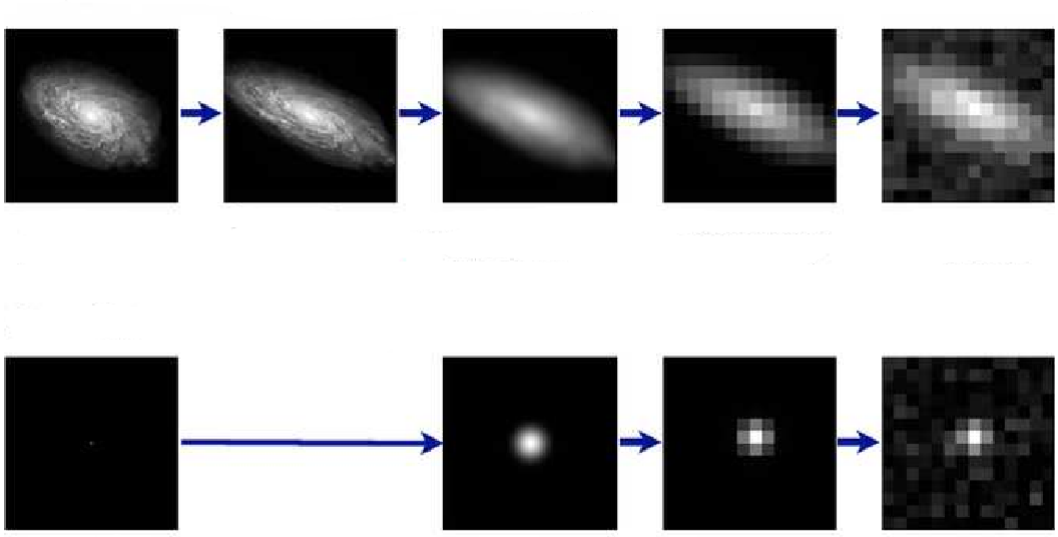


Figure 2.3: From Bridle et al. (2008). The *upper panel* shows the 'forward' process of gravitational lensing. A galaxy is sheared by the gravitational tidal and its image is convolved afterwards with the PSF. The pixelised detector output contains additionally a significant amount of noise. The *lower panel* shows the same effects on the image of a star, with the difference that it is not lensed.

A new class of model-fitting techniques was developed in the course of the GREAT08 challenge. The method by Lewis (2009) does not fit an elliptical model to a single image, but to a number of stacked images, thereby reducing the noise contamination. While this method still suffered from the fact that the centroid positions of single images has to be known quite precisely, Hosseini and Bethge (2009) cured this issue by stacking power spectra in Fourier space and comparing those to the elliptical model. It still has to be quantified how applicable those stacking techniques are in real observational applications, but they delivered the best results in the GREAT08 challenge (Bridle et al., 2009). The main drawback of the model-fitting techniques is the fact, that to measure a flexion signal, the model complexity has to be increased compared to simple elliptical models, which gives rise to the question if the runtime of such a method is still feasible.

2.3 Strong lensing

Let us now consider the case where image distortions due to lensing cannot be described any more with a linearised lens equation. In this so-called strong-lensing regime, image distortions are not restricted to slight deformations described by shear or flexion, but more spectacular effects like giant arcs, Einstein rings or multiply imaged sources are observed. As a rule of thumb one can say, that those effects only appear near the core of lenses, for which $\kappa > 1$ is true somewhere in the observed field but exceptions to this rule are possible. We will give a more stringent definition of the strong-lensing regime later on. An outstanding example of a gravitational lens producing hundreds of strong-lensing features is the Golden Lens, the galaxy cluster Abell 1689, situated at a redshift of $z_c = 0.181$. A composite image

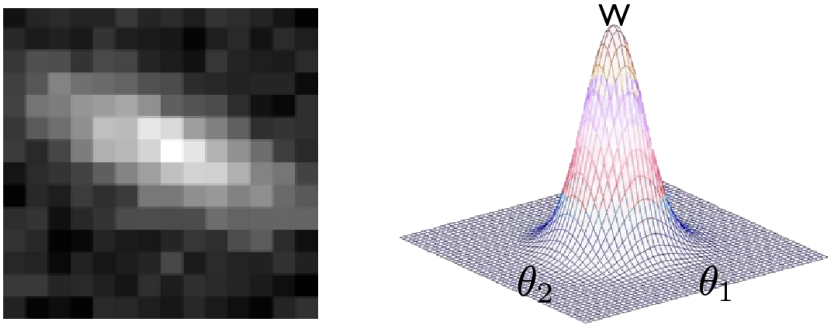


Figure 2.4: This figure, which was kindly provided by Catherine Heymans, shows the necessity to introduce an appropriate weighting function, in order to measure brightness moments, which are not contaminated by noise.

of `Abell 1689` is shown in Fig. 2.5.

2.3.1 Critical lens mapping

We go back to the lens Eq. 2.7 to quantitatively describe the strong-lensing regime, which is of course more difficult since the local expansion of the lens equation does not hold any more. We have already seen, that the lens mapping is described by its Jacobian matrix $\mathcal{A}$ defined as

$$\mathcal{A} = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}}, \qquad (2.65)$$

and an observed solid-angle element of an image $\delta\theta^2$ is related to the according unlensed solid-angle element $\delta\beta^2$ via the Jacobian determinant (Narayan and Bartelmann, 1996)

$$\mu := \frac{1}{\det\mathcal{A}} = \frac{\delta\theta^2}{\delta\beta^2}, \qquad (2.66)$$

where we have defined the magnification μ. Combining Eq. 2.65 with Eq. 2.7 gives

$$\det\mathcal{A} = (1-\kappa)^2 - \gamma^2, \qquad (2.67)$$

where κ and γ are generally functions of the angular coordinate θ. As one can see from Eq. 2.67 the lens mapping can become singular. At this point our initial assumptions on the lens configuration from Sec. 2.1.1 fail.

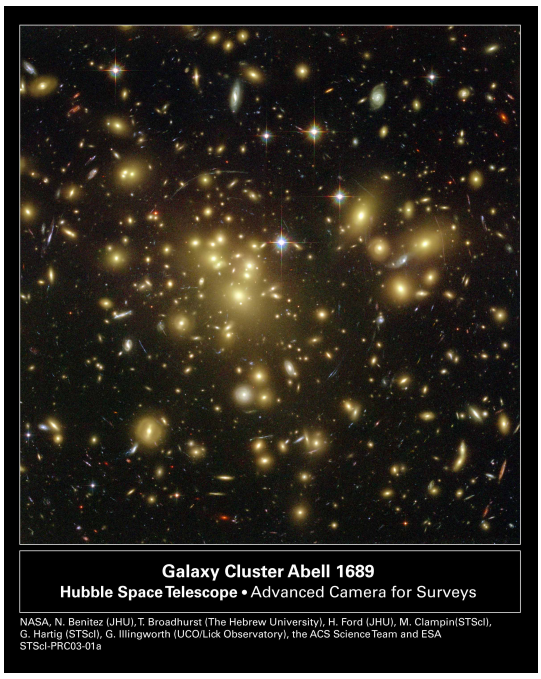


Figure 2.5: A composite image of the spectacular strong-lensing cluster `Abell 1689`. One can clearly identify several strong-lensing features like arcs and arclets in the centre of the image, which make the geometry of the dominating dark components of the lens almost visible by eye.

Sets of points, for which $\det\mathcal{A} = 0$ form closed curves in the lens plane and we call them critical curves or sometimes also critical lines. Their original positions in the source plane are called caustics. Images located near a critical curve appear strongly distorted and give rise to the term of a gravitational arc . Sources within a caustic are multiply imaged. Those effects, both related to the critical lines of a cluster are loosely called strong-lensing features and we will therefore define a cluster as strongly lensing if it produces at least one critical line .

2.3.2 A spherical model

In the following, we will assume lens models to describe typical observations in the strong-lensing regime. We start with the very simple case of a point source lensed through a circularly symmetric lens with otherwise arbitrary mass profile. In this case, the lens equation reads

$$\beta(\theta) = \theta - \frac{D_{\mathsf{ds}}}{D_{\mathsf{d}} D_{\mathsf{s}}} \frac{4GM(\theta)}{c^2\theta}, \tag{2.68}$$

where θ and β are just the distance from the origin in lens or source plane, respectively. For a source, sitting in the origin of the source plane $\beta = 0$ we find

$$\theta = \frac{D_{\mathsf{ds}}}{D_{\mathsf{d}} D_{\mathsf{s}}} \frac{4GM(\theta)}{c^2\theta}, \tag{2.69}$$

with the solution being a ring with radius

$$\theta_{\mathsf{E}} = \sqrt{\frac{4GM(\theta_{\mathsf{E}})}{c^2} \frac{D_{\mathsf{ds}}}{D_{\mathsf{d}} D_{\mathsf{s}}}}. \tag{2.70}$$

This special case, where a nearly ideal point source is lensed into the image of a ring with radius θ_{E} is called an Einstein ring and its radius is called the Einstein radius. This simple example is quite instructive for two reasons. First, it is observed in reality (see e.g. Gavazzi et al., 2008) and second it provides a natural scale for a lens system. The angular separation between multiple image systems is typically given by $2\theta_{\mathsf{E}}$ and images, which approach the Einstein radius of a lens are typically experiencing strong distortions.

A typical spherically symmetric lens model is derived by assuming a so-called Singular Isothermal Sphere (SIS) as the underlying mass distribution of the lens. In this model it is assumed that the 'mass elements' like stars or galaxies within the lens behave as particles of an ideal gas, bound by the gravitational potential of the whole structure. By assuming thermal and hydrostatic equilibrium , we can relate the density profile $\rho(r)$ of the structure to the velocity dispersion σ_{v} of its members (Binney and Tremaine, 2008)

$$\rho(r) = \frac{\sigma_{\mathsf{v}}^2}{2\pi G} \frac{1}{r^2}, \tag{2.71}$$

where r denotes the distance from the centre of mass. Projection along the line-of-sight gives the surface-mass density with the familiar impact parameter ξ

$$\Sigma(\xi) = \frac{\sigma_{\mathsf{v}}^2}{2G} \frac{1}{\xi}. \tag{2.72}$$

A peculiar feature of the SIS profile is the fact, that the absolute value of the deflection angle is equal to the Einstein radius of the lens and points towards the centre of the distribution due to the symmetry of the system

$$\theta_{\mathsf{E}} = 4\pi \frac{\sigma_{\mathsf{v}}^2}{c^2} \frac{D_{\mathsf{ds}}}{D_{\mathsf{s}}} = |\boldsymbol{\alpha}|. \tag{2.73}$$

The lensing potential is simply given by

$$\psi = \theta_{\mathsf{E}} |\boldsymbol{\theta}|. \tag{2.74}$$

2.3.3 Observables in elliptical models

We will drop now the assumption on circular symmetry and move towards elliptical lens models. Blandford and Kochanek (1987) suggested to introduce effective, elliptical lensing potentials by performing the substitution

$$|\boldsymbol{\theta}| \rightarrow \sqrt{\frac{\theta_1^2}{(1-\varepsilon)} + (1-\varepsilon)\theta_2^2} \tag{2.75}$$

in Eq. 2.74. The parameter $\varepsilon = 1 - b/a$ describes the ellipticity of the lens via the ratio of its minor and major axis. The derivation of all relevant lensing quantities, especially the deflection angle is then straight-forward (see e.g. Meneghetti, 2006). More interesting is the effect of such an elliptical lens on extended sources, as it is shown in Fig. 2.6. The following effects should be noted:

- An extended circular source, enclosed by the two caustics of the elliptical lens is imaged five times, including a faint image in the centre of the lens plane.
- The same source close to the inner caustic of the lens gives rise to an arc, which is tangentially oriented with respect to the centre of the lens.
- The same source enclosed by only one caustic of the lens is imaged three times, including a faint central image.
- A source crossing the outer caustic gives rise to an arc, which is radially oriented with respect to the centre of the lens.
- A source near a so-called cusp of the caustic, is imaged into a giant arc, whose saddle points in the light distribution cross the critical lines of the lens.

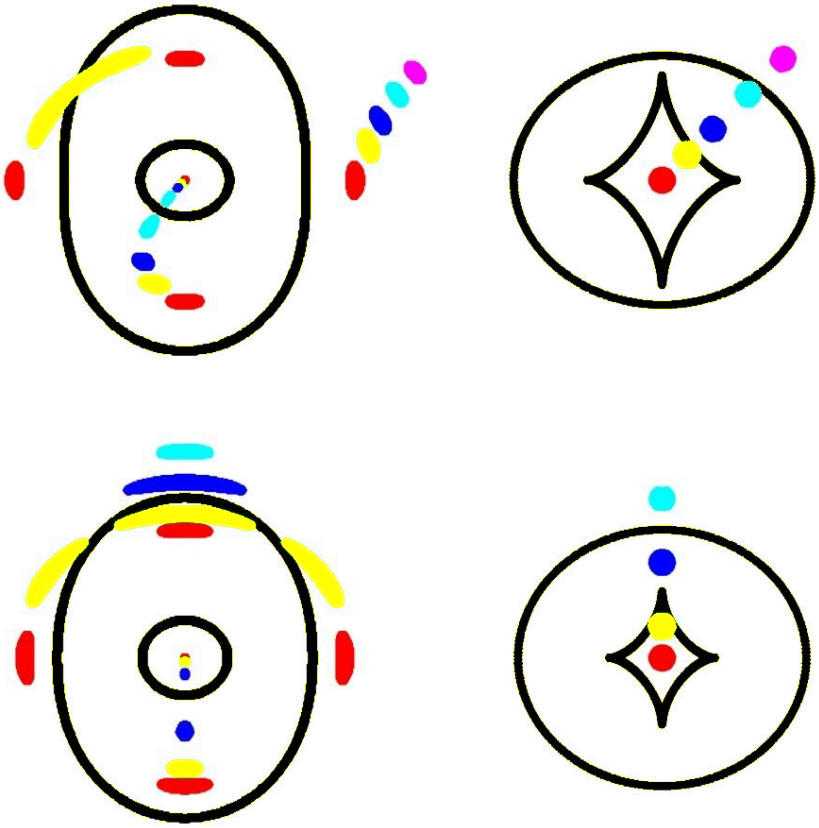


Figure 2.6: From Narayan and Bartelmann (1996), showing a lens system containing an elliptical lens. The *upper panel* shows a sequence of spherical sources crossing the smooth, so-called fold, part of the two caustics of the lens. The *lower panel* shows the crossing along the cusp part of the caustic.

3 Clusters of Galaxies

It is not obvious how to define the concrete entity of a galaxy cluster. Originally recognised as significant concentrations of galaxies, observed in the optical band (e.g. Abell, 1958; Abell et al., 1989), the picture has changed dramatically with observations in other wave bands and with the advent of the cosmological standard model and its description of structure formation in the Universe. Observations in the X-ray bands reveal, that only a very limited number of the baryons[1] in a galaxy cluster are constituted by stars and galaxies, but that most of them reside in a large, diffuse gas halo surrounding the centre of the cluster. In fact, also this so-called Intra-Cluster-Medium (ICM) constitutes only 10 - 20 % of the total mass of a galaxy cluster. Most of its mass is provided by a huge dark-matter halo with the result that the total mass of the most massive clusters can exceed $10^{15}\ M_{\odot}$.

The definition of a galaxy cluster should therefore start from the predictions of a cosmological model of structure formation. We have already seen, that the initial perturbations in the smooth cosmological background are presumably seeded by inflation. Due to gravitational attraction, those perturbations grow and might decouple from the Hubble flow, due to gravitational instabilities. This formation of gravitationally bound structures happens on very small scales first, amplifies and creates larger and larger objects due to an hierarchical merging process (see e.g. Bond et al., 1991; Lacey and Cole, 1993; Somerville and Kolatt, 1999). Galaxy clusters are the final result of this process, defining them as the high-mass tail of structures which decoupled from the Hubble expansion. Due to their formation via subsequent merging, clusters are relatively young structures, accessible to observations and encoding the full formation history of a structure within the cosmological context. The observed velocity dispersions of cluster members are of order $\sigma_{\mathrm{v}} \approx 1000$ km s^{-1}(Zwicky, 1933; Rines et al., 2010) and the typical scales of cluster are of order $R \approx 5$ Mpc. The crossing time of a cluster member defined as

$$t_{\mathrm{cr}} = \frac{R}{\sigma_{\mathrm{v}}} \simeq 1 \left(\frac{1}{1\ \mathrm{Mpc}}\right) \left(\frac{\sigma_{\mathrm{v}}}{10^3\ \mathrm{km\ s^{-1}}}\right)^{-1} \mathrm{Gyr} \tag{3.1}$$

is therefore shorter than the Hubble time $1/H_0 \approx 10\ h^{-1}$ Gyr, which implies that clusters still have enough time to reach a virialised state. We will have a closer look into that formation process in Sec. 3.1.

Due to their extremely high mass and deep potential wells, clusters do not lose their content to the environment. This is true for the dark-matter component, which is dominated by gravity and is the most significant component on scales > 1Mpc, but this is especially true for the content of baryonic gas inside a cluster, which starts playing an important role on scales < 1Mpc. Usually, this baryonic component makes it difficult to describe the formation and evolution of a structure theoretically since the underlying physics is much more complicated than the pure gravitational interaction of dark matter. This problem is less severe in galaxy clusters, since baryonic physics seems to be less important to describe the properties of an

[1] In Astrophysics, the term baryons stands for the particles described by the standard model of particle physics. Within this standard model, where the term baryons denotes noninteger-spin hadrons, the context of this term is wrong.

object, as massive as a cluster (Duffy et al., 2010). Nevertheless, the baryons are present and obviously observed in different wave bands as we will see in Sec. 3.2. Those peculiar aspects of cluster physics defines them as ideal laboratories to compare the predictions of the cosmological standard model to observations, as we will see in Sec. 3.3. Galaxy clusters define in a sense the border line between the cosmological, statistical description of large-scale structure, dominated by dark matter and dark energy and individual astrophysical objects, strongly affected by the physics of baryonic matter and therefore directly visible to observations. For a review on the interplay between those cosmological and astrophysical considerations, we recommend Voit (2005).

3.1 Cosmological structure formation

A particular interesting field in Cosmology is the formation of structures in the Universe. A number of excellent textbooks exist on the topic and we would like to mention Padmanabhan (1993) and Liddle and Lyth (2000). Structure formation aims to describe the evolution of structures from the initial, smooth background fixed by inflation.

It is useful to divide this description into an homogeneous part since observations of the Cosmic Microwave Background (CMB) suggest that the Universe is fantastically isotropic on large scales and into an inhomogeneous part, which describes the deviations from this homogeneity on smaller scales. The description of these scales is naturally important, due to the anisotropic character of our direct neighbourhood.

3.1.1 The homogeneous Universe

While thinking about the Universe as a whole, like in the case of cosmological considerations, gravity is the only relevant interaction in this picture. By now, there is no significant reason to believe that Einstein's general relativity is not the theory describing gravity. The description of an homogeneous Universe is therefore obtained by inserting a homogeneous and isotropic metric into the field equations of general relativity

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{3.2}$$

where the Einstein tensor $G_{\mu\nu}$ relates the space-time metric g to the content of the Universe through the energy-momentum tensor $T_{\mu\nu}$ and Newton's constant G. A metric describing such a Universe is given by the Robertson-Walker metric[2], with the line element in spherical coordinates (ω, θ, ϕ)

$$ds^2 = -cdt^2 + a^2(t)\left[d\omega^2 + f_K^2(\omega)\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right], \tag{3.3}$$

where $a(t)$ is called the scale factor $a(t)$ and the radial function

$$f_K(\omega) = \begin{cases} K^{-1/2}\sin(K^{1/2}\omega) & (K > 0) \\ \omega & (K = 0) \\ |K|^{-1/2}\sinh(|K|^{1/2}\omega) & (K < 0) \end{cases} \tag{3.4}$$

contains the curvature parameter K, describing the spatial geometry of the Universe.

[2] Howard Percy Robertson (1903-1961) was an American mathematician and physicist.
Arthur Geoffrey Walker (1909-2001) was an English mathematician.

By assuming that the content of the Universe is described as an ideal fluid, with density ρ and pressure p, containing a possible cosmological constant Λ, the Robertson-Walker metric simplifies Einstein's field equations to

$$\frac{\dot{a}^2}{a} = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3} \tag{3.5}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}. \tag{3.6}$$

The last two equations can be combined to obtain the dynamical relation between the scale factor and the density

$$\frac{d}{dt}\left(\rho c^2 a^3\right) + p\frac{d(a^3)}{dt} = 0. \tag{3.7}$$

To relate the density to the pressure of different matter components, one needs to provide an equation of state. For relativistic matter this equation fulfils $p = \rho c^2/3$, while for nonrelativistic matter $\rho c^2 = 0$ holds. The evolution of the different components is therefore given by

$$\rho_\mathrm{r} \propto a^{-4} \quad \text{for relativistic matter, also called radiation} \tag{3.8}$$

$$\rho_\mathrm{m} \propto a^{-3} \quad \text{for nonrelativistic matter, also called dust} \tag{3.9}$$

$$K \propto a^{-2} \quad \text{for the curvature parameter} \tag{3.10}$$

$$\Lambda \propto \text{const.} \quad \text{for the cosmological constant, hence the name} \tag{3.11}$$

We have good indications that the Universe is spatially flat (Komatsu et al., 2009), meaning $K = 0$. While incorporating the contribution from the cosmological constant in the total density ρ in Eq. 3.5, it is convenient to introduce the so-called critical density of the Universe, which ensures the flatness condition

$$\rho_\mathrm{cr} = \frac{3H^2}{8\pi G}, \tag{3.12}$$

where we defined the Hubble function $H(a) := \dot{a}/a$.

The scale factor today is conveniently set to one and the values of quantities at that particular point in cosmic evolution are denoted with the subscript 0. We obtain the Hubble constant $H_0 := H(1)$ and the critical density today

$$\rho_\mathrm{cr0} = \frac{3H_0^2}{8\pi G}. \tag{3.13}$$

H_0 is often parametrised by the dimensionless quantity h as

$$h := \frac{H_0}{100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}}. \tag{3.14}$$

By defining the density parameter $\Omega = \rho/\rho_\mathrm{cr}$, we rescale Eq. 3.7 to the critical density today

$$H^2 = H_0^2\left[\frac{\Omega_\mathrm{r0}}{a^4} + \frac{\Omega_\mathrm{m0}}{a^3} + \Omega_{\Lambda 0} + \frac{1 - \Omega_\mathrm{m0} - \Omega_\mathrm{r0} - \Omega_{\Lambda 0}}{a^2}\right]. \tag{3.15}$$

In this form, Eq. 3.15 is often called Friedmann[3] equation and describes the evolution of different energy-density components in an homogeneous Universe Finally, we will introduce so-called comoving coordinates

$$\boldsymbol{x} = \frac{1}{a}\boldsymbol{r}, \tag{3.16}$$

[3] Alexander Alexandrovich Friedmann (1888-1925) was a Russian cosmologist and mathematician.

which eliminate the cosmic expansion from the physical coordinates $\boldsymbol{r}$. The typical measure of time or distance during cosmic evolution is the so-called redshift given by

$$z := \frac{1}{a} - 1. \tag{3.17}$$

3.1.2 Linear structure formation

As we have seen in the last section, the density of radiation in the Universe scales as a^{-4} in Eq. 3.15, whereas the nonrelativistic matter scales as a^{-3}. Because of its steeper decay and its small value today, we can safely neglect the density contribution from the radiation at later times. If we furthermore assume the Universe to be spatially flat, we are left with a matter component and the cosmological constant.

Evolution of the density contrast
We consider now small perturbations in the smooth matter background, denoted as $\bar{\rho}(t)$ and governed by Eq. 3.15. The density contrast

$$\delta(\boldsymbol{x}, t) = \frac{\rho(\boldsymbol{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \tag{3.18}$$

depends on the comoving position $\boldsymbol{x}$. As is shown in e.g. Dodelson (2003), a linearly perturbed Robertson-Walker metric leads to a simple evolution equation for the density contrast

$$\ddot{\delta} + 2H\dot{\delta} - \left(4\pi G\bar{\rho} + \frac{c_\mathrm{s}^2\nabla^2}{a^2}\right)\delta = 0, \tag{3.19}$$

with the sound speed for the cosmic fluid defined as $c_\mathrm{s}^2 = \partial P/\partial\rho$. Solutions to this damped oscillator-like equation are found by writing the density contrast in Fourier space[4]

$$\ddot{\delta} + 2H\dot{\delta} - \left(4\pi G\bar{\rho} + \frac{c_\mathrm{s}^2 k^2}{a^2}\right)\delta = 0, \tag{3.20}$$

with the wave vector $\boldsymbol{k}$ describing the scale of a perturbation through its frequency in Fourier space. This remarkably simple equation holds already several cosmological implications:

1. The cosmic expansion, expressed by the Hubble function H, acts as a damping term in the equation above. This is not surprising, since it is working against gravitational collapse and smooths out over-densities.

2. If we neglect the Hubble flow for a moment, so $H = 0$, the third term of the evolution equation defines a typical length scale, which is called the Jeans[5] length $\lambda_\mathrm{j} = \frac{2\pi}{k_\mathrm{j}}$, through the Jeans wave number

   $$k_\mathrm{j} := \frac{2\sqrt{\pi G\bar{\rho}}}{c_\mathrm{s}}, \tag{3.21}$$

   since for $k \geq k_\mathrm{j}$ the frequency of the oscillator equation becomes real. The result is that perturbations smaller then the Jeans length are just oscillating.

3. In the case of $k < k_\mathrm{j}$ the oscillator frequency is complex and we find a growing and a decaying solution with time.

[4] We will suppress the usual hats over the Fourier transformations of quantities. If it is not clear from the context if an equation is written in the coordinate or frequency domain we will explicitly clarify the notation.
[5] Sir James Hopwood Jeans (1877-1946) was an English physicist, astronomer, and mathematician

4. The Jeans length depends on the sound speed c_s. This concept makes no sense for dark matter since we assume it to be pressureless. In this case of collisionless fluids, the Jeans length has to be replaced by a quantity which depends on the velocity dispersion of the fluid particles. For so-called cold dark matter (CDM), we assume this velocity dispersion to vanish. Therefore, linear perturbations can grow or decay on all scales[6].

Since the decaying solution is not interesting for structure formation, we focus on the growing solution which can be expressed by $\delta(a) = \delta_0 D_+(a)$, where the linear growth factor D_+ is given by

$$D_+(a) = \frac{G(a)}{G(1)} \tag{3.22}$$

with the approximative solution in a Universe containing cold dark matter and a cosmological constant (Lahav et al., 1991)

$$G(a) = a\Omega_\mathrm{m}\left[\Omega_\mathrm{m}^{4/7} - \Omega_\Lambda + \left(1 + \frac{\Omega_\mathrm{m}}{2}\right)\left(1 + \frac{\Omega_\Lambda}{70}\right)\right]^{-1}. \tag{3.23}$$

The power spectrum

One prediction of inflation is that the initial density contrast is nearly a Gaussian random field (see e.g. Liddle and Lyth, 2000, for a review). Such a Gaussian random field is completely described by its mean and variance. The mean of the density contrast is zero by construction, which leaves the variance to characterise the density contrast. Gaussian perturbations of different scales k grow independently in linear perturbation theory, as can be seen in Eq. 3.20. We will therefore describe the variance of the density contrast in Fourier space, which is called the power spectrum

$$\left\langle \delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')\right\rangle \equiv (2\pi)^3 P_\delta(\boldsymbol{k})\delta_\mathrm{D}(\boldsymbol{k} - \boldsymbol{k}'), \tag{3.24}$$

where Dirac's[7] delta function ensures, that different wave-vector modes remain uncoupled. The variance of the density contrast in coordinate space is then given by

$$\sigma^2 = 4\pi \int \frac{k^2\, dk}{(2\pi)^3} P(k). \tag{3.25}$$

If only particular, physical scales $> R$ are considered, the variance of the density contrast is given via a window function in Fourier space

$$\sigma_R^2 = 4\pi \int \frac{k^2\, dk}{(2\pi)^3} W_R^2(k) P(k). \tag{3.26}$$

A common normalisation of the power spectrum is obtained by the variance of the density contrast on a scale of 8 h^{-1} Mpc. This parameter σ_8 is one of the largest uncertainties within the framework of ΛCDM.

Before we can compute the power spectrum as a function of k, we have to take into account one more thing. So far we have assumed that perturbations grow in the matter-dominated era. During the radiation-dominated era, where the contribution of radiation to the total density of the Universe is larger than the contribution of matter, the growth of structure is heavily suppressed due to the fast expansion of the Universe. This only affects perturbation

[6] It is exactly this fact which highly disfavours neutrinos as the only constituent in the dark-matter puzzle

[7] Paul Adrien Maurice Dirac (1902-1984) was a British theoretical physicist.

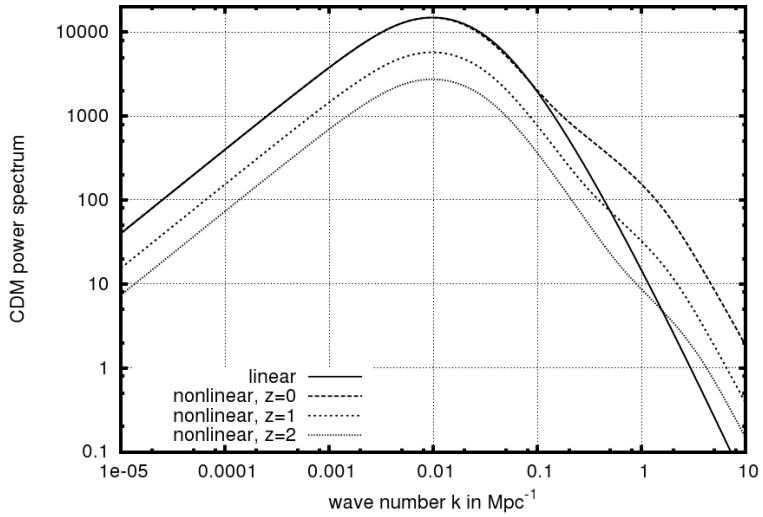


Figure 3.1: From Bartelmann (2010). Linearly and nonlinearly evolved CDM power spectra. The linear CDM spectrum is shown for $a = 1$ and scales like $D_+^2(a)$.

modes that were already in causal contact during the time of radiation domination. The minimum wave number refering to patches of the Universe causally connected at the time when the density parameter of radiation was equal to the density parameter of matter is given by

$$k_{\mathrm{eq}} = \frac{H_0}{c}\frac{\pi}{\sqrt{2}-1}\left(\frac{a_{\mathrm{eq}}}{\Omega_{\mathrm{m0}}}\right)^{-1/2} . \tag{3.27}$$

Thus, perturbation modes with $k \ll k_{\mathrm{eq}}$ have no imprint of the radiation-dominated era in their amplitude while we expect structures with $k \gg k_{\mathrm{eq}}$ to be significantly suppressed. Up to an unknown normalisation, the CDM power spectrum is then given by (see e.g. Harrison, 1970; Peebles and Yu, 1970; Zel'dovich, 1972).

$$P_\delta(k) \propto \begin{cases} k^{n_{\mathrm{s}}} & (k \ll k_{\mathrm{eq}}) \\ k^{n_{\mathrm{s}}-4} & (k \gg k_{\mathrm{eq}}) \end{cases} , \tag{3.28}$$

with the spectral index for the initial seed perturbations n_{s}. A plot of the linear power spectrum, compared with the fit to a nonlinear power spectrum found by numerical techniques, is presented in Fig. 3.1.

3.1.3 Non-linear structure formation and numerical simulations

Linear perturbation theory gives a good conceptual understanding of structure formation and describes the density contrast well, as long as its value stays well below unity. Nevertheless, even the largest individual objects in the Universe, decoupled from the Hubble flow, being clusters of galaxies easily exceed values of 200 in the density contrast. Also in Fig. 3.1 it can be seen that significant differences appear at the small-scale end of the power spectrum, when comparing the linear calculation with a numerical, nonlinear result. The reason for these differences is the transfer of power from larger to smaller scales. This is possible since perturbations of different scales do not evolve independently any more. In the literature this phenomenon goes under the name of mode coupling and can only be described in a nonlinear theory of structure formation.

Zel'dovich approximation

A first step towards a nonlinear theory was done by Zel'dovich (1970) who had the idea that the peculiar motion of particles in the presence of density perturbations can be described by a simple displacement in the comoving frame

$$\boldsymbol{x} = \boldsymbol{x}_0 + b(t)\boldsymbol{f}(\boldsymbol{x}_0), \tag{3.29}$$

where the initial position of the particle $\boldsymbol{x}_0$ is shifted by a displacement field $\boldsymbol{f}$. The time dependence of the shift is expressed by the function b.
Writing the formula in physical coordinates gives

$$\boldsymbol{r}(t) = a(t)\left[\boldsymbol{x}_0 + b(t)\boldsymbol{f}(\boldsymbol{x}_0)\right], \tag{3.30}$$

thus being a mapping between comoving and physical coordinates incorporating the particle displacement. While integrating over all particle trajectories and assuming mass conservation, we can relate the density at the initial comoving position $\boldsymbol{x}_0$ to the comoving mean density at some time t by the use of the Jacobian determinant of the mapping

$$\rho(\boldsymbol{x}_0, t)\left|\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}_0}\right| = \rho(\boldsymbol{x}_0, t)\left|a(t)\delta_{ij} + a(t)b(t)\frac{\partial^2 f}{\partial x_i \partial x_j}\right| = a(t)^3\bar{\rho}. \tag{3.31}$$

With $(\lambda_1, \lambda_2, \lambda_3)$ being the eigenvalues of the deformation tensor $f_{ij} := \partial^2 f/\partial x_i \partial x_j$, we find for the density contrast

$$\delta(\boldsymbol{x}_0, t) = \frac{1}{\left[1 + b(t)\lambda_1\right]\left[1 + b(t)\lambda_2\right]\left[1 + b(t)\lambda_3\right]}. \tag{3.32}$$

The Zel'dovich approximation gives already interesting insights into the formation of nonlinear structures. When the trajectories of particles cross the mapping above becomes singular, meaning that at least one of the eigenvalues becomes zero. It can be shown, while assuming that the perturbations arise from Gaussian random fields, that the probability of finding two identical eigenvalues is zero. Hence, gravitational collapse will be anisotropic. This is exactly what is observed in reality. Collapsed structures appear as filaments and 'blinies', as they were called by Zel'dovich.

Spherical collapse

We have just learned, that an isotropic collapse is excluded during structure formation. Nevertheless, we will now assume exactly this condition for the simple reason that it allows an analytic solution to the evolution of the density contrast.

We consider a spherically symmetric overdense region, embedded in the expanding Universe. The Newtonian equation of motion for a test particle within a thin shell at physical distance r from the centre of the sphere is given by

$$\ddot{r} = -\frac{GM(<r)}{r^2}, \tag{3.33}$$

with the total mass of the sphere

$$M(<r) = \frac{4}{3}\pi r^3 \bar{\rho}\left[1 + \frac{3}{r^3}\int_0^r \delta(\tau)\tau^2 d\tau\right] = \frac{4}{3}\pi r^3 \bar{\rho}\left(1 + \bar{\delta}\right), \tag{3.34}$$

given by the smooth background density $\bar{\rho}$ and the average overdensity within the sphere $\bar{\delta}$.
Integrating the equation of motion gives the total Energy per unit mass

$$E = \frac{1}{2}\dot{r}^2 - \frac{GM}{r}. \tag{3.35}$$

We can already distinguish three cases:

1. If $E > 0$, $\dot{r}$ cannot vanish and the sphere will expand forever together with the Universe.
2. If $E = 0$, $\dot{r}$ can only vanish for $r \to \infty$.
3. If $E < 0$, $\dot{r}$ vanishes at a given time t_{ta} marking a turn-around when the overdensity decouples from the expansion and collapses.

We can rewrite the kinetic and the potential term in Eq. 3.35 and relate them to a cosmological model, which only considers a dark-matter contribution through its density parameter Ω_{m}. In a flat Universe, the collapse condition $E < 0$ can be rewritten as $\bar{\delta} > \Omega_{\mathrm{m}}^{-1} - 1$.

$$\frac{\dot{r}^2}{2} = \frac{H^2 r^2}{2} \tag{3.36}$$

$$-\frac{GM}{r} = -\frac{1}{2} H^2 r^2 \left(1 + \bar{\delta}\right) \Omega_{\mathrm{m}}. \tag{3.37}$$

The maximum radius of the sphere before collapse at turn-around

$$r_{\mathrm{ta}} = r(t_{\mathrm{ta}}) = r_{\mathrm{i}} \frac{1 + \bar{\delta}_{\mathrm{i}}}{1 + \bar{\delta}_{\mathrm{i}} - \Omega_{\mathrm{m},i}^{-1}}, \tag{3.38}$$

contains the initial quantities at the onset of the sphere's evolution. The collapsing sphere will reach virial equilibrium relating kinetic energy K and potential energy U to the total energy E, which implies that the radius of the sphere in virial equilibrium is given by $r_{\mathrm{vir}} = r_{\mathrm{ta}}/2$. As a next step, it is useful to define the virial overdensity Δ_{vir} defined as the average density of the sphere at virialisation scaled with the critical density of the Universe.

In order to find the relation between the radius of the sphere and the time, one needs to solve the equation motion. In the case of an Einstein-de Sitter Universe[8], this can be done analytically and a value for the virial overdensity of $\Delta_{\mathrm{vir}} \backsimeq 178$ is found. If the linear density contrast is extrapolated to the according time when virial equilibrium is reached in the spherical collapse model, a value of $\delta_{\mathrm{c}} \backsimeq 1.686$ is found. This implies that linear theory is already falsifying itself, when predicting a density contrast for which its assumptions do not hold any more. We will use this result later on.

We have already mentioned, that a spherical collapse is not possible in nonlinear structure formation and the calculations above can only be seen as an instructive, but incomplete model. For a more rigorous analytic treatment one has to apply nonlinear perturbation theory. The mathematics of this extremely interesting subject include representations by Feynman-diagrams, known from quantum field theory, where linear perturbations of a certain mode are represented as propagators and nonlinear mode-coupling by vertices of different modes, while applying methods of renormalisation. The interested reader is referred to Crocce and Scoccimarro (2006); Matarrese and Pietroni (2007) and Pietroni (2008).

Numerical simulations

The commonly used approach to master the complexity of nonlinear structure formation exploits the tremendous numerical performance of modern computer systems. These so-called numerical simulations of cosmic structure formation can be performed on different length scales, ranging from volumes significant compared to the size of the Universe (Springel et al., 2005; Gottlöber et al., 2006; Crocce et al., 2010), down to the formation of the birth places where the first stars formed (Abel et al., 2000). Also the formation and evolution of galaxy

[8] Meaning a flat Universe containing only a matter component $\Omega_{\mathrm{m}} = 1$

clusters can be modeled via numerical simulations and we will briefly describe this process. An excellent review on the topic was written by Borgani and Kravtsov (2009).

The full dynamics of dark matter is governed by the collisionless Boltzmann[9] equation, also known as the Vlasov[10] equation. It describes the trajectories in the 6D phase space of coordinates and velocities. Of course, the complete phase space volume cannot be explored, but characteristic contours of phase-space trajectories are obtained by sampling the total volume with a number of N particles and integrating their equations of motions in the gravitational field created by the full ensemble. This method is therefore called an N-body simulation. Simulations can not only follow dark matter, but also baryonic matter, whose trajectories in phase space are described by the hydrodynamical equations (e.g. Landau and Lifshitz, 1987). Therefore, we can now start to leave the track of purely tracing dark matter in the Universe, which is the dominant component during structure formation, but does not give rise to the astrophysical processes coupling to the photons that we actually observe.

The first step which has to be fixed before running a cosmological simulation are the initial conditions deciding how the dark matter and baryonic particles are initially placed within the simulation box. Fortunately, the initial power spectrum of matter in the Universe at a very early time is known to a fascinating precision due to the observations of the CMB. The power spectrum at this time of recombination can be evolved by linear, or mildly nonlinear techniques like the Zel'dovich approximation, to the initial time step of the simulation. The particles within the box are placed according to the statistics of this power spectrum (see Prunet et al., 2008; Sirko, 2005, for recent reviews).

Once the initial conditions are fixed, gravitational forces on the individual dark matter and baryonic particles in the simulation have to be calculated. This can be done by e.g. direct summation over all particles in the box, but the number of computational operations would scale with the square of the number of particles in the box. This numerical scheme is therefore not suited for cosmological simulations involving a large number of particles. Approximating summation schemes have been developed, representing different levels of trade-off between numerical accuracy and computational time. We want to mention particle-mesh and particle-particle/particle-mesh methods (Hockney and Eastwood, 1988; Couchman, 1991), which calculate the forces between particles and mass-averaged areas of the box, tree-codes which hierarchically increase the resolution of the force calculation by an adaptive multipole-expansion scheme (Barnes and Hut, 1986; Bouchet and Hernquist, 1988) and hybrid methods, like the widely spread code `GADGET` by Springel (2005), which combines both approaches.

More involved is the calculation of the hydrodynamic forces. Two approaches are commonly used, smooth particle hydrodynamics (SPH) on the one hand, grid based methods on the other hand. In SPH (Monaghan, 2005) the fluid is represented by particles and the governing equations are obtained from the Lagrangian form of the hydrodynamical conservation laws. In grid based methods a Eulerian approach is followed by discretising the hydrodynamical equations. Both methods have advantages and problems, SPH is fast in its computations and naturally follows regions of interest with increased resolution, nevertheless it needs to make use of an artificially introduced viscosity, which prevents the development of shear motions and usually fails in correctly describing transient areas between different phases of the fluid. Possible solutions to this problem are e.g. given in Dolag et al. (2005). Grid methods do not suffer from such shortcomings but are computationally more demanding. Furthermore, their spatial resolution is somewhat arbitrarily set by the size of the discretising grid. Those

[9] Ludwig Eduard Boltzmann (1844-1906), Austrian physicist.
[10] Anatoly Alexandrovich Vlasov (1908-1975), Russian theoretical physicist

problems are cured while making use of an adaptive-mesh-refinement (AMR) technique (e.g. Norman and Bryan, 1999), which increases the resolution of the grid in areas of interest, e.g. in areas of high particle density. This increases the spatial resolution in those areas and reduces the runtime by lowering the resolution in the other areas. A detailed comparison between the different hydrodynamical approaches can be found in Agertz et al. (2007). A new Lagrangian method, using a dynamic Voronoi tessellation was suggested by Springel (2010) and seems not to suffer from the problems attached to SPH.

In order to make numerical simulations of galaxy clusters more realistic, additional physical processes are added by introducing sink and source terms on the right-hand side of the hydrodynamical equations. State-of-the-art simulations of galaxy clusters include:

- **Star formation**
  Whenever a group of baryonic particles reaches a density exceeding a certain threshold, these particles are extracted from the collisional fluid and converted into a collisionless, so-called sink particle representing a population of new-born stars. Of course the spatial resolution of a simulation is not high enough to resolve the processes of star formation, which is why the model describing this process approximately on the grid is assigned to the group of so-called sub-grid models.

- **Radiative cooling**
  Since we are only assuming gravitational and hydrodynamical interactions, the loss of energy due to the thermal and nonthermal emission of particles has to be modeled separately. This effect of cooling plays an important role in the description of the intracluster gas, as we will see later.

- **Heating by feedback processes**
  Particles can not only loose energy due to cooling, they can also gain energy via heating processes. Usually, energy injection due to Supernova explosions and violent accretion onto the supermassive black hole of an active galactic nucleus (AGN) (see McNamara and Nulsen, 2007, for a review) is taken into account.

- **Chemical enrichment**
  Observations indicate (see e.g. Werner et al., 2008, for a review), that clusters contain a significant amount of metals[11]. This metallicity, which alters heating and cooling within the cluster, is due to the nuclear processes within stars and enriches the cluster gas e.g. via supernova explosions.

We note that the influence and evolution of magnetic fields within a cluster simulation, described by magneto-hydrodynamics (MHD) , is out of the scope of this work (see Dolag and Stasyszyn, 2009, and references therein).

An example for a typical result of a state-of-the-art cosmological simulation of the formation of a galaxy clusters is shown in Fig. 3.2. In general, the results can summarised as follows:

1. Structure formation is dominated by the dark matter component and follows an hierarchical merging process.

2. The gas component follows globally the potential wells of the dark matter, but shows a smoother structure due to its additional pressure support.

[11] Defined as any element despite hydrogen

3. Stars form in high-density regions of the dark matter halo.

4. Turbulent motion plays also an important role in relaxed structures and can significantly bias the assumption of hydrostatic equilibrium in virialised structures (Rasia et al., 2004, 2006; Iapichino and Niemeyer, 2008). A trend which seems to be indirectly confirmed observationally by Mahdavi et al. (2008) and direct observation of turbulent motion within the gas component of galaxy cluster might be possible with future high-resolution instruments (Inogamov and Sunyaev, 2003).

Despite their tremendous success in reproducing several observed quantities of real clusters, there are still some open questions in the comparison between simulations and observations, which are mostly related to baryonic physics. We will mention the most pressing: The first puzzle is the absence of indicators for gas cooling in the observations of dense and relaxed, so-called cool-core clusters. Because of their high X-ray emissivity the core should cool relatively fast in the centre. Nevertheless, high-resolution X-ray observations with `Chandra` or `XMM-Newton` do not show any presence of cool gas. A more severe problem refers to the fact that the derived star formation rate in simulated galaxy clusters differs from observations by an order of magnitude. This is the manifestation of the so-called overcooling problem which is, not like the cool-core problem, only an issue at the present time but during the whole evolution of a simulated galaxy cluster. Both, the cool-core problem and the overcooling problem, probably related, show that the gas physics in numerical simulations of galaxy clusters does not obey yet the right balance between heating and cooling. We will mention additional problems later on.

3.2 Observations of galaxy clusters

We have learned from the theory of the formation of galaxy clusters, that they consist of three main components: A large and extended dark-matter halo, a halo of diffuse intra-cluster gas and the collapsed gas component, forming stars and galaxies. Those three components give rise to different kinds of observations. The observational analysis of clusters splits into two important steps. First, clusters have to be detected and objects have to be assigned as cluster members. Once a cluster is identified, the derivation of astrophysical properties can be performed, possibly in several wave bands.

3.2.1 Optical

The first detections of galaxy clusters were naturally obtained in the optical wave bands and go back to early observations of the Virgo and the Coma clusters by Messier and Herschel in the late 18th century. The first stringent approach to define and categorise clusters of galaxies was done by Abell (1958), where he was very careful in defining clusters because projection effects may easily dilute the cluster membership association. Abell's first step was to estimate the distance to a cluster candidate through the magnitude of its tenth-brightest member galaxy. For a given radius, which turned out to be ~ 2 Mpc for all clusters, he counted the galaxies with a magnitude not fainter than the magnitude of the cluster candidate's third-brightest galaxy. After subtracting a background level, to account for projection effects, he defined candidates as a cluster if it exceeded a count of 50. Modern versions of this heuristic approach to optically detect clusters use additional colour information by exploiting the so-called red sequence, which reflects the fact that the old stellar population of cluster galaxies fall into a distinct region within a colour-magnitude diagram (e.g. Gladders

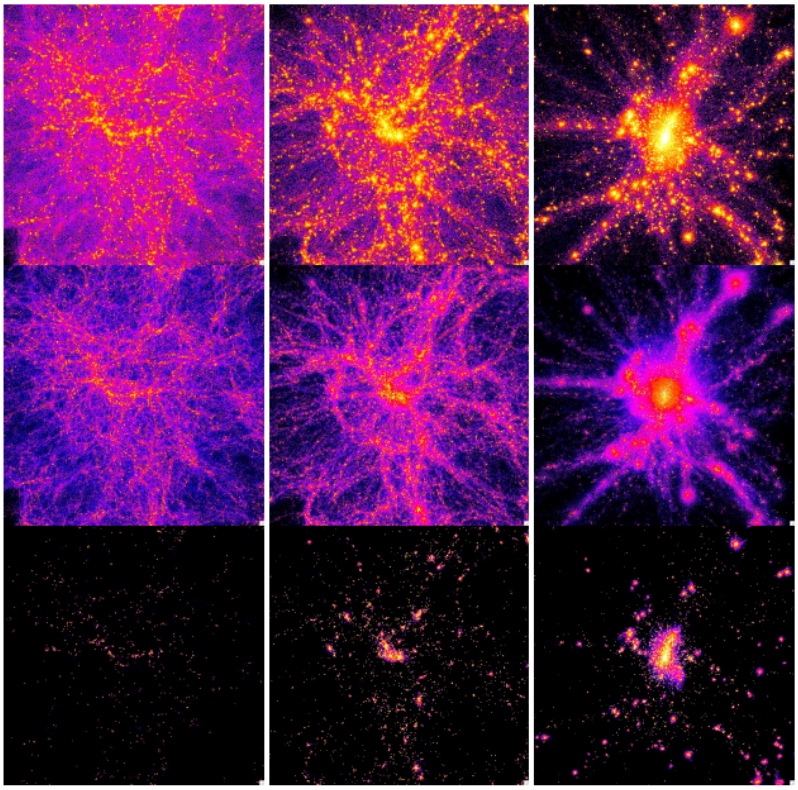


Figure 3.2: From Borgani and Kravtsov (2009). The formation of a galaxy cluster in a cosmological context, as described by a hydrodynamical simulation carried out with the Tree-SPH GADGET code (Springel et al., 2005). *Upper*, *central* and *bottom panels* refer to the density maps of dark matter, gas and stellar distributions, respectively. From left to right we show the snapshots at $z = 4$, where putative proto-cluster regions are traced by the observed concentrations of Lyman-break galaxies and Lyman-α emitters (e.g. Overzier et al., 2008), at $z = 2$, where highly star-forming radio-galaxies should trace the early stage of cluster formation (Miley et al., 2006; Saro et al., 2009), and at $z = 0$. This cluster has a total virial mass $M_{\rm vir} = 10^{15}\ h^{-1}\ M_{\odot}$ at $z = 0$ (Dolag et al., 2009). Each panel covers a comoving scale of about $24\ h^{-1}$ Mpc, while the cluster virialised region at $z = 0$ is nearly spherical with a radius of about $3\ h^{-1}$ Mpc.

and Yee, 2000). Once a cluster is optically detected, already the observed high-luminosity tail of its brightest member galaxies yields a mass estimate since their luminosity distribution generally follows a Schechter function (Schechter, 1976). This allows the definition of richness criteria (Postman et al., 1996; Yee and López-Cruz, 1999), which can be related to the total mass of the cluster and represent therefore a first mass estimate.

More elaborate methods of deriving cluster masses from optical observations involve spectroscopy. Once radial velocities of member galaxies are determined, it can be measured if their distribution follows a Gaussian profile with the 1D velocity dispersion σ_v (Binney and Tremaine, 2008). These measurements were initially performed by Zwicky (1933, 1937) for the Coma cluster and he acquired a value of $\sigma_\mathrm{v} \sim 700$ km s^{-1}. By applying the virial theorem of classical mechanics, a mass estimate of the total mass for the virialised structure, which has to assumed as a closed physical system with radius R, yields

$$M \approx \frac{3R\sigma_\mathrm{v}^2}{G} \approx 10^{15} M_\odot \, h^{-1}\mathrm{Mpc} \left(\frac{R}{1.5h^{-1}\mathrm{Mpc}}\right)\left(\frac{\sigma_\mathrm{v}}{1000 \text{ km s}^{-1}}\right). \tag{3.39}$$

Mass estimates obtained with the virial theorem are questionable since clusters are not separate systems with respect to their environment. Nevertheless, those mass estimates where the first hint towards dark matter, because the calculated cluster masses exceeded the expectations based on the observed luminosity by orders of magnitude.

Another aspect of the optical analysis of galaxy clusters is provided by the effect of gravitational lensing, where the cluster acts as the lens and bends the light of distant background sources. This effect has already been discussed and the way to obtain cluster properties from this method will be the content of the next chapters.

3.2.2 X-ray

We have seen before that clusters contain a large amount ($\sim$ 15% of their total mass) of baryonic intra-cluster gas. If we assume that the gas follows the same kinematics as the cluster members from the previous section, we expect a temperature of

$$kT \approx \mu m_\mathrm{p} \sigma_\mathrm{v}^2 \approx 6 \left(\frac{\sigma_\mathrm{v}}{10^3 \text{ km s}^{-1}}\right)^2 \text{keV}, \tag{3.40}$$

with the mean molecular density μ, the proton mass m_p and the Boltzmann constant k[12].

Indeed, an X-ray continuum due to thermal bremsstrahlung, indicating a gas temperature of $kT \gtrsim 2$ keV is observed. Also a line spectrum contributes, since clusters contain a significant amount of heavy elements ($\sim$30 % of the solar metallicity). This line spectrum dominates for cluster temperatures $\lesssim 2$ keV. The observed X-ray luminosities of clusters span a range from $10^{43} - 10^{45}$ erg s^{-1}, making them easily detectable lighthouses in the X-ray sky. A review on the observations of clusters in the X-ray band and their interpretation is given by Rosati et al. (2002).

To obtain an X-ray mass estimate, one has to assume spherical symmetry and hydrostatic equilibrium of the emitting system. While doing so, pressure p and density ρ_g of the gas are related by

$$\frac{dp}{dR} = -\frac{GM(<R)\rho_\mathrm{g}(R)}{R^2}. \tag{3.41}$$

We assume an ideal cluster gas and insert the equation of state

$$p = \frac{\rho_\mathrm{g} kT}{\mu m_\mathrm{p}}, \tag{3.42}$$

[12] $k = 8.617342(15) \times 10^{-5}$ eV K^{-1}.

to obtain the X-ray mass estimate

$$M(< R) = -\frac{kTR}{G\mu m_\mathrm{p}} \left(\frac{d\ \log \rho_\mathrm{g}}{d\ \log R} + \frac{d\ \log T}{d\ \log R} \right). \tag{3.43}$$

X-ray observations provide not only the X-ray luminosity but also the X-ray temperature through spectral fitting. Both can be obtained nowadays with high angular resolution, using the space observatories `Chandra`, `XMM-Newton` and `Suzaku`.

In order to obtain the mass from those observed quantities in several radial bins, two techniques are usually applied, labeled forward and backward methods. The forward approach (Vikhlinin et al., 2006; Pratt and Arnaud, 2003) assumes a parametric form of the gas-density profile like the β-model (Cavaliere and Fusco-Femiano, 1978)

$$\rho_\mathrm{g}(r) = \rho_{\mathrm{g},0} \left[1 + \left(\frac{r}{r_\mathrm{c}} \right)^2 \right]^{-3\beta/2}, \tag{3.44}$$

or generalisations of it (Pratt and Arnaud, 2002, and references therein) and a parametric form of the temperature profile (Allen et al., 2001)

$$T(r) = T_0 + \frac{(r/r_\mathrm{t})^{-a}}{[1 + (r/r_\mathrm{t})^b]^{c/b}} T_\mathrm{cool}, \tag{3.45}$$

projects them properly along the line-of-sight (Mazzotta et al., 2004) and fits them to the observations. The backward approach (Ettori et al., 2002; Morandi et al., 2007) deprojects the observed quantities and compares them to the predictions of an analytical 3D mass profile of the cluster. A comparison between the different methods can be seen in Meneghetti et al. (2009).

The mass estimates obtained with these methods seem to be quite robust but critical questions focus on the strong assumptions which went into their derivation, especially the assumption of hydrostatic equilibrium and spherical symmetry. Numerical simulations show, that these assumptions are not generally a good description of the shape and the dynamical state of a galaxy cluster, which might contribute to differences as large as factors of two in the mass estimate obtained with X-ray techniques, compared to more direct methods like gravitational lensing (Morandi et al., 2010, and references therein).

3.2.3 Microwave

Clusters of galaxies are also observable in the microwave band, due to the the Sunyaev-Zel'dovich (SZ) effect (Sunyaev and Zeldovich, 1970, 1972). CMB photons pass through the electrons of the hot cluster gas and experience a shift in frequency by inverse Compton scattering. This shift can be directly measured in the temperature distribution of the CMB as a secondary anisotropy. This effect is called the thermal Sunyaev-Zel'dovich effect and the derivation of the temperature shift

$$\frac{\Delta T_\mathrm{SZE}}{T_\mathrm{CMB}} = f(x) y = f(x) \int n_\mathrm{e} \frac{kT_\mathrm{e}}{m_\mathrm{e} c^2} \sigma_\mathrm{T}\ dl \tag{3.46}$$

can be followed in the extensive review of Birkinshaw (1999). In the formula above, T_CMB is the measured temperature of the CMB today, y is called the Compton y-parameter, n_e is the electron density in the cluster, T_e their temperature and m_e their mass, σ_T the Thomson

scattering cross-section and dl denotes integration along the line-of-sight. $x = h\nu/(kT_{\mathrm{CMB}})$ encodes the frequency dependence of the effect, which depends also on relativistic effects (Shimon and Rephaeli, 2004; Itoh and Nozawa, 2004) expressed by the correction function δ_{SZE}

$$f(x) = \left(x \frac{e^x + 1}{e^x - 1} - 4 \right) \left[1 + \delta_{\mathrm{SZE}}(x, T_{\mathrm{e}}) \right] . \tag{3.47}$$

The overall change in the CMB photon intensity is shown in Fig. 3.3. The so-called kinetic Sunyaev-Zel'dovich effect is caused by the clusters' peculiar motion against the CMB background and gives therefore the possibility of measuring it. So far, the sensibility of current instruments is not high enough to measure peculiar cluster velocities reliably through the tiny signal of the kinetic SZ effect as results of Benson et al. (2003) indicate.

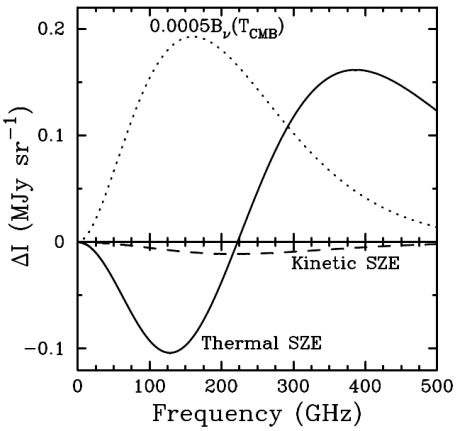


Figure 3.3: From Carlstrom (2002). Spectral intensity distortion of the cosmic microwave background (CMB) radiation due to the Sunyaev-Zel'dovich effect. The thick solid line is the thermal SZ effect and the dashed line is the kinetic SZ effect. For reference the 2.7 K thermal spectrum for the CMB intensity scaled by 0.0005 is shown by the dotted line. The cluster properties used to calculate the spectra are an electron temperature of 10 keV, a Compton y parameter of 10^{-4} , and a peculiar velocity of 500 km s^{-1}.

Three things are especially noticeable about the use of the SZ effect for understanding galaxy clusters. First, the effect is independent of the cluster's redshift and gives therefore the opportunity to detect also extremely distant clusters. Second, the effect is complementary with X-ray observations and could help to provide better estimates for the gas-density profile of the cluster because both effects are sensitive to the electron density of the ICM. Third, it allows a direct test of the Universe's geometry since one can derive the physical scale of the cluster through the line-of-sight integral (Eq. 3.46) and compare it to the observed scale (Birkinshaw et al., 1991).

3.3 Constraining the universe with galaxy clusters

After we have reviewed the theoretical description of clusters and their actual observations, we should combine both aspects to put constraints on cosmological models. Interestingly, clusters of galaxies still provide riddles within the cosmological standard model, which is in excellent agreement with other cosmological observations by now.

3.3.1 Mass function

An obvious cosmological test with clusters is counting their abundance in the sky and compare it with the theoretical prediction. The number of objects of a certain mass and redshift range is given by the so-called mass function $n(M, z)$. The pioneering steps of deriving the mass function from cosmological theories of structure formation were done by Press and

Schechter (1974), with substantial contributions later by Bond et al. (1991) and Lacey and Cole (1993).

If $Q(M,z)$ describes that fraction of the volume in the Universe, which contains per unit redshift structures heavier in mass than M, the mass function is given by

$$n = -\frac{\bar{\rho}}{M}\frac{\partial Q}{\partial M} = \frac{\bar{\rho}}{M}\left|\frac{dS}{dM}\right|\frac{\partial Q}{\partial S}, \tag{3.48}$$

where we have identified the variance of the density contrast with S for simplicity. We know that the variance is a decreasing function of mass, as can be seen in Fig. 3.1 and cancelled the minus sign. As we have seen in Sec. 3.1.2, the variance of the density contrast in coordinate space depends on a filtering function. If we choose a top-hat filter in Fourier space, the evolution of the filtered density contrast δ_f at a given coordinate, as a function of the filtered variance $S = S(M)$, will be a random walk. The reason for this is that increasing or decreasing the filtering length and therefore varying the mass scale in which we are interested, will add or eliminate k-modes in Fourier space, which are independent of each other in linear perturbation theory.

Following the spherical-collapse model of Sec. 3.1.3, we assume an overdensity as collapsed if the random walk of the density contrast passes over a critical barrier $\delta_\mathsf{c}(z)$[13]. We denote the value of δ_f when it passes the barrier with $w(z) = \delta_\mathsf{c}(z)/D_+(z)$. In this picture the problem is boiled down to finding the probability distribution of a random walk in the (δ_s, S) plane to be absorbed by a density barrier defined by $\delta_\mathsf{c}(z)$. This problem in a more general context has been solved by Chandrasekhar (1943), finding the fraction of collapsed structures

$$Q(S,w) = 1 - \int_{-\infty}^{w} W(\delta_\mathsf{f}, S, w)\, d\delta_\mathsf{f}, \tag{3.49}$$

and the fraction of trajectories per unit interval in the random walk, that for given variance S have a density contrast of δ_f

$$W(\delta_\mathsf{f}, S, w) = \frac{1}{\sqrt{2\pi S}}\left[\exp\left(-\frac{\delta_\mathsf{f}^2}{2S}\right) - \exp\left(-\frac{2w-\delta_\mathsf{f}^2}{2S}\right)\right]. \tag{3.50}$$

By plugging the derivative with respect to S of Eq. 3.49 into Eq. 3.48 we find the final result for the mass function

$$n(M,z) = \frac{\bar{\rho}(z)}{M}\frac{w(z)}{\sqrt{2\pi}S^{3/2}}\exp\left[-\frac{w^2(z)}{2S}\right]\left|\frac{dS}{dM}\right|. \tag{3.51}$$

If we assume the initial power spectrum to be scale free as in Sec. 3.1.2 we find

$$S(M) = \left(\frac{M}{M_0}\right)^{-\alpha}, \tag{3.52}$$

with the normalisation M_0 and $\alpha := (n_\mathsf{s}+3)/3$. The mass function is then given by

$$n(M,z) = \alpha\bar{\rho}(z)\frac{w(z)}{\sqrt{2\pi}}\frac{M^{\alpha/2-2}}{M_0^{\alpha/2}}\exp\left[-\frac{w^2}{2}\left(\frac{M}{M_0}\right)^{\alpha}\right]. \tag{3.53}$$

[13] In the case of spherical collapse in an Einstein-de Sitter Universe, this barrier is given by the critical value of $\delta_\mathsf{c} = 1.686$

This result is already in excellent agreement with the abundances of structures in simulations, despite its significant amount of assumptions. Nevertheless, the agreement can be improved by introducing the fitting parameters (A, B, p) in the mass function

$$n(M,z) = A\sqrt{B}\frac{\bar{\rho}(z)}{M}\frac{w(z)}{\sqrt{2\pi}S^{3/2}}\left[1+\left(\frac{S}{Bw(z)}\right)^{2p}\right]\exp\left[-B\frac{w^2(z)}{2S}\right]\left|\frac{dS}{dM}\right|. \tag{3.54}$$

The original mass function by Press and Schechter (1974) is recovered by setting $A = 1/2$, $B = 1$ and $p = 0$. By using a more general model of ellipsoidal collapse (Sheth et al., 2001), which introduces a moving barrier in the random walk, Sheth and Tormen (2002) find $A = 0.3222$, $B = 0.707$ and $p = 0.3$. Another good agreement with numerical results is found by Jenkins et al. (2001) with $A = 0.353$, $B = 0.73$ and $p = 0.175$. More recently the mass function was calibrated by Warren et al. (2006) and Tinker et al. (2008), where the latter also accounts for several different mass definitions, as they can be obtained with observational techniques.

Measuring the mass function is obviously an excellent tracer to determine the normalisation of the power spectrum, since it exponentially depends on the variance of the density contrast for a given mass and redshift. Furthermore, Ω_m enters directly into the mass function through the density evolution of the background. In order to determine the mass function for a large number of mass bins, one needs a convenient mass-observable relation. X-ray observables seem to be best suited for that purpose since single X-ray luminosity or temperature measurements give in principle already a robust mass estimate. Furthermore, by assuming that the process of cluster formation is self-similar, meaning that clusters of different masses are just scaled versions of one another, simple scaling relations can be derived for a given redshift between mass and X-ray temperature, between X-ray temperature and X-ray luminosity and consequently also between mass and X-ray luminosity. Unfortunately, these simple scaling relations do not hold in reality, due to the complicated gas physics that we have discussed in Sec. 3.1.3. They have to be parametrised and calibrated by precision mass measurements from other methods like gravitational lensing or by numerical simulations. A promising feature of large future surveys is the method of self-calibration, which intrinsically finds the most-likely parameters in the mass-observable relation (Levine et al., 2002; Hu, 2003; Majumdar and Mohr, 2003, 2004). As was already mentioned, a mass-observable relation from other methods, like suitable richness measurements of clusters, gravitational lensing or the observed velocity dispersion is usually more difficult to obtain, but the derived masses might be more accurate. This allows calibrating the X-ray scaling relations.

3.3.2 Evolution of the mass function

Much tighter constraints on the cosmological model can be placed if the evolution of the mass function with redshift is taken into account. The number of structures dN with mass M within a given solid angle $d\Omega$ and redshift interval dz is given by

$$\frac{d^3N}{dMd\Omega dz}(M,z) = \frac{dn}{dM}(M,z)\frac{d^2V_\mathrm{co}}{dzd\Omega}(z), \tag{3.55}$$

where the geometry of the Universe is probed by the comoving volume factor

$$\frac{d^2V_\mathrm{co}}{dzd\Omega}(z) = \frac{cf_k(z)}{H(z)}. \tag{3.56}$$

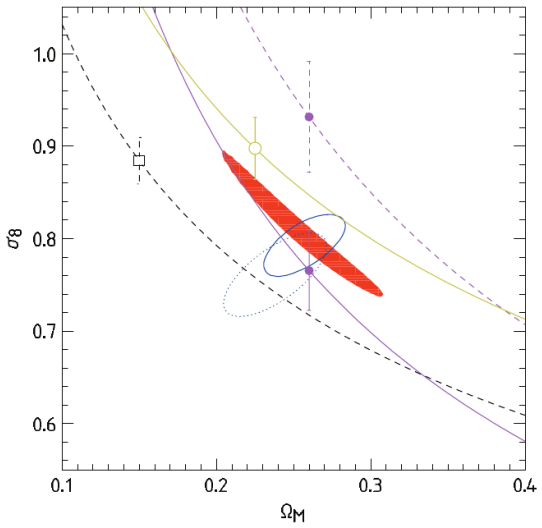


Figure 3.5: From Vikhlinin et al. (2009). Constraints on the σ_8 and Ω_m parameters in a flat ΛCDM cosmology. The solid region is the 68% confidence level region reproduced by Vikhlinin et al. (2009). Blue contours show the WMAP 3 and 5 year results from Spergel et al. (2007) and Dunkley et al. (2009) (dotted and solid contours, respectively). For other measurements, the general direction of degeneracy as a solid line and a 68% uncertainty in σ_8 at a representative value of Ω_m is shown. Filled circles show the weak-lensing shear results from Hoekstra et al. (2006) and Fu et al. (2008) (dashed and solid lines, respectively). Open circle shows results from a cluster sample with galaxy dynamics mass measurements (Rines et al., 2007). Finally, the open square shows the results from Reiprich and Böhringer (2002) (approximately the lower bound of recently published X-ray cluster measurements).

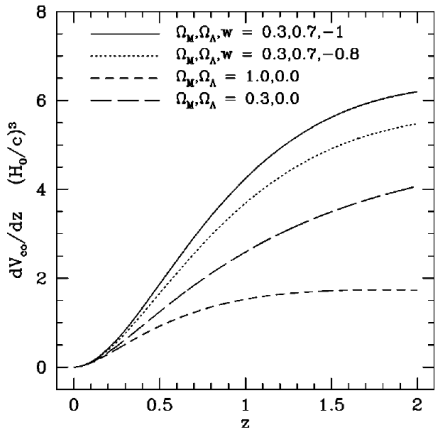


Figure 3.4: From Voit (2005). Redshift dependence of the comoving volume in various cosmologies. The parameter w describes the dark energy equation of state.

the constraints on dark energy in Fig. 3.6.

The redshift dependence of the volume factor in different cosmological models is shown in Fig. 3.4. A big problem with a redshift survey, necessary for this kind of analysis, is the fact that also high-redshift clusters have to be detected and their mass has to be determined accurately. The mass-observable relations are also evolving with time, which makes the calibration or self-calibration more difficult. The most recent analysis of an redshift-evolving mass function was done by Vikhlinin et al. (2009), who used a set of 37 galaxy clusters with an average redshift of $z = 0.55$ and a sample of 40 low-redshift clusters with a redshift of $z \approx 0.05$. All cluster where observed in X-rays. The obtained cosmological constraints in the Ω_m–σ_8 plane are shown in Fig. 3.5 and

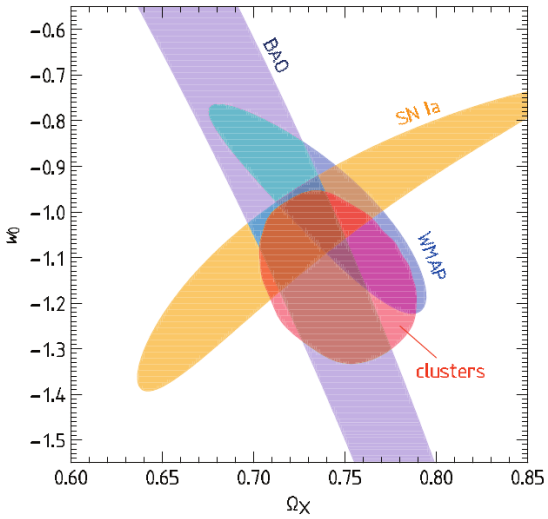


Figure 3.6: From Vikhlinin et al. (2009). Comparison of the dark energy constraints from X-ray clusters and from other individual methods, like supernovae (SN), baryon-acoustic oscillations (BAO), and WMAP. In this figure, $\Omega_\Lambda = 1 - \Omega_m$ and w_0 describes the dark energy equation of state today.

3.3.3 Individual systems as cosmological laboratories

The derivation of cosmological parameters, by means of the cluster mass function is an interesting topic, but clusters do not provide the only possibility to do so. Of course clusters give valuable input when combined with other datasets, but especially when compared to observations of distant supernovae or the CMB, the constraints from clusters alone appear weaker. The real domain of clusters as cosmic tracers lies in individual systems, thereby comparing observations directly to numerical simulations. In the following, we will list the most interesting properties which, when observed reliably, can be used to understand the physics involved in the formation of clusters.

Mass-density profiles

Numerical simulation show that the Navarro, Frenk and White (NFW) profile provides an accurate fit to the simulated density field (Navarro et al., 1996). The parametric form of this profile is given by

$$\rho_{\mathrm{NFW}}(r) = \frac{\rho_s}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}. \tag{3.57}$$

This density profile has two parameters, the characteristic density ρ_s and the scaling radius r_s, which describes the transition between the steep outer profile an the flattening towards the core. This profile in different variations is now largely used to model the mass content in galaxy clusters. More recently Merritt et al. (2006) have shown that the Einasto profile (Einasto and Haud, 1989, and references therein) fits the density field in simulations even better. Its functional form is given by

$$\rho_{\mathrm{E}}(r) = \rho_e \exp\left\{-d_n\left[(r/r_e)^{1/n} - 1\right]\right\}, \tag{3.58}$$

where d_n is defined as a function of n such, that ρ_e becomes the density at radius r_e with the help of the complete gamma function

$$\Gamma(3n) = 2\gamma(3n, d_n) \tag{3.59}$$

and the incomplete gamma function

$$\gamma(3n, x) = \int_0^x e^{-t} t^{3n-1}\, dt. \tag{3.60}$$

Observations of matter-density profiles are still controversial since they usually cannot exclude the, within the standard model unfavoured, SIS profile (Shu et al., 2008) in the strong-lensing regime and recently Broadhurst et al. (2008) found NFW parameters in tension with the predictions of ΛCDM.

In the context of dark matter density profiles it seems to be the case that simulations of clusters converge towards a decisive result. Therefore, more accurate observational methods have to be developed to confirm or reject these predictions from simulations.

Temperature profiles

This situation changes when more involved gas physics becomes important, as is the case in the temperature profiles of observed and simulated galaxy clusters. We already mentioned the problem of cool-core clusters (see e.g. Peterson and Fabian, 2006, for a review), where the rapid cooling of the high-density core is suppressed due to a not fully understood heating process. Furthermore, simulated clusters overestimate the star formation rate with respect to what is observed. Also the observed scaling relations of X-ray observables show a significant deviation from self-similar models (Pratt et al., 2009; Sun et al., 2009; Boehringer et al., 2009) and suggest the inclusion of sophisticated models of gas physics, which have to be tested and calibrated against observations.

Extreme dynamical conditions

An ideal kinematic test site for numerical simulations are merging events, like in the case of the "Bullet Cluster" 1E 0657-558 (Clowe et al., 2004b, 2006), the "Baby Bullet" MACS J0025.4-1222 (Bradač et al., 2008) or "The Cosmic Train Wreck" Abell 520 (Mahdavi et al., 2007a).

Milosavljević et al. (2007); Springel and Farrar (2007); Mastropietro and Burkert (2008) tried to simulate mergers, similar to the "Bullet Cluster" and only the latter work is able to reproduce the observed X-ray and lensing features of the cluster. They derive an initial velocity of the merging subclump (bullet) of the order of ≈ 3000 km s^{-1}. Recently, Lee and Komatsu (2010) show that the largest available cosmological simulation MICE (Crocce et al., 2010) disagrees at a high significance level with the existence of such a system. This fact indicates, that also our understanding of merging events within the framework of the standard model is still incomplete.

Strong lensing features

Another puzzle where simulations disagree with actual observations of clusters is related to the production of strong-lensing features. The so-called arc-statistics problem, first addressed by Bartelmann et al. (1998) and since then controversially discussed (Wambsganss et al., 2004; Horesh et al., 2005; Li et al., 2005; Fedeli et al., 2006, 2008) states the fact that the observed number of arcs in the sky exceeds the predicted production of arcs in the standard model by roughly an order of magnitude. A similar problem was addressed by Broadhurst et al. (2008) and Zitrin et al. (2010), who claim that the size of observed Einstein radii

are also incompatible with ΛCDM. This last example of cosmology with galaxy clusters via their strong lensing features shows instructively how the interplay between simulations and observations is fruitful to improve the quality of scientific statements on both sides. The arc-statistics problem, though conceptually simple, is unsolved for a decade by now and helped to understand better the effects of triaxiality, gas physics and merging events on the lensing properties of galaxy clusters (Oguri et al., 2003; Torri et al., 2004; Puchwein et al., 2005; Fedeli et al., 2006; Fedeli and Bartelmann, 2007; Puchwein and Hilbert, 2009; Meneghetti et al., 2010).

With our work that we present from now on, we hope to contribute to the question of the mass-density distribution of galaxy clusters and to the interpretation of their strong-lensing properties.

4 A joint reconstruction method

There are several methods to recover the properties of a galaxy cluster, like the total mass, the mass distribution and the baryon fraction, from observations. Among them are classical methods, which use the kinematics of cluster member galaxies, while reconstructions based on X-ray observations investigate the diffuse, thermal emission of the Intra-Cluster-Medium (ICM). Gravitational lensing traces the mass distribution of the cluster in a very direct way and distinguishes between a weak and a strong-lensing regime.
Usually, all these methods are used separately and we intend to demonstrate the possibility to combine the particular strengths of individual methods into a joint reconstruction, thereby eliminating the typical limitations of reconstructions based on a single class of observational constraints.
After describing the basic principle of such an approach in Sec. 4.1, we focus on the weak-lensing regime in Sec. 4.2, with an excursion towards the special case of flexion in Sec. 4.3. The strong-lensing regime is covered in Sec. 4.4. These are the areas, which can be combined with our reconstruction method in its actual state. We summarise the advantages and drawbacks of a joint reconstruction method, as we present it in this work, in Sec. 4.5 and give a short overview on other observables of galaxy clusters, as they could be included in our method, in Sec. 4.6.

4.1 A grid-based maximum-likelihood approach

The original idea of such a method goes back to Bartelmann et al. (1996) and has been extensively developed since then (Bradač et al., 2004, 2005, 2009; Cacciato et al., 2006). There were also several other attempts to combine weak and strong lensing (Diego et al., 2007; Deb et al., 2008; Oguri et al., 2009), weak lensing and flexion (Leonard et al., 2007; Okura et al., 2008), strong lensing and kinematics (Sand et al., 2008) or lensing, X-Ray and Sunyaev-Zel'dovich observations (Puchwein and Bartelmann, 2006, 2007; Mahdavi et al., 2007b) into a joint reconstruction method.
The method presented in this work delivers for the first time a consistent framework to combine weak lensing, flexion and several strong-lensing constraints, in order to recover optimally the properties of a galaxy cluster and justifies the development of a new algorithm. Furthermore, combined methods usually suffer from severe problems regarding implementation and runtime due to the complexity of the problem, which forces them to accept compromises regarding the inclusion of the constraints or in the error estimation. We will show how to overcome these limitations in our implementation.
Another advanced feature of our method, despite its ability to combine a wide range of constraints, is the fact that it is fully nonparametric. This means, that no a-priori assumptions on e.g. the mass profile of the galaxy cluster are necessary. The reconstruction relies purely on the observational data.

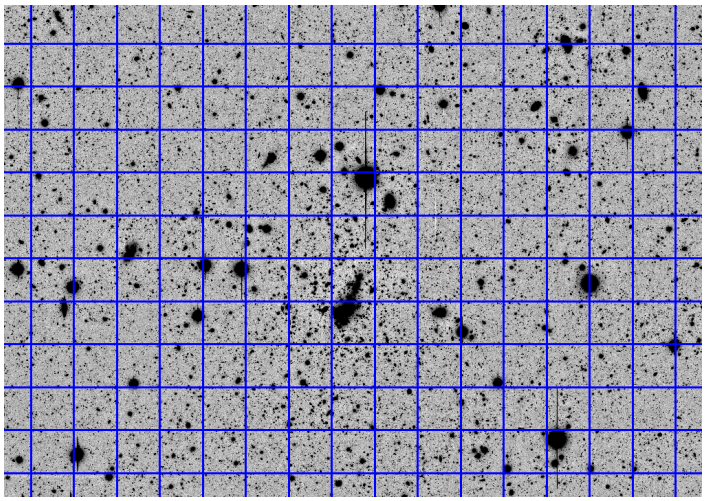


Figure 4.1: The discretised observational field. The full observational field of the galaxy cluster is divided into a grid and the basic quantity of the reconstruction, the lensing potential, is defined in each individual grid cell. Here you see a slight zoom on a SUBARU r-band exposure of Abell 1703 with a field size of $\sim 26\farcm8 \times 17\farcm7$

4.1.1 The discretised lensing potential

The first decision, which has to be made while developing a combined reconstruction method, is what quantity actually describes the cluster best in this context and should therefore be reconstructed. Possible candidates would be the surface-mass density or, related to it, the convergence of the cluster. These quantities have clear physical meanings, but nevertheless they are not ideally suited for a combined reconstruction, since they are not fully general regarding the observables. As we have seen in Sec. 2.2.2, all effects changing the shapes of background objects can be described as spin-fields by applying derivative operators to the lensing potential ψ. Since ψ is the rescaled and projected Newtonian potential of the galaxy cluster, it can be related to kinematic, X-ray or SZ observations. Those points convinced us to choose the lensing potential as the reconstruction quantity.
For reasons, which will become more clear in the next section, we will not reconstruct the continuous lensing potential as a function, but its discretised equivalent. As indicated in Fig. 4.1, we divide the field, which contains all observational constraints, into a grid. The lensing potential is then defined on each individual grid cell. This procedure is key to obtain a fully nonparametric reconstruction, as we will show now.

4.1.2 Combining different constraints

To recover the discretised lensing potential, we perform model-fitting in a least-squares fashion (see e.g. Press et al., 2007; Verde, 2009).
We define a multi-component χ^2-function

$$\chi^2(\psi) = \chi^2_1(\psi) + \chi^2_2(\psi) + \chi^2_3(\psi) + ..., \tag{4.1}$$

where all individual terms depend on the lensing potential ψ and refer to different observational constraints. We minimise the overall χ^2-function with respect to the lensing potential in every grid cell and obtain the discretised lensing potential at every grid position with index

l as a result

$$\frac{\partial \chi^2(\psi)}{\partial \psi_l} \stackrel{!}{=} 0. \tag{4.2}$$

Every discrete value of ψ is a parameter in this optimisation procedure and if the grid resolution is suitably high, we will obtain a smooth representation of the galaxy cluster. Therefore, we will continue calling the method nonparametric and we deal with concrete realisation of Eq. 4.2 in Sec. 5.4.
Since we are fitting a very high-dimensional model to a probably ill-defined problem, we will have to deal with the issue of overfitting and may introduce additional regularisation terms in Eq. 4.1. This is discussed in Sec. 5.5.

4.1.3 Connecting the observables

From now on it is clear how to proceed. We connect the observational quantities, from lensing or other constraints, to the discretised lensing potential and while doing so, define suitable terms to be inserted in Eq. 4.1.
While having a look at Eqs. 2.27-2.31, differential operators on the reconstruction grid are indispensable and we obtain them by following this recipe:

- We use Taylor's theorem (Rannacher, 2003a,b) to derive the forward (Δ_h), backward (∇_h) and central (δ_h) differences to express the partial derivatives in x-direction, with the horizontal distance between neighbouring grid-cell centres being defined as h[1]

$$\frac{\partial f(x,y)}{\partial x} = \frac{\Delta_{h_x}[f](x,y)}{h} - \mathcal{O}(h) = \frac{f(x+h,y)+f(x,y)}{h} - \mathcal{O}(h) \tag{4.3}$$
$$\frac{\partial f(x,y)}{\partial x} = \frac{\nabla_{h_x}[f](x,y)}{h} - \mathcal{O}(h) = \frac{f(x,y)-f(x-h,y)}{h} - \mathcal{O}(h) \tag{4.4}$$
$$\frac{\partial f(x,y)}{\partial x} = \frac{\delta_{h_x}[f](x,y)}{h} - \mathcal{O}(h^2) = \frac{f(x+\frac{1}{2}h,y)-f(x-\frac{1}{2}h,y)}{h} - \mathcal{O}(h^2) \tag{4.5}$$
$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{\Delta^2_{h_x}[f](x,y)}{h^2} - \mathcal{O}(h) = \frac{f(x+2h,y)-2f(x+h,y)+f(x,y)}{h^2} - \mathcal{O}(h) \tag{4.6}$$
$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{\nabla^2_{h_x}[f](x,y)}{h^2} - \mathcal{O}(h) = \frac{f(x-2h,y)-2f(x-h,y)+f(x,y)}{h^2} - \mathcal{O}(h) \tag{4.7}$$
$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{\delta^2_{h_x}[f](x,y)}{h^2} - \mathcal{O}(h^2) = \frac{f(x+h,y)-2f(x,y)+f(x-h,y)}{h^2} - \mathcal{O}(h^2). \tag{4.8}$$

  With equivalent formulas for the partial derivatives in y-direction.

- By using these finite differences, we can express the deflection angle, convergence, shear and flexion fields via the discrete representation of ψ. For example three of the

[1] Not ot be confused with the Hubble parameter.

nine possible quotient combinations for the convergence are given by

$$\begin{aligned}\kappa_{\delta\delta}(x,y) \approx h^{-2}\Big\{&-\frac{2}{3}\psi(x,y)-\frac{1}{6}\psi(x-h,y)-\frac{1}{6}\psi(x+h,y)\\&-\frac{1}{6}\psi(x,y-h)+\frac{1}{3}\psi(x-h,y-h)+\frac{1}{3}\psi(x+h,y-h)\\&-\frac{1}{6}\psi(x,y+h)+\frac{1}{3}\psi(x-h,y+h)+\frac{1}{3}\psi(x+h,y+h)\Big\},\end{aligned} \tag{4.9}$$

$$\begin{aligned}\kappa_{\delta\nabla}(x,y) \approx h^{-2}\Big\{&\frac{1}{2}\psi(x-h,y)-\frac{1}{2}\psi(x,y)+\frac{1}{2}\psi(x+h,y)\\&-\psi(x,y-h)+\frac{1}{2}\psi(x,y-2h)\Big\},\end{aligned} \tag{4.10}$$

$$\begin{aligned}\kappa_{\Delta\nabla}(x,y) \approx h^{-2}\Big\{&\psi(x,y)-\psi(x+h,y)+\frac{1}{2}\psi(x+2h,y)\\&-\psi(x,y-h)+\frac{1}{2}\psi(x,y-2h)\Big\},\end{aligned} \tag{4.11}$$

where the subscripts indicates the quotient (forward, backward or central) for the x and y-direction, respectively.

- This calculation is visualised for all quantities, which we will need to define χ^2-functions on the grid in Figs. 4.2, 4.3, 4.4, 4.5 and 4.6.
- When assuming, that all lensing potential values on the grid are represented by a vector ψ_i of length N, where N is total number of grid cells, we can write all relevant lensing quantities at a grid position with index i as matrix multiplications

$$\alpha_i^1 = \mathcal{D}_{ij}^1\psi_j \tag{4.12}$$
$$\alpha_i^2 = \mathcal{D}_{ij}^2\psi_j \tag{4.13}$$
$$\kappa_i = \mathcal{K}_{ij}\psi_j \tag{4.14}$$
$$\gamma_i^1 = \mathcal{G}_{ij}^1\psi_j \tag{4.15}$$
$$\gamma_i^2 = \mathcal{G}_{ij}^2\psi_j \tag{4.16}$$
$$F_i^1 = \mathbf{F}_{ij}^1\psi_j \tag{4.17}$$
$$F_i^2 = \mathbf{F}_{ij}^2\psi_j \tag{4.18}$$
$$G_i^1 = \mathbf{G}_{ij}^1\psi_j \tag{4.19}$$
$$G_i^2 = \mathbf{G}_{ij}^2\psi_j, \tag{4.20}$$

where all matrix quantities are sparse band matrices. Only those entries per row with index i, which represent a coefficient as shown in Figs. 4.2– 4.6 are non-zero.

This gives us all ingredients at hand to look at several possibilities, which might contribute a term in Eq. 4.1.

4.2 Weak lensing

We will focus on weak lensing first and the reader is referred to the excellent review of Bartelmann and Schneider (2001) for a more complete overview. The main advantage of weak

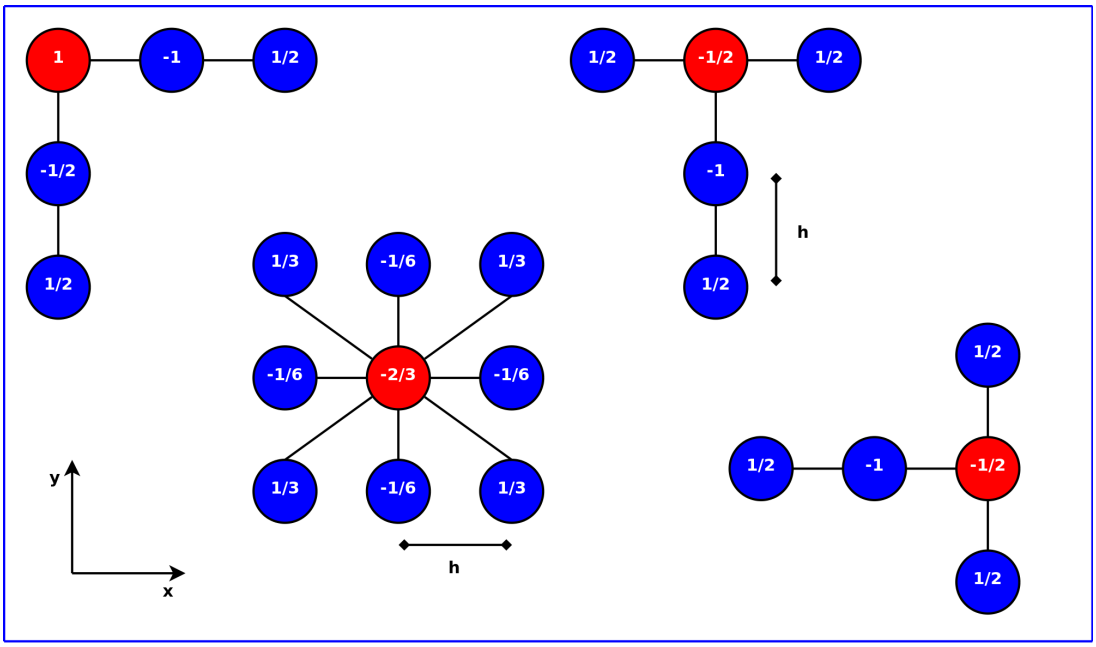


Figure 4.2: The finite-differences scheme, which belongs to Eqs. 4.9, 4.10, 4.11, 4.14. Marked in red is the grid cell, for which the convergence is calculated. The coefficients are given in units of h^{-2}.

gravitational lensing for cluster mass reconstruction is the fact, that it allows for a complete reconstruction of the whole observed field as long as shape measurements of background galaxies are available. We have seen in Sec. 2.2.2 that local averages of these measurements are approximately equal to the reduced shear of the lens.
Hence a sufficient number of ellipticity values has been measured, a second advantage of a weak-lensing reconstruction is the fact, that it allows for a nonparametric recovery of the mass distribution.
Kaiser and Squires (1993) were the first to notice, that a direct mapping between shear and convergence exists.
Starting from

$$2\gamma = \partial\partial\psi, \tag{4.21}$$

we apply the spin-lowering operator to both sides and use the definition of the convergence

$$2\partial^*\gamma = \partial\partial^*\partial\psi \tag{4.22}$$

$$\partial^*\gamma = \partial\kappa. \tag{4.23}$$

This Cauchy-Riemann equation (Straumann, 1997) is solved with the help of the Green's function $\tilde{G}$ of the spin-raising operator, as defined by

$$\partial\tilde{G}(\theta) = \partial\left(\frac{1}{\theta\pi}\right) = \delta(\theta), \tag{4.24}$$

up to an additive constant by the complex field

$$\tilde{\kappa}(\theta) - \tilde{\kappa}_0 = \partial^{-1}\partial^*\gamma, \tag{4.25}$$

with the inverse differential operator

$$\partial^{-1}(\cdot) = -\frac{1}{\pi}\frac{1}{\theta^2} * (\cdot), \tag{4.26}$$

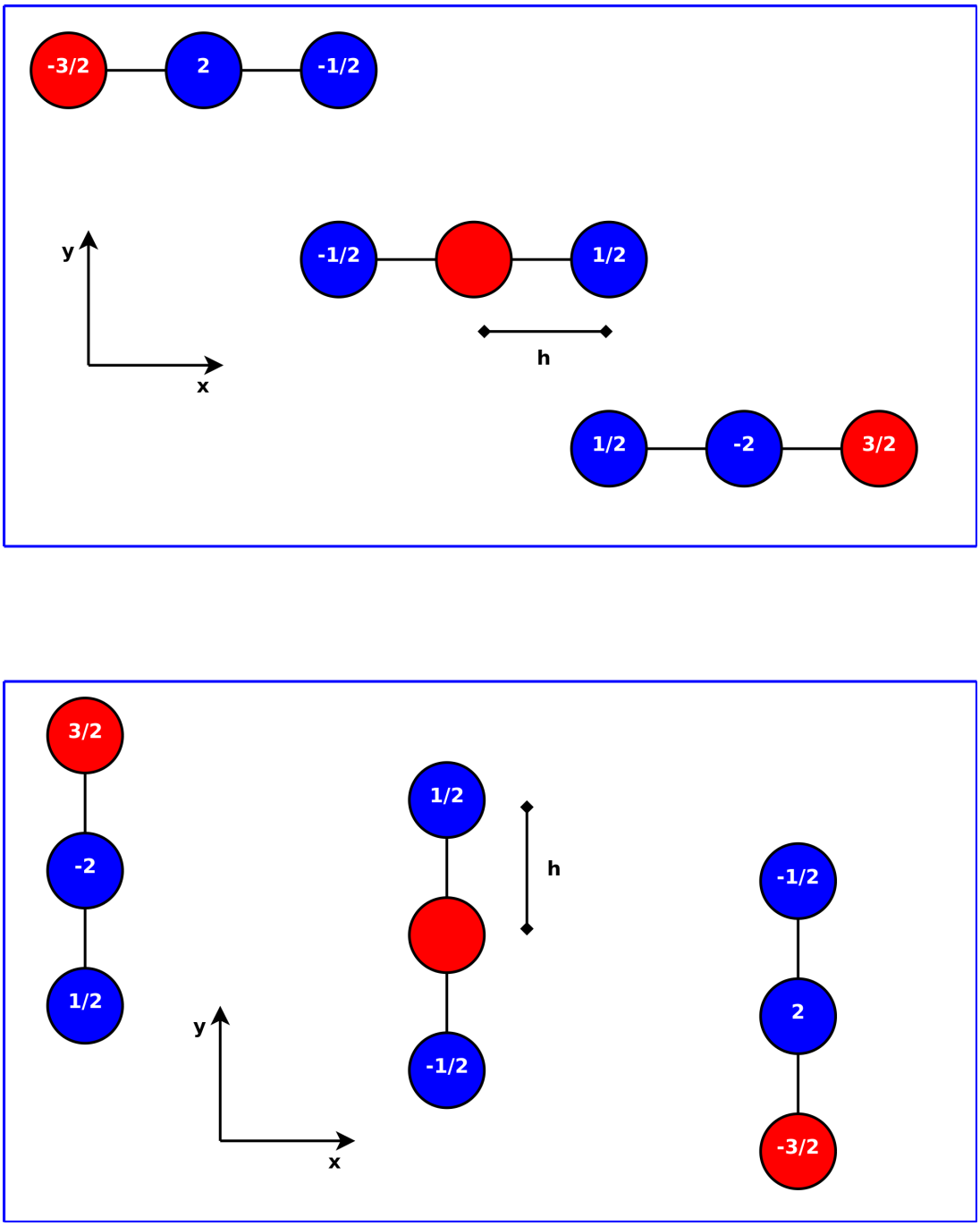


Figure 4.3: The finite-differences scheme for the first component of the deflection angle (Eq. 4.12) is shown in the *top panel*. The *bottom panel* shows the scheme for the second component (Eq. 4.13). Marked in red is the grid cell, for which the deflection angle is calculated. The coefficients are given in units of h. Please note, that a higher-order Taylor expansion has been used, which is why the schemes in the figure do not match exactly with Eqs. 4.3, 4.4.

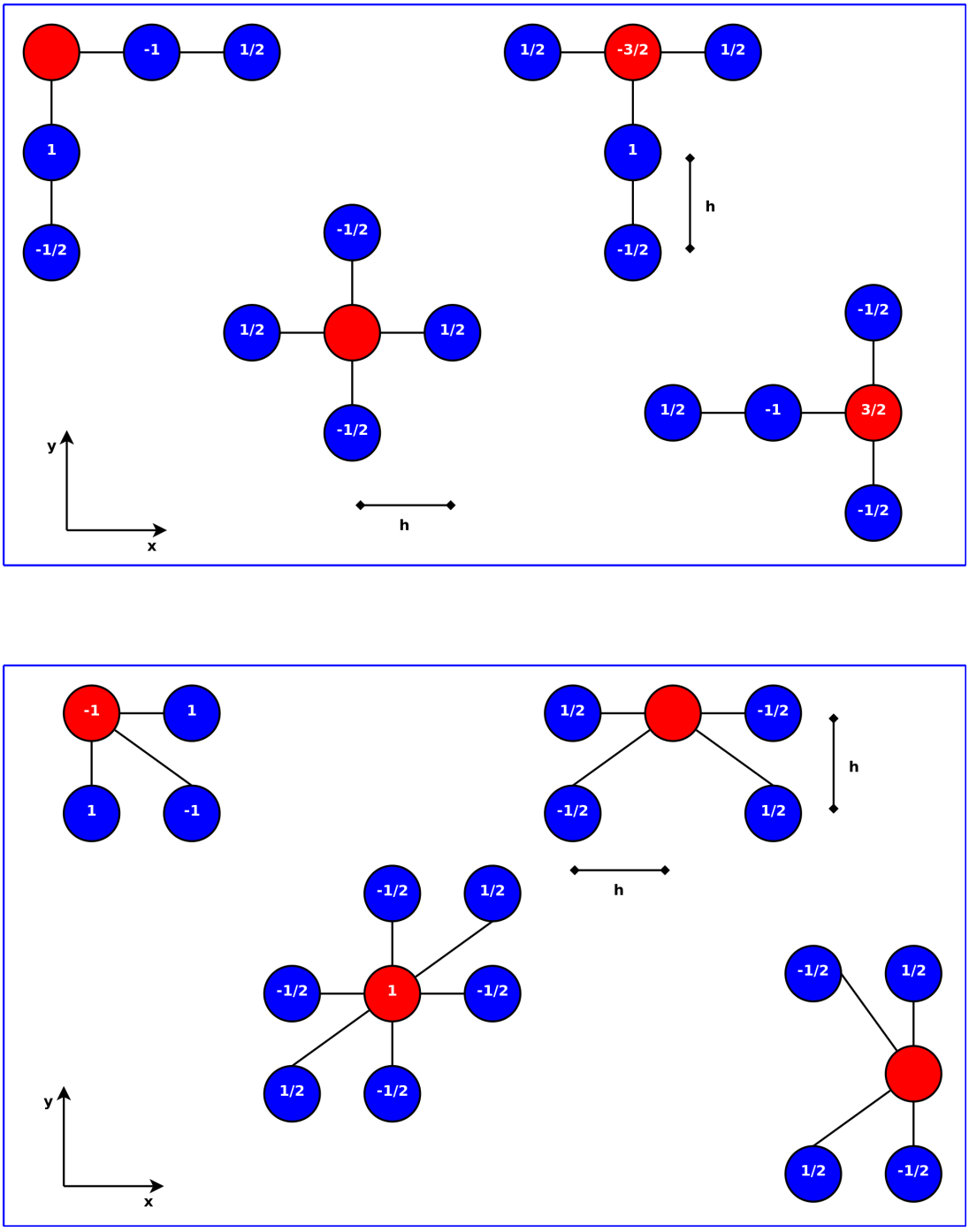


Figure 4.4: The finite-differences scheme for the first component of the shear (Eq. 4.15) is shown in the *top panel*. The *bottom panel* shows the scheme for the second component (Eq. 4.16). Marked in red is the grid cell, for which the shear is calculated. The coefficients are given in units of h^{-2}.

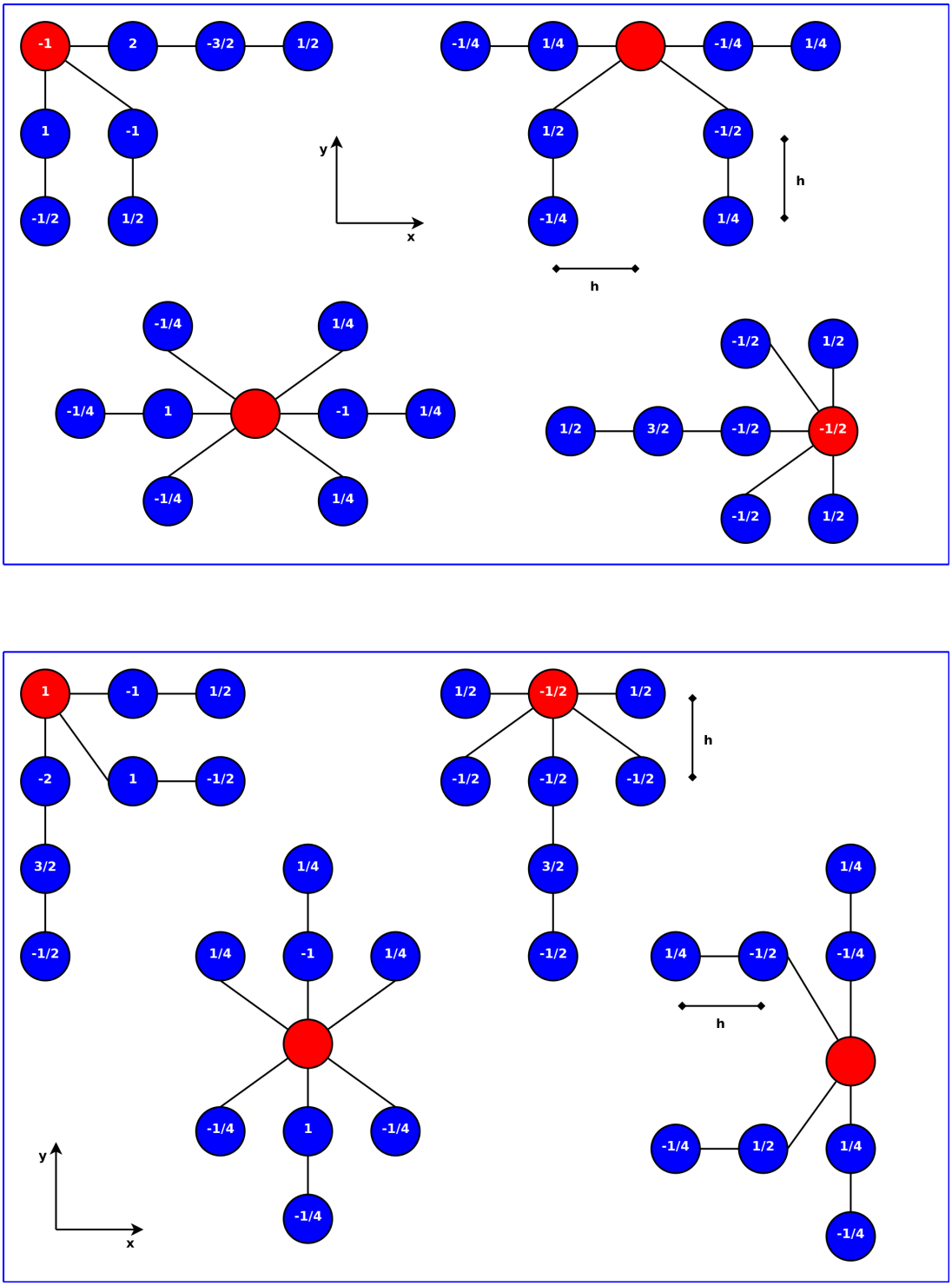
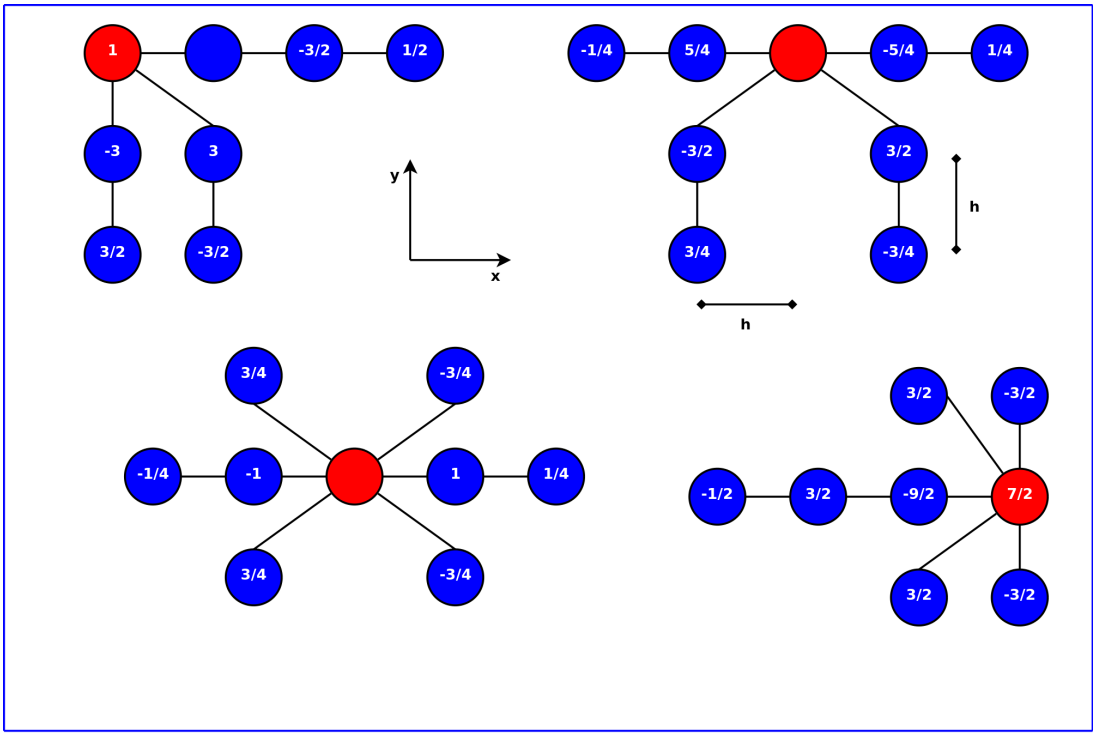


Figure 4.5: The finite-differences scheme for the first component of the F-flexion (Eq. 4.17) is shown in the *top panel*. The *bottom panel* shows the scheme for the second component (Eq. 4.18). Marked in red is the grid cell, for which the F-flexion is calculated. The coefficients are given in units of h^{-3}.



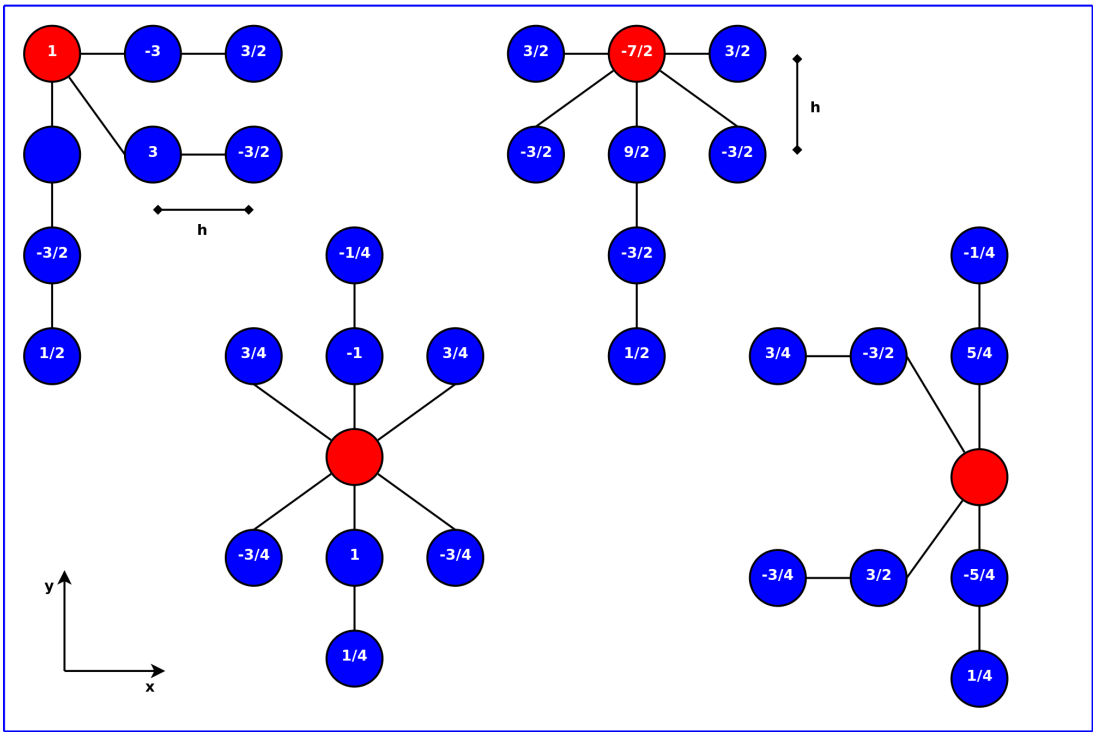


Figure 4.6: The finite-differences scheme for the first component of the G-flexion (Eq. 4.19) is shown in the *top panel*. The *bottom panel* shows the scheme for the second component (Eq. 4.20). Marked in red is the grid cell, for which the G-flexion is calculated. The coefficients are given in units of h^{-3}.

and $*$ being the convolution.
Since lensing effects alone cannot create any imaginary parts in this field (Crittenden et al., 2002), the real part of $\tilde{\kappa}(\theta) - \tilde{\kappa}_0$ recovers the convergence as shown by Kaiser and Squires (1993)

$$\kappa(\boldsymbol{\theta}) - \kappa_0 = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \Re[\mathcal{D}^*(\boldsymbol{\theta} - \boldsymbol{\theta}')\gamma(\boldsymbol{\theta}')], \tag{4.27}$$

with a complex convolution kernel $\mathcal{D}$.
A few comments on Eq. 4.27:

1. $\gamma(\boldsymbol{\theta})$ can at best be determined at discrete points within the reconstruction field, where shape measurements are available. Thus, smoothing with e.g. a Gaussian kernel $\mathcal{D}$ is required (Seitz and Schneider, 1995), which introduces covariances in the resulting mass map (van Waerbeke, 2000; Lombardi and Bertin, 1998).

2. Not the shear, but the reduced shear can be related to observations. This problem has been solved by Seitz and Schneider (1995), while using an iterative approach.

3. Reduced-shear measurements are only available on a finite field, but Eq. 4.27 requires an integration over the whole sky. This problem has been fixed by a finite-field-inversion technique (Seitz and Schneider, 1995).

4. A problem, which cannot be fixed in a fully satisfactory way by using a direct-inversion technique, is that the convergence can only be recovered up to an additive constant. This problem is called the mass-sheet degeneracy (Falco et al., 1985).

4.2.1 Defining a χ^2-function

A way to deal with the last problem is to perform the weak-lensing reconstruction within an maximum-likelihood approach, as we have presented it in Sec. 4.1.3. This allows for the introduction of terms, which break the mass-sheet degeneracy.
We know from Eq. 2.58, that we can relate the expectation value of the ellipticity in one grid cell to the reduced shear of the lens. This suggests the following definition of the weak-lensing χ^2-function

$$\chi^2_{\mathrm{w}} = \sum_{i,j} (\varepsilon - g(\psi))_i \mathcal{C}^{-1}_{ij} (\varepsilon - g(\psi))_j, \tag{4.28}$$

where the indices i and j run over all cells of the reconstruction grid.
In the case of $|g| \leq 1$, we find

$$\chi^2_{\mathrm{w}}(\psi) = \sum_{i,j} \left(\varepsilon - \frac{Z(z)\gamma(\psi)}{1 - Z(z)\kappa(\psi)} \right)_i \mathcal{C}^{-1}_{ij} \left(\varepsilon - \frac{Z(z)\gamma(\psi)}{1 - Z(z)\kappa(\psi)} \right)_j . \tag{4.29}$$

The covariance matrix $\mathcal{C}_{ij}$ describes the noise in the reduced-shear measurements and its concrete form depends on the averaging scheme that we apply. Generally, it will depend on the standard deviation σ of the averaging process within each grid cell. This standard deviation is composed of three different contributions

$$\sigma = \sigma_{\mathrm{int}} + \sigma_{\mathrm{sys}} + \sigma_{\mathrm{l}}, \tag{4.30}$$

with the noise due to intrinsic ellipticity σ_{int}, systematic error (measurement noise) σ_{sys} and lensing noise σ_{l}. Under lensing noise we understand the noise which arises from the fact

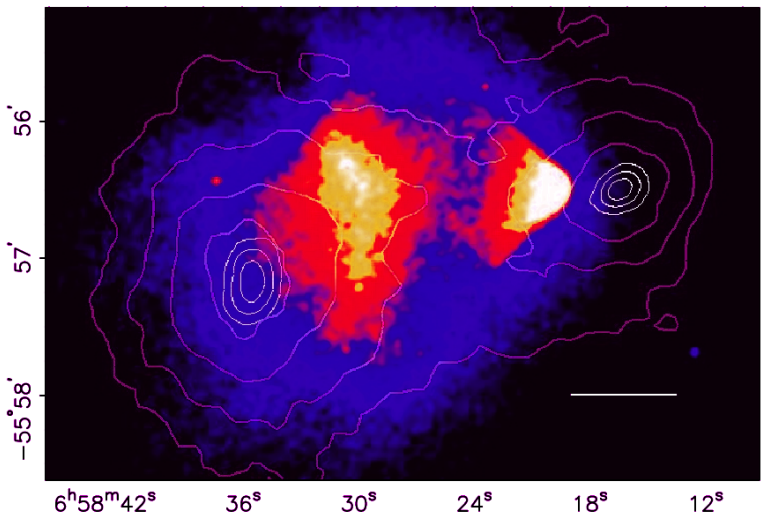


Figure 4.7: From Clowe et al. (2006). An X-ray view on the "Bullet Cluster" using Chandra. Encoded in the background-colour map is the X-ray luminosity. The purple contours represent the convergence levels, obtained by weak lensing and the white contours show the positions of the convergence centres of the two halos, including error intervals. This figure shows a clear offset between the X-ray emitting gas components and the dark matter in the merging cluster 1E 0657-558. The white bar indicates a scale of 200 kpc at the cluster's redshift of $z_c = 0.296$.

that the galaxies over which we average, are spatially separated, so the properties of the lens change inside a grid cell. We assume that for a sample of sufficient size, the intrinsic ellipticity noise tends to zero, which leaves the lensing and the systematic noise and we will calculate $\mathcal{C}$ for the concrete averaging scheme of our reconstruction method in Sec. 5.3.1.

4.2.2 Problems with weak lensing

The application of weak-lensing reconstruction techniques has become standard in the analysis of galaxy clusters by now. Outstanding breakthroughs were achieved while using this technique, with the strong evidence that the "Bullet Cluster" (1E 0657-558) contains collisionless dark matter (Clowe et al., 2004b, 2006; Bradač et al., 2006) being presumably the most prominent among them (Fig. 4.7). Nevertheless, weak lensing relies on a linearised lens equation and its assumptions become less and less accurate towards the central regions of efficient lenses. Whenever the shapes of background sources become distorted in a different way, than the effects of convergence and shear can explain, weak lensing alone is not able to explain them any more. Obviously, those effects are indeed observed in the form of arclets, arcs or multiple-image systems in several galaxy clusters (see e.g. Broadhurst et al., 2005, as an extreme example).

4.3 Flexion

Flexion is sensitive to the third derivatives of the lensing potential and is therefore ideal to trace the structure of the deflecting galaxy cluster on a much smaller scale than this is possible with shear measurements alone. Once a flexion signal was measured with one

of the techniques that we have described in Sec. 2.2.3, it can contribute important information in the transient area between the weak and the strong-lensing regime and unravel sub-structures, to which the shear alone is blind.

Nonparametric flexion reconstructions

Direct-inversion formulas for flexion are derived from the definitions of the two complex fields

$$2F = \partial\partial\partial^*\psi \tag{4.31}$$
$$2G = \partial\partial\partial\psi, \tag{4.32}$$

by inserting the convergence

$$F = \partial\kappa \tag{4.33}$$
$$\partial^*\partial^*\partial^*G = \partial^4\kappa. \tag{4.34}$$

Solving those equations in the same as Eq. 4.23 leads to two new complex fields, which real part is the convergence and which imaginary part is a quantity, that cannot be due to lensing and is therefore used to check for systematics in the analysis. These systematics are often called B-Modes (Crittenden et al., 2002).

$$(\kappa + \mathrm{iB})_F = \partial^{-1}F \tag{4.35}$$
$$(\kappa + \mathrm{iB})_G = \partial^{-4}\partial^*\partial^*\partial^*G \tag{4.36}$$

Parametric flexion fitting

Another way of using the flexion signal in mass reconstructions was chosen by Leonard et al. (2007). To every flexion measurement an analytic mass profile like e.g. the Singular-Isothermal-Sphere (SIS) (Binney and Tremaine, 2008) is fitted. The two complex flexion fields, produced by such an analytic profile, were derived by Bacon et al. (2006)

$$F(\theta) = -\frac{\theta_\mathsf{E}}{2\theta^2}e^{\mathrm{i}\phi}, \tag{4.37}$$
$$G(\theta) = \frac{3\theta_\mathsf{E}}{2\theta^2}e^{3\mathrm{i}\phi}, \tag{4.38}$$

in polar representation and with the Einstein radius

$$\theta_\mathsf{E} = 4\pi\left(\frac{\sigma_\mathsf{v}}{c}\right)^2\frac{D_\mathsf{ds}}{D_\mathsf{s}}. \tag{4.39}$$

σ_v is the line-of-sight velocity dispersion. The convergence map, like shown in Fig. 4.8, is then obtained by summing over all SIS-contributions with index i

$$\kappa(\theta) = \sum_i \frac{\theta_{\mathsf{E},i}}{\theta_i - \theta}. \tag{4.40}$$

4.3.1 Defining a χ^2-function

We now define the flexion χ^2-contributions of Eq. 4.1, which are quite similar to the case of shear. By assuming, that local averages of flexion measurements (f and g) in individual grid

cells are equal to the flexion of the model, based on the lensing potential, we define

$$\chi^2_F(\psi) = \sum_{i,j} \left(f - Z(z)F(\psi)\right)_i \mathcal{C}^{-1}_{ij} \left(f - Z(z)F(\psi)\right)_j \tag{4.41}$$

$$\chi^2_G(\psi) = \sum_{i,j} \left(g - Z(z)G(\psi)\right)_i \mathcal{C}^{-1}_{ij} \left(g - Z(z)G(\psi)\right)_j . \tag{4.42}$$

The same considerations regarding the covariance matrices apply. Again, their exact shapes depend on the averaging scheme which is applied. This point leads us directly to the first of some problems, which are still connected with the use of gravitational flexion.

4.3.2 Problems with flexion

Several things are still unclear and problematic in the use of flexion, which is mostly due to the fact, that the whole field is relatively young and still evolving. The following problems should be noted:

1. As a second-order contribution, the flexion signal is lower in amplitude and unlike in ellipticity measurements (Brainerd et al., 1996) the level of intrinsic flexion is not well-known up to now. Therefore, it is unclear how large the averaging sample for individual grid cells in Eqs. 4.41, 4.42 should be. Goldberg and Bacon (2005) estimated that the intrinsic noise level of the flexion signal is lower than in the shear signal, which is a promising result and suggests that less galaxies have to be averaged to obtain a signal. This might increase the resolution of reconstructions involving flexion.

2. Schneider and Er (2008) raised the point, that not the originally defined flexion fields F and G are actually measurable by observations, but the reduced flexion fields

$$G_1 = \frac{F + gF^*}{(1-\kappa)} \quad \text{and} \quad G_3 = \frac{G + gF^*}{(1-\kappa)}. \tag{4.43}$$

   Our method would not suffer from that point, since it only demands for a slight redefinition of the flexion χ^2-function, similar to the case of the reduced shear.

3. As we already pointed out in Sec. 2.2.3, the measurement of flexion from images is extremely difficult and even if a reliable measurement is given, also the second-order lens equation, which gives rise to the flexion fields, fails for highly non-linear critical lensing events near the cluster core. Furthermore, almost no background galaxies are observed at all in this region, because of foreground light of the brightest cluster galaxy (BCG) and demagnification effects. This problem also manifests itself in Fig. 4.8, where a flexion reconstruction of `Abell 1689` is compared to a reconstruction, which also includes strong-lensing constraints.

4.4 Strong lensing

As we have seen, both reconstruction methods based on shear or flexion, have problems in describing the innermost core of a galaxy cluster due to a lack of constraints and an approximated lens equation. Indeed, tests with simulations show, that pure weak lensing reconstructions usually recover a mass-profile which is too shallow in the cluster centre (Cacciato et al., 2006; Merten et al., 2009; Meneghetti et al., 2009). These limitations are cured with

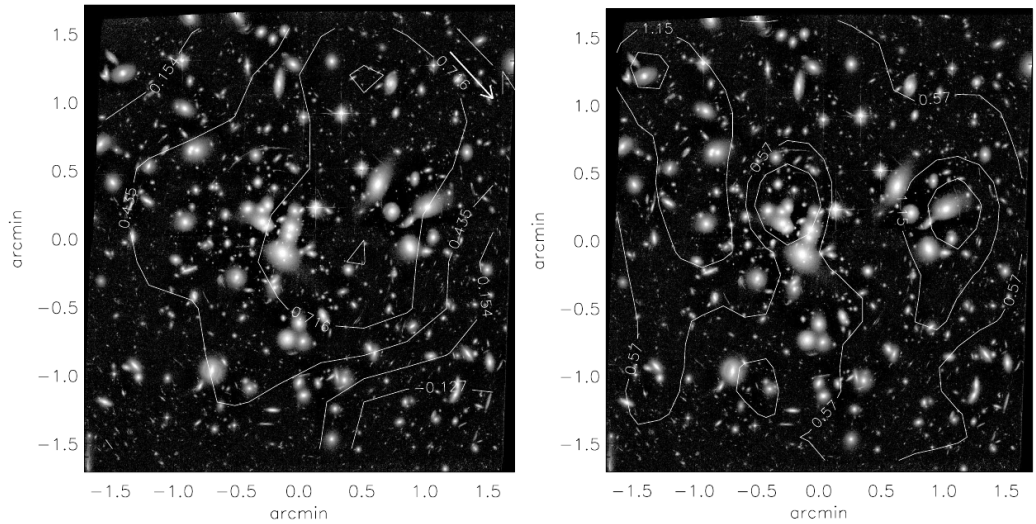


Figure 4.8: From Leonard et al. (2007). `HST/ACS` exposure of `Abell 1689`. *Left Panel:* The contours show the convergence levels, obtained by a combined weak and strong-lensing reconstruction. *Right Panel:* The contours show the convergence levels of a reconstruction based on the F-flexion alone. It can be clearly seen, that the reconstruction follows substructure better, but shows a unexpected under-density in the cluster centre. This was explained in Leonard et al. (2007) as a shot-noise effect, due to a lack of signal in this area.

methods, which rely on the constraints attached to the strong-lensing regime, like multiple-images systems, giant arcs or Einstein rings. Within the regions where those features appear, strong-lensing methods deliver excellent results since the constraints are very decisive as we will see. Nevertheless, the number of those constraints is usually two to three orders of magnitude smaller than in the case of shear or flexion, which means that parametric models are applied to prevent the problem from being under-constrained.

Parametric methods

As the name already implies, parametric, analytic models like the SIS or the Navarro, Frenk and White (NFW) (Navarro et al., 1996) profile are used to parametrise the lens and to fit the observations. As Meneghetti et al. (2007b) have shown, during the modelling process it is very important to account for asymmetries and substructure in the profile of the main dark matter halo (see also Keeton, 2003; Meneghetti et al., 2007a). Therefore, generalised, elliptical profiles are used to model the main halo. Significant mass contributions of single cluster galaxies are modelled separately by e.g. a Pseudo-Isothermal-Elliptical-Mass-Distribution (PIEMD) (Kassiola and Kovner, 1993).
Together, these different mass profiles model the deflector with a parameter vector $\mathbf{p}$ of length N_p, where N_p is the sum of all parameters from the different contributions to the overall mass model.
This parameter vector is then optimised to fit the input observations best. In the case of multiple-image systems we use the lens equation to derive the source position β_i of one image, observed at position θ_i

$$\beta_i = \theta_i - \alpha(\theta_i, \mathbf{p})^2. \tag{4.44}$$

[2]The reader should remember the complex notation, which applies to the lensing quantities.

Of course, the original source position is not known, but we can define an average source position, to which the sample of all images of an multiple-image system is imaged back on average

$$\langle\beta\rangle = \frac{1}{N}\sum_{i=1}^{N}\beta_i, \tag{4.45}$$

where N is the number of images in an individual system. Afterwards, we demand that the deviation of the source position of an individual image from the average source position is as small as possible in a least-squares interpretation

$$\chi^2_{\mathsf{sp}} = \sum_{i=1}^{N}\left(\frac{\beta_i(\theta_i,\mathbf{p}) - \langle\beta\rangle(\theta_i,\mathbf{p})}{\sigma_i}\right)^2, \tag{4.46}$$

where σ_i describes the error in the knowledge of the image position θ_i. The procedure above is called source-plane optimisation. Lens-plane optimisation is defined by

$$\chi^2_{\mathsf{lp}} = \sum_{i=1}^{N}\left(\frac{\theta_i - \theta_i(\beta)}{\sigma_i}\right)^2, \tag{4.47}$$

but the minimisation of this χ^2-function is a lot more expensive regarding CPU time.
Another possible constraint in the strong lensing regime are giant gravitational arcs. As indicated in Fig. 4.9, the position of saddle points in the surface-brightness distribution of arcs gives an estimate on the position of critical lines as they are found in strongly lensing galaxy clusters. If also the redshift of the arc is known, this can be an extremely powerful constraint for the reconstruction. The constraint is translated, in terms of model fitting, again into a likelihood function

$$\chi^2 = \sum_{i=1}^{N}\left(\frac{||\mathbf{O}_i - \mathbf{D}_i(\mathbf{p})||}{\sigma_i}\right)^2, \tag{4.48}$$

where $\mathbf{O}$ describes the predicted position of the critical line, $\mathbf{D}$ the position of the critical line given the model and σ_i the error in the prediction. N is the total number of critical-line predictions.
In the final analysis, the posterior probability, meaning the probability to find the parameters $\mathbf{p}$ given the observational data is optimised by applying Monte-Carlo-Markov-Chains (MCMC) and Bayesian model-selection techniques. Verde (2009) gives a decent review on these methods, especially in cosmological applications.
With the publicly available code `LensTool` (Jullo et al., 2007; Kneib et al., 1996) and the soon[3] publicly available code `CLens` (Halkola et al., 2006, 2008) very powerful tools exist, which apply the constraints mentioned above and allow reliable, parametric reconstructions of galaxy clusters based on strong gravitational lensing.

A nonparametric method, genetic algorithms
Liesenborgs et al. (2006, 2007) suggested a particularly interesting nonparametric, strong-lensing-inversion method based on a genetic algorithm (see e.g. Deb, 2001; Charbonneau, 1995, for reviews). Genetic algorithms are a special class of evolutionary algorithms, which mimic natural selection schemes as they are found in biological systems. The solutions to a certain problem are encoded in a so-called genome[4]. Several trial-solutions are created in an

[3] private communication, the author is involved in the further development of the package
[4] Also the vocabulary of Biology is used in the context of these algorithms

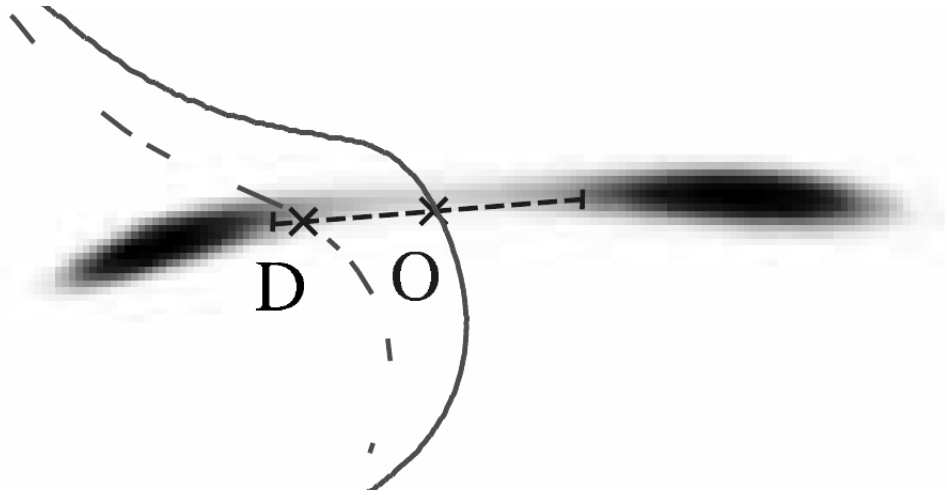


Figure 4.9: From Jullo et al. (2007). The position of the critical line of a galaxy cluster can be predicted from the existence of saddle points in the surface-brightness distribution of gravitational arcs.

initial step and generations of preceding genomes are bred by sexual reproduction, meaning by combining parts of different genomes, or by asexual reproduction, meaning existing genomes are simply copied. To ensure genetic diversity also mutation is applied during the evolutionary process. To ensure the survival-of-the-fittest concept, selection pressure has to be applied by defining a fitness criterion, which gives certain genomes a higher probability to reproduce themselves. Evolution stops when a specified fitness-criterion is reached by the genome. The main advantage of genetic algorithms is their ability to solve under-determined problems, as in the case of nonparametric strong lensing, without any additional constraints. In Liesenborgs et al. (2006, 2007) the reconstructed lens is represented by a dynamic grid, which automatically increases its resolution in high-density areas of the reconstruction during the evolution. To each grid cell a projected Plummer sphere (Plummer, 1911) is assigned, which encodes the mass within that cell. The whole lens is defined by the sum of all these spheres. The lens equation in terms of N Plummer spheres reads

$$\beta(\theta) = \theta - \frac{D_{\mathsf{ds}}}{D_{\mathsf{s}} D_{\mathsf{d}}} \frac{4G}{c^2} \sum_{i=1}^{N} \frac{\theta - \theta_{\mathsf{s},i}}{|\theta - \theta_{\mathsf{s},i}|^2 + \theta_{\mathsf{P},j}^2} M_i, \tag{4.49}$$

where M is a measure of the mass within a sphere, θ_{s} the centre of this sphere and θ_{P} a measure of its width. The fitness criterion is then defined such, that when remapped into the source plane, the source-positions of all multiple images within a multiple-images system overlap as much as possible, given the lens configuration of Eq. 4.49.

4.4.1 Defining χ^2-functions

We will implement the strong-lensing constraints into our grid-based, nonparametric method in a similar way as in the parametric strong-lensing methods. We define a χ^2-function for the multiple-image systems as in Eq. 4.46 with the difference, that the model parameters are not

given by an analytic model, but by the lensing potential on the grid ψ

$$\begin{aligned}\chi^2_{\mathsf{m}} &= \sum_{i=1}^{N}\left(\frac{\beta_i(\psi)-\langle\beta\rangle(\psi)}{\sigma_{i,\mathsf{m}}}\right)^2 = \sum_{i=1}^{N}\left(\frac{(\theta_i - Z(z)\alpha_i(\psi)) - \langle\beta\rangle(\psi)}{\sigma_{i,\mathsf{m}}}\right)^2 \\ &= \sum_{i=1}^{N}\frac{1}{\sigma^2_{i,\mathsf{m}}}\left(\theta_i - Z(z)\alpha_i(\psi) - \frac{1}{N}\sum_{j=1}^{N}\theta_j - Z(z)\alpha_j(\psi)\right)^2,\end{aligned} \tag{4.50}$$

where the error σ_{m} is given by the grid-cell size.
If we assume a grid cell with index i to be part of the critical line it has to minimise

$$\chi^2_{\mathsf{s}}(\psi) = \frac{|\det A(\psi)|^2_i}{\sigma^2_{i,\mathsf{s}}} = \frac{|(1-Z(z)\kappa(\psi))^2 - |Z(z)\gamma(\psi)|^2|^2_i}{\sigma^2_{i,\mathsf{s}}}, \tag{4.51}$$

where the error estimate σ_{sfs} is related to the inaccuracy in predicting the position of the critical curve. We approximate this uncertainty to first order with the help of the Einstein radius (see Cacciato et al., 2006)

$$\sigma_{\mathsf{s}} \approx \left.\frac{\partial \det A}{\partial\theta}\right|_{\theta_{\mathsf{c}}} \delta\theta \approx \frac{\delta\theta}{\theta_{\mathsf{E}}}, \tag{4.52}$$

with the angular inaccuracy due to pixelisation $\delta\theta$.
Strictly speaking, the formula above only holds for an isothermal sphere, but nevertheless it gives us a good approximation for the noise in the critical-curve position.

4.4.2 Problems with strong lensing

Pure strong lensing reconstructions have been applied to several clusters and delivered comprehensive results. We will just list some of the more recent. Reconstructions of the "`Golden Lens`" `Abell 1689` with over 100 strong lensing features were performed by Broadhurst et al. (2005) and Halkola et al. (2006). Limousin et al. (2008) reconstructed `Abell 1703`, which is not less spectacular. Zitrin et al. (2009) identified several unknown multiple-image systems in `CL0024+1654` by using a parametric strong-lensing reconstruction. A sample of 10 clusters observed by the `HST` was reconstructed by Comerford et al. (2006).
Despite the detailed recovery of strong-lensing features, like critical lines as shown in Fig. 4.10, two aspects should be noted while talking about pure strong-lensing reconstructions:

1. Parametric strong-lensing reconstructions can only be as good as the given input model. The results do not only rely on the data, but also on the assumptions and complexity of the parametric model. Nonparametric strong-lensing methods exist and have already been applied to real data (Liesenborgs et al., 2008), but their reliability have still to be tested carefully.

2. A maybe more severe problem is the fact, that strong-lensing reconstructions rely on constraints, which are confined to a very limited area of the cluster. It is therefore dangerous to extrapolate the result from a reconstruction of the cluster core to its outskirts (Meneghetti et al., 2009; Donnarumma et al., 2010). Weak-lensing constraints should be added in order to constrain cluster properties over the complete field-of-view.

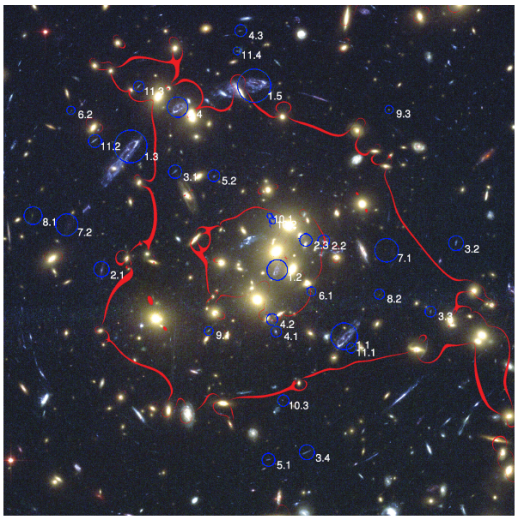


Figure 4.10: From Zitrin et al. (2009). A parametric strong-lensing reconstruction of CL0024+1654. Overlaid on top of the HST/ACS exposure are in blue the positions of the multiple images and in red the position of the reconstructed critical line, assuming a source redshift of $z_s = 1.675$. The field size is $\sim 100'' \times 100''$.

4.5 Advantages and problems of the approach

The section above finishes the summary of all constraints, which are implemented in our reconstruction method. We would like to highlight the advantages and problems with our approach of combining all these constraints into a joint reconstruction method.

Advantages

1. The main advantage is obvious. While combining different observations, one can use their individual strengths, while eliminating their weaknesses:

   - Weak lensing allows a reconstruction over the whole observational field but is based on a first-order approximation of the lens equation and therefore rather insensitive to substructure and fails in the highly non-linear regime of the cluster core.
   - Flexion is sensitive to substructure, especially in the transient regime between weak and strong lensing, but it also does not allow a reliable reconstruction of the innermost core of the cluster.
   - Strong lensing alone should not be used to determine the cluster profile too far beyond its Einstein radius, which is perfectly covered by weak lensing and flexion in our method, but it provides excellent reconstructions of the cluster core if strong lensing constraints are present. A problem of strong lensing can be its dependence on parametric models of the cluster-mass profile.

2. Since our method is fully nonparametric, also this problem is cured. The problem of under-determination from which nonparametric strong lensing models usually suffer does not appear, since the additional constraints maintain the problem fully determined.

Problems

1. The biggest problem of an approach, as we have presented it, is the question of error estimation on the derived quantities. An analytic derivation of the error budget has not been found yet, so error bars are attached by statistical methods. Input catalogues are bootstrapped (Bradač et al., 2005; Davison A.C., 1997) and the reconstruction is repeated for each (~ 100) resampling of the catalogues. The scatter in the subsequent analysis provides the error bars.
   An even more attractive method to obtain error bars, would be an MCMC sampling of the parameter space during the optimisation process. Unfortunately, this method is still too costly with respect to CPU time.

2. This leads to the second problem of our approach. Combined reconstruction methods call for rather complicated implementations, which are quite computationally demanding. The combination of different lensing regimes, asks for reconstruction on different length scales within the observational field and nonparametric reconstructions need a very complex optimisation strategy (compare Liesenborgs et al., 2007).

To explain how to tackle these problems, we spend the whole next chapter on the concrete, runtime-efficient implementation of our method.

4.6 Additional constraints

We will complete this chapter with an overview over some other observational constraints, which are not implemented in our algorithm yet, but which are likely to contribute in the future. For some of them, an implementation, which should still be nonparametric, is quite straight-forward, while for other constraints the implementation might face significant difficulties.

Magnification
Another lensing observable, which could be included in our code is magnification. Broadhurst et al. (1995) suggested the use of magnification in mass reconstructions of galaxy clusters and Bartelmann et al. (1996) even implemented it already in a maximum-likelihood approach. Mass reconstructions would profit strongly from the use of a constraint based on magnification, because this quantity is not invariant under the mass-sheet degeneracy and could therefore be used to break it (see Sec. 5.6.2). The problem are the difficulties in measuring a magnification signal since the intrinsic, unlensed flux of an object is usually unknown.

Kinematics
A constraint, which is not based on gravitational lensing are kinematics of cluster member galaxies. The velocity dispersion of those objects also reflects very directly the potential wells within the hosting galaxy cluster. Sand et al. (2002, 2008) suggested a method based on the velocity dispersion of the BCG and combined this constraint with strong-lensing constraints. Rines et al. (2010) studied a spectroscopic sample of 15 clusters and infered their virial mass from kinematic constraints.

Tracers of the Intra-Cluster-Medium (ICM)
Cluster constraints, which are not based on optical observations are mainly X-ray and Sunyaev-Zel'dovich observations. X-ray reconstructions have proven several times to deliver excellent results, especially in the case of relaxed objects (Schmidt and Allen, 2007; Donnarumma

et al., 2009; Vikhlinin et al., 2006; Ettori et al., 2002). We saw in Sec. 3.3, that the derivation of cosmological parameters from large samples of galaxy clusters by applying well-calibrated scaling relations is a particular strength of X-ray observations (Vikhlinin et al., 2009). Those scaling relations could be derived from a joint reconstruction method (Meneghetti et al., 2009) as we present it in this work.
The same is true for observations based on the SZ-effect, which have become more interesting recently, with the first SZ cluster detections (Vanderlinde et al., 2010) by the `South Pole Telescope` (Carlstrom et al., 2009).
Both methods could deliver valuable, nonoptical contributions to our method, especially since they are directly tracing the baryonic component of the ICM and could therefore give insight to the interplay between the dominant dark-matter component, which lensing is mostly sensitive to, and the baryons in the ICM (Powell et al., 2009).

5 Implementation

This rather technical chapter is dedicated to the concrete implementation of our reconstruction method and therefore represents the essence of this work. The algorithm encompasses several different input constraints, confined to different length scales. This calls for several numerical techniques which are not straight-forward to implement. We will describe them in quite some detail.
After an overview of modern high-performance computing in Sec. 5.1, we will define a static version of adaptive-mesh-refinement (AMR) grids in Sec. 5.2. The connection between observable quantities and their appearance on those AMR grids is made in Sec. 5.3. How to minimise a multi-component χ^2-function, defined on AMR grids, in an extremely fast way is shown in Sec. 5.4, by using the iterative approach explained in Sec. 5.5. An overview of the whole reconstruction method, the description of the numerical package and the usage of the output is presented in Sec. 5.6.

5.1 Modern high-performance computing

Computers are definitely one of the main tools in modern science. Most of the more recent results and insights would not have been possible without their tremendous speed in numerical calculations and ability to model complex, physical processes. The ability to store vast amounts of experimental data has just become available in modern times and improved the understanding of natural laws, by exploiting the increased statistics of the observed samples.
Also Cosmology is footing heavily on modern high-performance computing for two reasons. First, the observational point of view. Modern surveys will have a huge, almost full, sky coverage and the data that the telescopes take must be selected, stored and analysed. A survey which has just started to deliver first data products is the `Panoramic Survey Telescope & Rapid Response System` (`Pan-Starrs`) (Kaiser et al., 2002). The `Pan-Starrs` survey delivers several TB's of data every night. An even more extreme case is the planned `Large Synoptic Survey Telescope` (`LSST`) facility (LSST Science Collaborations, 2009). Its enormous GigaPixel camera will deliver even more data per night and dedicated computation centres will be assigned just for the storage, analysis and distribution of this data.
Second, the end of the last century saw the advent of large cosmological simulations as reliable scientific tools. These attempts to model the whole[1] Universe, or parts of it, on different spatial and temporal resolutions would not be possible without the computational resources that we will describe. The most prominent examples in cosmology were the `Millennium` (Springel et al., 2005) and `MareNostrum` (Gottlöber et al., 2006) simulations with huge cosmological volumes, or more recently the `Aquarius` project (Springel et al., 2008), which tried to model the creation and evolution of a structure compatible in size with the Milky Way.
Of course scientists are interested in an improvement of the numerical performance on modern computer systems. Basically, this improvement can be achieved in the following ways[2]:

[1] Meaning those parts, which are of interest for a certain scientific field
[2] In principal one should add, that most scientific codes are written in an extremely inefficient way with respect to memory access and clock rate adaptation and do not use the available resources.

1. Increase individual CPU performance
2. The use of specialised hardware
3. Massive parallelisation.

As the first point sounds rather trivial, it is indeed the main driver for the rise of modern super computing. The numerical performance of CPU's increased exponentially since the first production series in the 1960's. This performance is mainly given by the number of transistors in the CPU, used to execute instructions, and by its clock rate. Unfortunately, it seems that both aspects of CPU performance are approaching natural limits dictated by solid state physics due to limitations regarding further miniaturisation and cooling problems. Ways out of this dead end are multi-core designs and completely different architectures like quantum computers in the future.
Another solution to increase the numerical performance in scientific applications is the use of specialised hardware, which is designed for a specific purpose and not for general purposes like standard CPU's. An example are custom-built hardware boards called `Gravity Pipe` (`GRAPE`) (Makino and Taiji, 1998). These devices are specifically designed to evaluate the force equations in astrophysical N-Body simulations in an extremely efficient way. Another, more general, example are `Field Programmable Gate Arrays` (`FPGA's`). Those devices allow for a direct on-chip implementation of numerical algorithms after the chip was actually manufactured. Since the calculations are basically hard-wired an extremely high performance can be reached. `FPGA's` are used in several scientific areas. The most prominent example are the detectors of the `Large Hadron Collider` where they are used for triggering and event selection. Also in astrophysical applications they contribute again mostly to accelerate N-body simulations (Spurzem et al., 2008). Despite their high numerical performance, there is a severe drawback in specialised hardware. The fact that they are not produced for the mass market makes those devices expensive in acquirement, and maintenance. Furthermore, special knowledge for programming and assembling has to be available.
The standard way to increase the numerical performance on ordinary hardware is massive parallelisation. In this case a numerical task is split into separate problems, which are then executed simultaneously. We will focus on parallelisation from now on, since this is also the way in which our reconstruction method is implemented. We will have to introduce some vocabulary first[3].
A numerical problem we will call a task. If this task can be subdivided into a certain amount of sub-problems we will call them threads. If a CPU has several independent calculation units which can execute threads, we will call them compute cores or just cores. If a certain area of the main memory is accessible only to one thread, we will call it distributed memory, if it is accessible by several threads we will call it shared memory. A single computer with one or several cores and a certain amount of main memory we will call a node. If we talk about process communication, we mean exchange of data between the dedicated memory spaces of different threads.

5.1.1 Single node computing

Already on single-node systems parallelisation is possible since modern CPU's are assembled in a multi-core architecture, usually including two, or more recently even four or six, compute cores. Additionally, powerful server or workstation nodes are mostly equipped with

[3] The used terms may very well differ from common definitions or from the expressions used by other authors

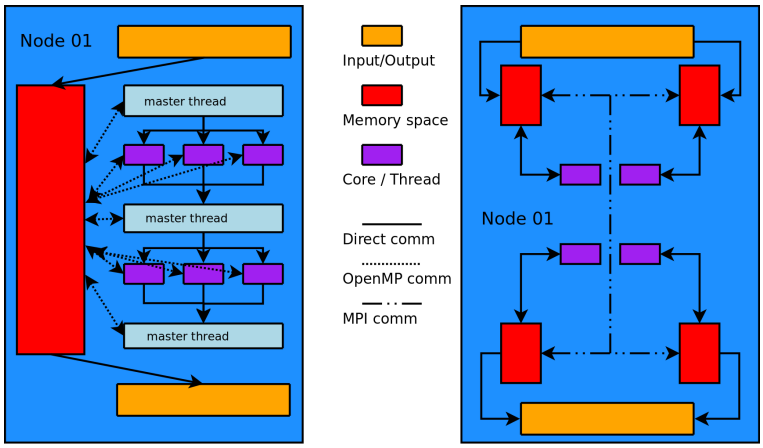


Figure 5.1: Parallelisation on a single node. *Left Panel:* The workflow using e.g. `OpenMP`. The programme runs on a single memory space, which can be accessed by all compute nodes. A so-called master thread performs the tasks, which cannot be parallelised like e.g. I/O operations and calls the different compute cores (in this example three) whenever a task profits from their combined computational power. *Right Panel:* The typical workflow when using `MPI`. All compute cores (in this example four) access their own, separate memory space. Message passing (dash-dotted line) is used for communication. I/O operations are performed by only one or all of the available threads.

more than one CPUs. This gives the user the possibility to split his problem into several threads and the number of threads usually coincides with the number of compute cores because overloading the node does generally decrease the numerical performance. A specific advantage of single node parallelisation is the usage of shared memory. Obviously, all threads can work on the same memory space and process communication usually just reduces to synchronisation instructions. The most commonly-used standard to implement this parallelisation model is called `OpenMP`[4] and has already been included in the most important compilers for programming languages used in science like `C(++)` and `Fortran`. A graphical overview of the `OpenMP` parallelisation model is show in Fig. 5.1.
There might be certain situations where it is eligible to establish a shared-memory parallelisation set-up on a single node. This can be the case when more flexibility within the process communication is desired. The widely-accepted standard for this programming model is the `Message Passing Interface (MPI)` [5]. The basic workflow of `MPI` can also be seen in Fig. 5.1. All threads maintain their own, separate part of the main memory and communication is performed via the internal bus of the system.
The main disadvantage of single-node parallelisation is simply given by the limited number of reasonable threads, since the number of available compute cores is also limited.

5.1.2 Cluster computing

The fundamental limitation of single node computation can be overcome by combining a number of nodes into a large compute cluster. Of course, communication between the different nodes must be possible, which is why they are usually connected via a high-performance,

[4] http://openmp.org/wp/
[5] http://www.mcs.anl.gov/research/projects/mpi/

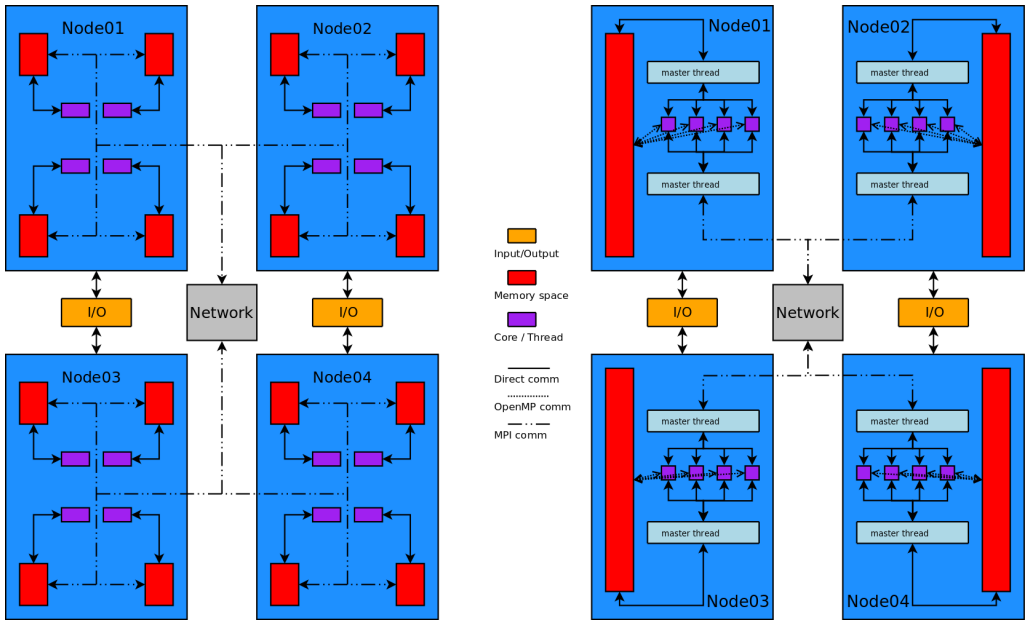


Figure 5.2: Parallelisation on a compute cluster. *Left Panel:* The basic workflow of an implementation which is just based on `MPI`. Several nodes (in this example four) are connected via an external network and communicate via message passing. Within the nodes, several cores (in this example four) work on distributed memory spaces and communicate, also by message passing, via the system bus. *Right panel:* As before, but the different cores work on shared memory spaces. One master thread per node communicates with the other nodes within the cluster.

low-latency network. Most big clusters are using the `InfiniBand`-specification[6] but solutions using e.g. the standard `GigaBit Ethernet` specification are also in use. It is clear that in this way it is possible to construct huge computers, which can execute thousands of threads in parallel. In fact, the fastest computers in the world are Linux clusters of the form that we have just described and consist of tens of thousands of nodes, which combine hundreds of thousands of compute cores[7].

Naturally, such a machine cannot be embedded in a single shared-memory environment any more. The memory spaces of different nodes are physically separated and e.g. `MPI` has to be used to allow for process communication between the threads on different nodes. Within individual nodes one can choose to use either a distributed-memory model to access the memory of the node or a shared-memory model. Both types of set-ups are illustrated in Fig. 5.2.

The biggest difficulty in cluster computing is the complexity of the communication between the possibly huge number of threads. Usually, threads are organised in a number of sub-groups and process communication is performed at first just within the sub-group and at a certain point of the task, sub-groups start communicating with each other. Generally, the relatively slow communication between nodes, using the external network, has to be minimised to achieve the best result possible.

[6] http://www.infinibandta.org/
[7] http://top500.org

5.1.3 GPU computing

The last possible form of parallelisation, that we intend to mention, is somewhat related to the use of specialised hardware as described in Sec. 5.1. Driven by the multi-billion Dollar computer-gaming industry, more and more powerful graphics cards were produced. Those are based on so called `Graphics Processing Units` (GPU's), which are a special form of processors specialised on parallel calculations. While on normal CPU's a large number of the available transistors is dedicated to cache memory, so temporal saving of data and to data-flow control, only a limited number of transistors is actually dedicated to calculation work. This picture completely changes on GPU's. Most of the available resources on the chip are dedicated to calculations and therefore hundreds of compute cores find space on one single GPU chip. This opens the door for massive on-chip parallelisation.
This has also been recognised in the scientific community, which is the reason why GPU implementations gain more and more momentum also in astrophysics (Barsdell et al., 2010; Hassan et al., 2010; Hu et al., 2010; Sainio, 2009; Wang et al., 2009). Of course, there is a drawback in GPU computing due to the lack of auxiliary transistors on the chip. A specific task has to be specially suited for parallelisation to take full advantage of the tremendous computational resources already available with a single GPU. More specifically, a problem must be data-parallel (NVIDIA, 2009b), which means that all its threads have exactly the same form with the only difference that they address different positions in memory. Also, the GPU uses its own on-board memory. As a consequence, data transfers have to be made between the main memory of the node to the memory of the GPU.
The programming model for parallelsing a task on a GPU distinguishes between a so-called host code, which runs normally on the node and a so-called device code, which runs massively-parallel on the GPU. Because of their different architecture, GPU's need to be addressed by using a special programming language, which in the case of `NVIDIA`[8] graphics boards, is either `C for CUDA` or the `CUDA Driver API` (Kirk and Hwu, 2010; NVIDIA, 2009b,a,c). `C for CUDA` is a high-level programming language and is easy to learn for an experienced programmer. Still, it offers a broad amount of flexibility and enables the usage of several different memory levels on the Graphics board. This ranges from the global GPU memory down to on-chip shared memory, which allows small groups of threads memory access on the register level.
But GPU parallelisation does not end with hundreds of cores on one GPU, several GPU's can be used in the same node and several nodes can be combined to form GPU clusters with thousands of compute cores at extremely low cost. Both models are shown in Fig. 5.3. Also our reconstruction method is implemented on GPU's and we will mention how their resources are used in the course of this chapter.

5.2 Adaptively refined grids

From the computers on which we intend to implement our algorithm, we will now turn the focus to the basic elements of this implementation. As mentioned in Sec. 4.1.1, we perform the reconstruction on a grid and we will call the grid cells pixels. Since we have to combine constraints, which are confined on quite different length scales and therefore on different resolutions in terms of grids, we will not use a regular grid with fixed pixel size. Instead, we will make use of a technique well-known as adaptive-mesh-refinement (AMR).
The main feature of AMR grids is their varying pixel size within the full grid area. This enables

[8] http://www.nvidia.com

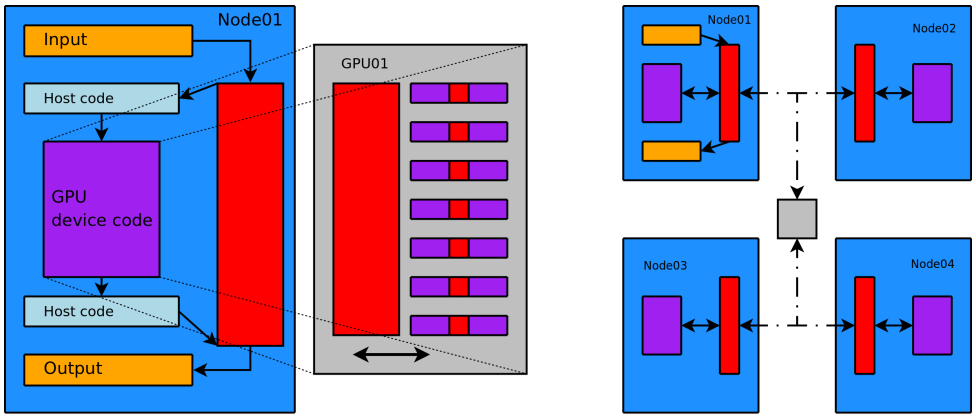


Figure 5.3: GPU computing. *Left panel:* The workflow of a GPU implementation on a single node. The programme is started normally in the form of a so-called host code using the nodes I/O and memory capabilities. At a certain point, data is transfered from the host memory to the device memory of the GPU. On the GPU the device code uses the resources of hundreds of available cores for data-parallel code execution. Also, extremely fast on-chip shared memory can be used. This is indicated by the red areas between the GPU compute cores. After the execution of the device code, the result is transfered back to the host, where it is processed further. *Right panel:* Several GPU-powered nodes can be combined into a GPU cluster via message passing.

us to increase the resolution in the strong-lensing regime and to follow the strong-lensing features with high accuracy. Since we are not limited to a single refinement level, we will also introduce intermediate resolutions to describe the transition between the weak -and the strong-lensing regime, which is particularly interesting with respect to gravitational flexion. Another advantage of AMR grids arises from the fact, that the total number of pixels can be reduced, because the resolution is only high in those areas where this is really necessary. Runtime improvements are a direct consequence. We will define and explain our implementation of AMR grids, which simplifies with respect to dynamical AMR implementations (see e.g. Fryxell et al., 2000) since lensing reconstructions are a static problem. We will explain our implementation on one projection of a simulated, merging galaxy clusters, which is shown in Fig. 5.4.

5.2.1 Refinement criterion

We decided that for our purposes it is sufficient to implement three different refinement levels. We call them refinement level 0, 1 and 2, where level 0 refers to the lowest pixel resolution and large pixel size, level 1 describes an intermediate resolution and level 2 stands for the highest pixel resolution and a small pixel size. The resolution doubles between each refinement level, which is visualised in Fig. 5.5. A crucial ingredient of an AMR grid is a refinement criterion, so a rule which defines what resolution is assigned to a certain area of the reconstruction grid. Since we want to use the refinement for increasing the resolution of areas in the reconstruction where strong lensing takes place, the quantity which traces these areas best is the determinant of the lensing Jacobian as defined in Eq. 2.65. We calculate the mean and the signal-to-noise values of this quantity on the whole reconstruction grid and assign a pixel, which exceeds a first S/N-threshold, to refinement level 1 and a pixel, which exceeds a second, higher SN-threshold, to refinement level 2. The outcome of this proce-

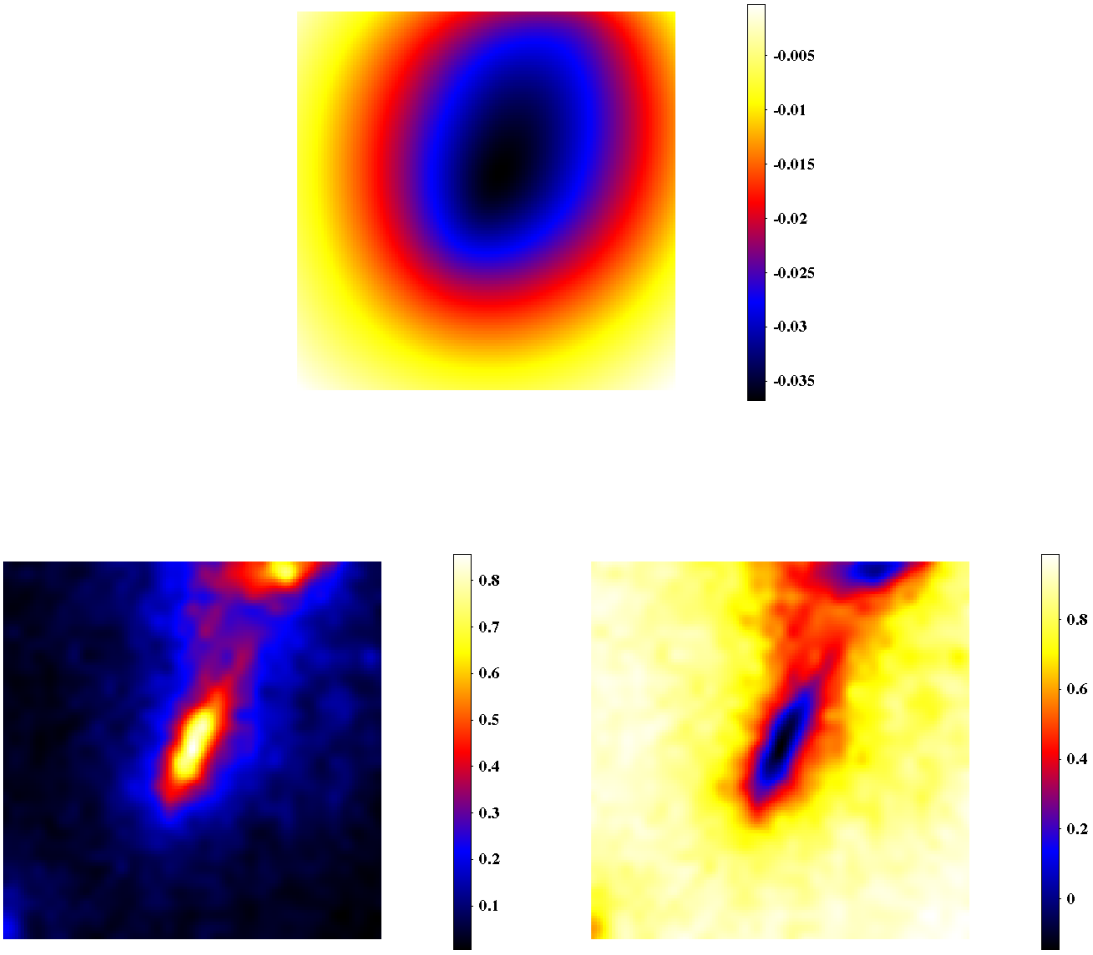


Figure 5.4: One projection of a simulated, merging galaxy cluster. *Upper panel:* The lensing potential on a regular grid with a resolution of 128×128 pixels. *Lower left panel:* The associated convergence map of the cluster on the same pixel resolution. *Lower right panel:* The associated determinant of the lensing Jacobian.

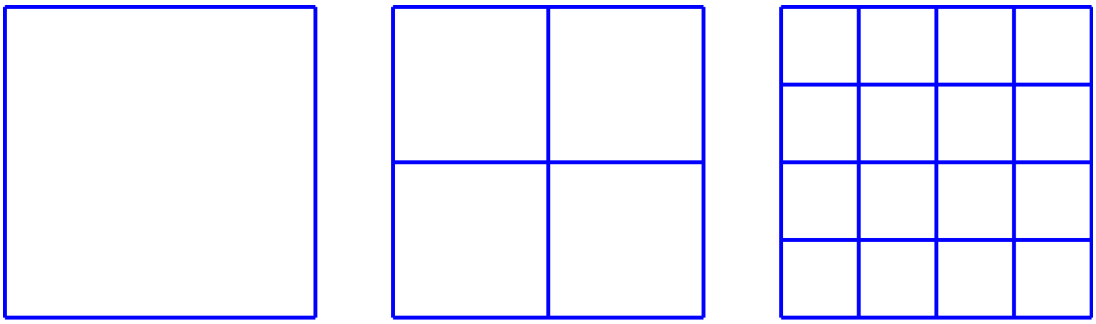


Figure 5.5: The three different refinement levels of our AMR implementation. While a given area on the reconstruction grid refers to one pixel on refinement level 0, it refers to four or 16 pixels on refinement level 1 or 2, respectively.

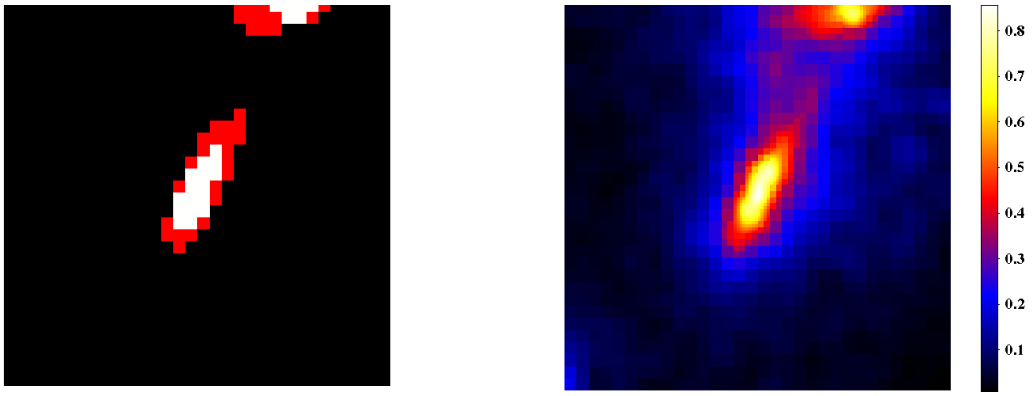


Figure 5.6: *Left panel:* The refinement map for the cluster shown in Fig. 5.4, based on the signal-to-noise level of the lensing Jacobian. Black colour refers to a S/N-level $<= 2$, red colour to $2 <$ S/N-level < 3 and white colour to a S/N-level $=> 3$. *Right panel:* The respective convergence map. One can clearly identify the three different resolutions in the map, especially when comparing to the refinement map on the left.

dure is shown in Fig. 5.6, together with the resulting adaptively refined convergence map. Alternative refinement criterions might be the convergence map of the reconstruction, or an a-priori user definition of the refinement levels in the different areas of the reconstructed field. The former following the density distribution of the reconstruction but not necessarily the strong-lensing features, the latter being somewhat arbitrary and not flexibly changed during the reconstruction.
One may raise the question from what to derive the signal-to-noise map of the reconstruction in the first place and we will answer it with an iterative approach that we address in Sec. 5.5.

5.2.2 Pixel indexing

While having defined three different refinement levels and a suitable refinement criterion, we will now describe the concrete entity of an AMR grid within our reconstruction algorithm. Every quantity which will be situated on the adaptive grid, will be stored in form of a vector, having a length identical to the number of pixels in the grid. Thereby is plays no role what kind of refinement level a certain pixel has, so we have to save this information elsewhere.

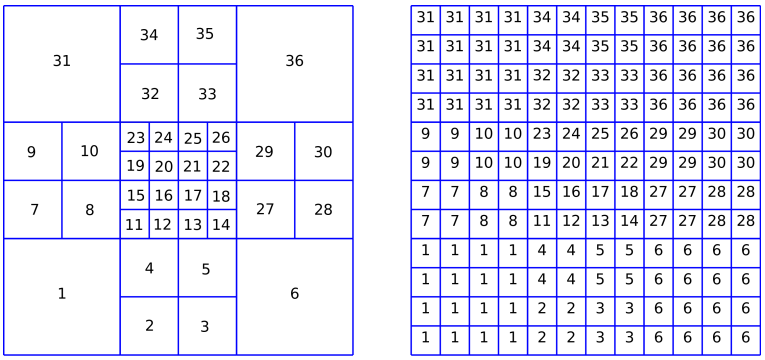


Figure 5.7: *Left panel:* The scheme for ordering the pixels within the AMR grid vector. Counting starts in the bottom-left corner, continues in x-direction and after reaching the right border, continues one line above. The pixels of refined grid areas are counted with the same scheme sequentially. *Right panel:* The index grid on full (refinement level 2) resolution to find next neighbours in a simple way.

Following this idea, the AMR grid should be defined by the number of pixels that it contains and by additional information, unique for each pixel, that we will describe later.
Starting with a refinement map, as it is shown in Fig. 5.6 the number of pixels n_{pix} in the grid is simply calculated by

$$n_{\text{pix}} = n_o + 4n_1 + 16n_2, \tag{5.1}$$

where n_o is the number of pixels with refinement level 0, n_1 is the number of pixels with refinement level 1 and n_2 is the number of pixels with refinement level 2. Having obtained the number of pixels in the grid one can now define a property vector of length n_{pix}, which contains for each pixel the following information:

1. The x-coordinate of the pixel centre with respect to the total grid size,
2. The y-coordinate of the pixel centre with respect to the total grid size,
3. The refinement level of the pixel,
4. The positions of its 48 neighbours within a box of 7×7 pixels around the pixel centre.

The last information is necessary to perform finite differencing on the AMR grid , that we will describe in the next section. The process of finding the neighbours of a pixel on the grid is by no means trivial and is performed in the following way: All grid pixels are ordered as shown in the left panel of Fig. 5.7. Additionally, an index grid is produced with a pixel size refering to refinement level 2 (see Fig. 5.7). The 48 neighbours of an individual pixel with a certain refinement level are found by stepping from the centre of this pixel in the index grid by one, two or three steps up, down, left or right. The step size reflects the refinement level of the pixel.
It should be noted, that our AMR implementation also allows for masking the grid. The procedure described above is not affected by this, besides that certain areas in the index grid are marked to be not part of the grid.
If a certain neighbour of a pixel does not exist for a pixel, because it would reside outside the grid or within a masked area, this is remembered by the property vector.

5.2.3 Noise-reduced finite differencing

A particularly difficult task on AMR grids is finite differencing as we need it for our method. The problem of this technique, which can be implemented in a very straight-forward way on regular grids, are the different pixel sizes within the grid. To perform a finite-differencing operation for a particular pixel, one needs, depending on the differential operator, several neighbour positions and the distance to these neighbours. Also on regular grids, a problem appears when a pixel lies near the edges of the grid or near a masked area. Several neighbours might not be available. Therefore, we defined several different forward, backward, central and mixed finite-differencing schemes (Figs. 4.2-4.6). Depending on the neighbours, which are available on the grid, the best finite-differencing scheme is applied. The position, or absence of neighbours is known as described in the previous section.

Priority and rejection

We apply the same idea for finite-differencing on AMR grids , but adding one additional check and one additional priority. Depending on the refinement of the pixel we are interested in, the distances to the neighbours are different. Especially if a highly-refined pixel sits next to a low-resolution area, it might very well be the case that different neighbour positions in the index grid (Fig. 5.7) belong to the same pixel. If this is the case, while applying a certain finite-difference scheme, this scheme is rejected by the AMR grid . Furthermore, to ensure a smooth derivative, we demand all used neighbours for a finite-differencing scheme to be of the same or higher refinement level as the pixel of interest. If not all of these requirements can be fulfilled, the distance to the next neighbours is increased. This priorities for the choice of the finite-differencing scheme are visualised in Fig. 5.8. After two of these distance increments the demand for the same or higher refinement level in all neighbours is dropped. This ensures that the finite differencing procedure can be performed at all for peculiar cases, which may appear at the border between areas with a high and a low refinement level.

To demonstrate the performance of this priority-and-rejection procedure we also convert the lensing potential of our example cluster to a representation on the AMR grid . Afterwards, we perform finite differencing to obtain the convergence map of the cluster. In Fig. 5.9 we show the result with and without the priority-and-rejection scheme.

5.3 Constraints translation

As described in Sec. 5.2.2 all quantities, which are used in a reconstruction, have to be defined on the AMR grid . Of course, this applies also to the observational quantities serving as the main input of a reconstruction. The way in which they are translated onto the AMR grid depends on the type of the constraint and one should also note that not all of these different types must be used during the reconstruction. The input formats are defined as follows:

- **Ellipticity measurements of background galaxies**
  must be given in the form of an `ASCII`-catalogue with the following columns:

  `x-position` `y-position` ε_1 ε_2 `weighting-factor`,

  where one row represents one individual galaxy. The units for the x -and y-positions are arbitrary, as long as they coincide with the coordinate frame of the other input quantities. Galaxy positions are expected to lie on a plane patch in the sky, so R.A. and DEC positions need to be converted, accordingly.

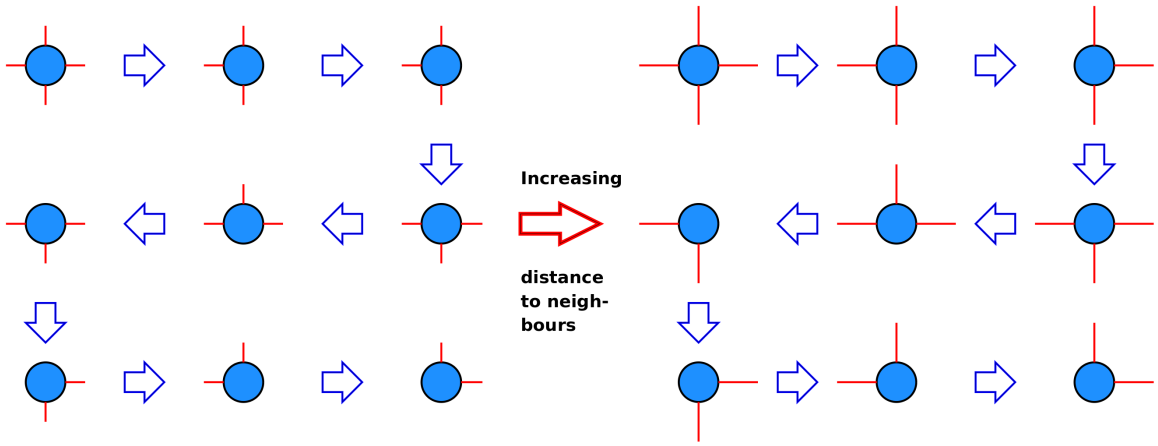


Figure 5.8: The sequence of finite-differencing schemes, ordered by their priority within the priority-and-rejection scheme. The blue spheres mark the pixel of interest and the red bars show the forward, backward or central scheme for which neighbours in the grid are searched. If no suitable neighbours are found, according to the rules described in the text, the distance within which neighbours are searched is increased.

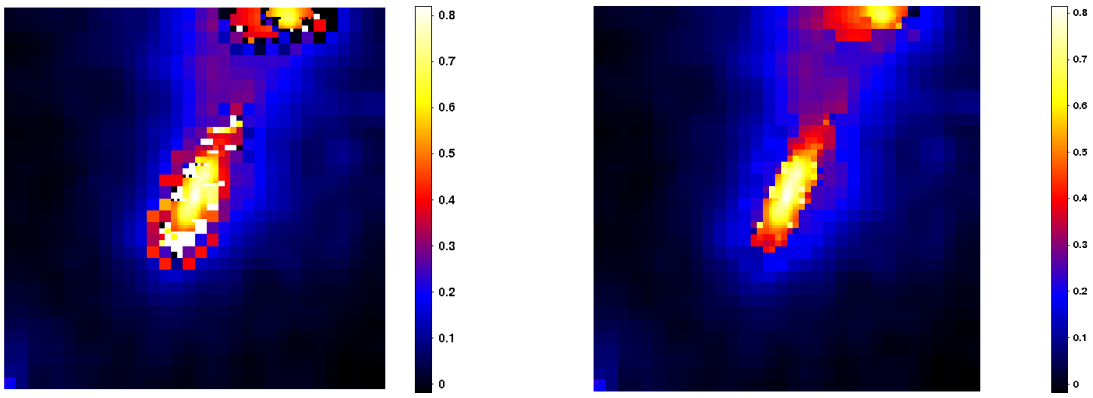


Figure 5.9: The recovery of the convergence from the lensing potential on an AMR grid . *Left Panel:* The result without our rejection technique, which delivers a quite noisy result. *Right Panel:* The result with the rejection technique described in the text. The noise has almost disappeared completely.

- **Flexion measurements of background galaxies**
  also given in the form of an ASCII-catalogue consisting of the following columns:

  x-position y-position f_1 f_2 g_1 g_2 weighting-factor,

  where the same rules on the positions apply.

- **Multiple-image-system positions**
  which are a little bit more complicated to describe, are also given in ASCII-format consisting of separated blocks, each of which contains information on one individual system:

#1	System identifier
N_{sys} z	number of images and redshift of the system
x-position y-position	of image one
x-position y-position	of image two
x-position y-position	of image three
x-position y-position	of image four
x-position y-position	of image five,

  and the next block following.

- **Critical-curve-position estimators**
  are simply provided as an ASCII-catalogue with three columns per estimator and one estimator per row:

  x-position y-position redshift.

5.3.1 Adaptive shape averaging

Given the input catalogues of the previous section, these quantities need to be translated into grid vectors. As we have mentioned in Sec. 2.2.2, weak lensing has to be treated statistically, which is why we will apply an adaptive-averaging technique to the ellipticity and flexion fields. This technique is used identically for the two ellipticity and four flexion components .
After the reconstruction field is defined by the coordinates of the four field corners in a coordinate frame which is consistent with the input catalogues, the centres of each pixel in the AMR grid are calculated. The adaptive averaging starts by defining a circle around each pixel centre, whose initial radius refers to the pixel size. The number of background galaxies within this circle is counted and if it does not match or exceed a certain number, which is given by the user or by the variance in the circle sample, the radius is increased by e.g. 10 %. This procedure, visualised in Fig. 5.10, is repeated until a sufficient number of background galaxies are contained within the circle around one pixel. Once this is achieved, the following information is calculated or saved by the adaptive-averaging routine:

1. Mean of the component,
2. Standard deviation of the component,
3. Galaxies associated to this particular pixel.

Means and standard deviations are saved as grid vectors and the galaxy associations are saved for further processing (Sec. 5.3.2).

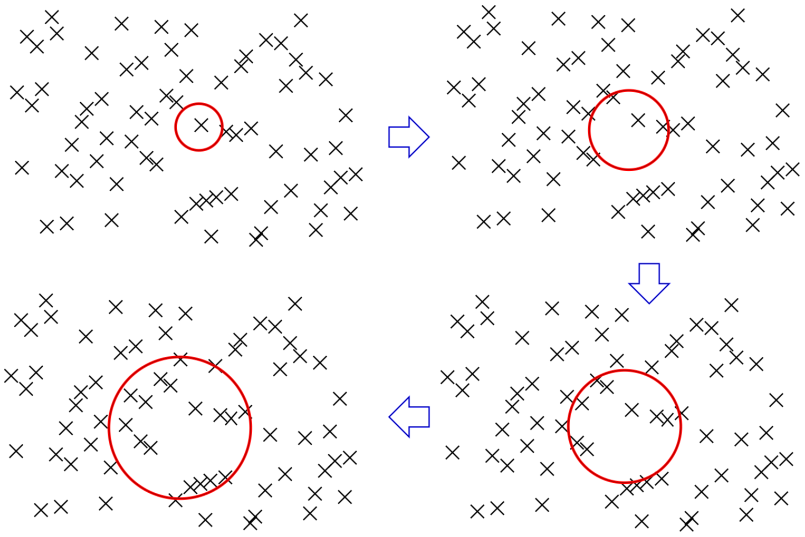


Figure 5.10: The concept of the adaptive-averaging technique.

It should be mentioned at this point, that the adaptive-averaging process can be done independently for each individual pixel and is therefore ideal for parallelisation. We implemented this step, even if it is not that computationally demanding, on GPU's by processing individual pixels on separated cores.

5.3.2 Covariances

Since the circles, within which we calculate the averages for an individual pixel, might overlap (as indicated in Fig. 5.11), we have to take into account the correlations which arise between the AMR grid pixels. Let, in the course of the argument, $\mathcal{S}_i$ be the set of galaxies which were used to calculate the mean and standard deviation in the grid pixel i.
Starting from the considerations in Sec. 4.2.1 we calculate each entry of the covariance matrix $\mathcal{C}$ for the whole reconstruction grid and for all averaged quantities by

$$\mathcal{C}_{ij} = w_{ij}\sigma_i\sigma_j, \tag{5.2}$$

where $\sigma_{i,j}$ are the standard deviations of the i-th and j-th pixel. The quantity w_{ij} is a weighting factor, which describes the overlap between two pixels. It is defined as

$$w_{ij} = \frac{\#\mathcal{S}_{ij}}{\frac{1}{2}\left(\#\mathcal{S}_i + \#\mathcal{S}_j\right)}, \tag{5.3}$$

with

$$\mathcal{S}_{ij} = \mathcal{S}_i \cap \mathcal{S}_j. \tag{5.4}$$

In this way, the weighting factor w retains the following, required properties

$$w_{ij} = 1 \quad \text{for} \quad i = j \tag{5.5}$$

$$0 < w_{ij} < 1 \quad \text{for} \quad \mathcal{S}_{ij} \neq \emptyset \tag{5.6}$$

$$w_{ij} = 0 \quad \text{for} \quad \mathcal{S}_{ij} = \emptyset. \tag{5.7}$$

This step is computationally rather demanding, especially if a large catalogue of background galaxies is used. To be more precise, the step of calculating the weighting factor in

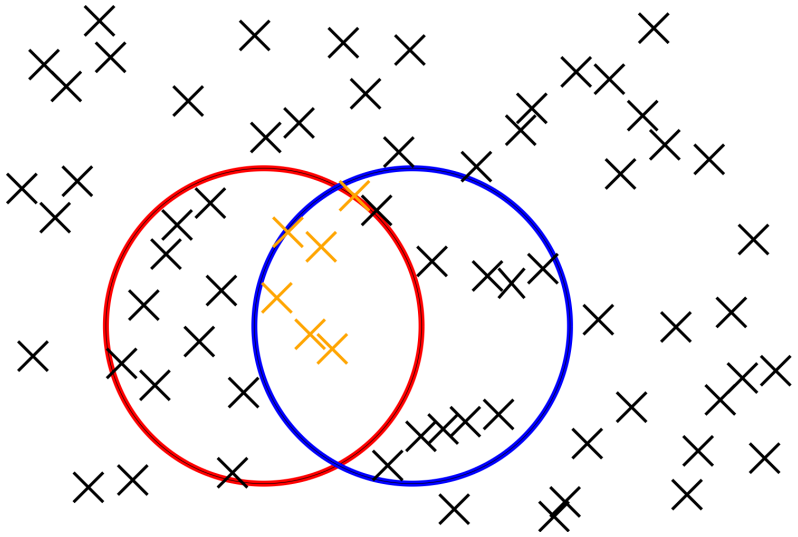


Figure 5.11: Covariances are determined by the number of galaxies which are shared between neighbouring pixels. For the example of the two pixels in the figure, those galaxies are marked in orange.

the covariance matrices takes the second most CPU-time in the whole reconstruction process, because correlations for all pixel pairs in the grid have to be calculated. There are two possibilities to speed this operation up. First, the catalogue of background galaxies can be spatially ordered and overlap searches can be constricted to limited patches. The second idea is to exploit the brute computational power of GPU's and calculate each entry w_{ij} on a separate computation core. The second idea is implemented in our routine and the performance of the `CUDA` implementation is evaluated in Sec. 6.1. If the number of used background galaxies should increase further in the future, we will also implement a suitable ordering scheme.

5.3.3 Strong lensing

The translation of strong-lensing constraints onto the AMR grid is much easier since no additional operations have to be performed on them. Their positions within the observational field need just to be associated with pixel positions on the AMR grid . Again, the coordinates of each pixel centre are compared to the coordinates of a multiple image or of a critical curve estimate. If they match, this pixel is marked as being a carrier of strong lensing information and will be considered as such during the reconstruction. Also, additional information like the redshift of the constraint or its association to a certain multiple image system are saved separately.

5.4 χ^2-minimisation

Having collected all ingredients to actually perform a nonparametric reconstruction based on multiple input constraints, it is time to summarise what we have achieved so far:
We defined all necessary χ^2-functions and connected them to the lensing potential. The core of the whole reconstruction, a flexible AMR grid , was described in Sec. 5.2 and we translated the input catalogues onto that grid in Sec. 5.3. All these parts will be used to convert the

χ^2-minimisation into a linear system of equations (LSE) by building up a coefficient matrix and a result vector, to invert the problem and to recover the lensing potential as its solution.

5.4.1 Linearisation in grid space

The overall χ^2 (Eq. 4.1) shall be the starting point

$$\chi^2(\psi) = \chi^2_{\mathsf{w}}(\psi) + \chi^2_{\mathsf{s}}(\psi). \tag{5.8}$$

The weak-lensing term splits into a shear and a flexion contribution

$$\chi^2_{\mathsf{w}}(\psi) = \chi^2_{\gamma}(\psi) + \chi^2_{F}(\psi), \tag{5.9}$$

which consist of two, respectively four components

$$\chi^2_{\gamma}(\psi) = \chi^2_{\gamma 1}(\psi) + \chi^2_{\gamma 2}(\psi) \tag{5.10}$$
$$\chi^2_{F}(\psi) = \chi^2_{F1}(\psi) + \chi^2_{F2}(\psi) + \chi^2_{G1}(\psi) + \chi^2_{G2}(\psi) \tag{5.11}$$

The strong-lensing term splits into the contributions from the multiple-image systems and from the critical-curve estimators

$$\chi^2_{\mathsf{s}}(\psi) = \chi^2_{\mathsf{m}}(\psi) + \chi^2_{\mathsf{c}}(\psi) \tag{5.12}$$

We end up with a sum of several terms, which we will minimise for each term χ_x separately, where x substitutes all the different terms above

$$\frac{\partial \chi^2_x(\psi)}{\partial \psi} \stackrel{!}{=} 0 \tag{5.13}$$

Now comes the simple, but very important step, which has already been introduced in Sec. 4.1.2 but we invest the effort to point it out again due to its central role in the whole reconstruction concept.
We make use of the fact that all our quantities are defined on grid positions, including the lensing potential. The parameter with respect to which we want to minimise the χ^2-function is the lensing potential itself at every grid position. The solution of this minimisation is also our final result, since it is that potential, defined on an AMR grid , which is most likely to have caused the multiple input constraints. Translated into a formula this reads like

$$\frac{\partial \chi^2_x(\psi)}{\partial \psi_l} \stackrel{!}{=} 0 \quad \text{with} \quad l \in [0, ..., N_{\mathsf{pix}}]. \tag{5.14}$$

We will show now for all terms, that this minimisation can be linearised and performed by solving a LSE. Please note that from now on Einstein's sum convention applies.

Shear term

We start with $x = \gamma$. We only show one of the two components and substitute the quantities $\varepsilon^1, \gamma^1, \mathcal{C}^1$ and G^1 or $\varepsilon^2, \gamma^2, \mathcal{C}^2$ and G^2, which were defined in Sec. 4.2.1, with just $\varepsilon, \gamma, \mathcal{C}$ and G. By separating the non-linear $(1 - Z\kappa)$ factors and putting them in the non-linearity matrix $\mathcal{P}_{ij}$

we rewrite the χ^2-function

$$\begin{aligned}\chi^2_\gamma &= \left(\varepsilon_i - \frac{Z_i\gamma_i}{1-Z_i\kappa_i}\right)\mathcal{C}^{-1}_{ij}\left(\varepsilon_j - \frac{Z_j\gamma_j}{1-Z_j\kappa_j}\right)\\ &= \underbrace{\frac{\mathcal{C}^{-1}_{ij}}{(1-Z_i\kappa_i)(1-Z_j\kappa_j)}}_{\mathcal{P}_{ij}}\left(\varepsilon_i(1-Z_i\kappa_i)-Z_i\gamma_i\right)\left(\varepsilon_j(1-Z_j\kappa_j)-Z_j\gamma_j\right)\\ &= \mathcal{P}_{ij}\left[(\varepsilon_i-\varepsilon_iZ_i\kappa_i-Z_i\gamma_i)(\varepsilon_j-\varepsilon_jZ_j\kappa_j-Z_j\gamma_j)\right]\\ &= \mathcal{P}_{ij}\left[\varepsilon_i\varepsilon_j-\varepsilon_i\varepsilon_jZ_j\kappa_j-\varepsilon_iZ_j\gamma_j-\varepsilon_i\varepsilon_jZ_i\kappa_i+\varepsilon_i\varepsilon_jZ_iZ_j\kappa_i\kappa_j\right.\\ &\qquad \left.+\varepsilon_iZ_iZ_j\kappa_i\gamma_j-\varepsilon_jZ_i\gamma_i+\varepsilon_jZ_iZ_j\kappa_j\gamma_i+Z_iZ_j\gamma_i\gamma_j\right],\end{aligned} \tag{5.15}$$

These non-linearities are just kept constant during a reconstruction, which on the one hand asks for an initial value and on the other hand compromises the accuracy of our calculation. This problem can be solved in an elegant way and we will deal with it in Sec. 5.5.3.
By applying the partial derivative of Eq. 5.13 to Eq. 5.15 we find

$$\begin{aligned}\frac{\partial\chi^2_\gamma(\psi)}{\partial\psi_l} = \mathcal{P}_{ij}\Bigg[&-\varepsilon_i\varepsilon_jZ_j\frac{\partial}{\partial\psi_l}\kappa_j(\psi)-\varepsilon_iZ_j\frac{\partial}{\partial\psi_l}\gamma_j(\psi)-\varepsilon_i\varepsilon_jZ_i\frac{\partial}{\partial\psi_l}\kappa_i(\psi)\\ &+\varepsilon_i\varepsilon_jZ_iZ_j\kappa_i\frac{\partial}{\partial\psi_l}\kappa_j(\psi)+\varepsilon_i\varepsilon_jZ_iZ_j\kappa_j\frac{\partial}{\partial\psi_l}\kappa_i(\psi)\\ &+\varepsilon_iZ_iZ_j\kappa_i\frac{\partial}{\partial\psi_l}\gamma_j(\psi)+\varepsilon_iZ_iZ_j\gamma_j\frac{\partial}{\partial\psi_l}\kappa_i(\psi)\\ &-\varepsilon_jZ_iZ_j\frac{\partial}{\partial\psi_l}\gamma_i(\psi)+\varepsilon_jZ_iZ_j\kappa_j\frac{\partial}{\partial\psi_l}\gamma_i(\psi)\\ &+\varepsilon_jZ_iZ_j\gamma_i\frac{\partial}{\partial\psi_l}\kappa_j(\psi)+Z_iZ_j\gamma_i\frac{\partial}{\partial\psi_l}\gamma_j(\psi)+Z_iZ_j\gamma_j\frac{\partial}{\partial\psi_l}\gamma_i(\psi)\Bigg].\end{aligned} \tag{5.16}$$

From inserting the representation of the shear and the convergence by the matrix representation of finite differences applied to the lensing potential $\gamma_i = G_{ik}\psi_k$, $\kappa_i = \mathcal{K}_{ik}\psi_k$ and the simple relations $\frac{\partial}{\partial\psi_l}\mathcal{K}_{ik}\psi_k = \mathcal{K}_{ik}\delta_{kl}$ follows

$$\begin{aligned}\frac{\partial\chi^2_\gamma(\psi_k)}{\partial\psi_l} = \mathcal{P}_{ij}\Big[&-\varepsilon_i\varepsilon_jZ_j\mathcal{K}_{jk}\delta_{kl}-\varepsilon_iZ_jG_{jk}\delta_{kl}-\varepsilon_i\varepsilon_jZ_i\mathcal{K}_{ik}\delta_{kl}\\ &+\varepsilon_i\varepsilon_jZ_iZ_j\mathcal{K}_{ik}\psi_k\mathcal{K}_{jk}\delta_{kl}++\varepsilon_i\varepsilon_jZ_iZ_j\mathcal{K}_{jk}\psi_k\mathcal{K}_{ik}\delta_{kl}\\ &+\varepsilon_iZ_iZ_j\mathcal{K}_{ik}\psi_kG_{jk}\delta_{kl}+\varepsilon_iZ_iZ_jG_{jk}\psi_k\mathcal{K}_{ik}\delta_{kl}\\ &-\varepsilon_jZ_iZ_jG_{ik}\delta_{kl}+\varepsilon_jZ_iZ_j\mathcal{K}_{jk}\psi_kG_{ik}\delta_{kl}\\ &+\varepsilon_jZ_iZ_jG_{ik}\psi_k\mathcal{K}_{jk}\delta_{kl}+Z_iZ_jG_{ik}\psi_kG_{jk}\delta_{kl}+Z_iZ_jG_{jk}\psi_kG_{ik}\delta_{kl}\Big].\end{aligned} \tag{5.17}$$

From the last equation one can see that we can now write Eq. 5.13 as a linear system of equations

$$\mathcal{B}_{lk}\psi_k = \mathcal{V}_l, \tag{5.18}$$

with the coefficient matrix

$$\begin{aligned}\mathcal{B}_{lk} = \mathcal{P}_{ij}\Big[&\varepsilon_i\varepsilon_jZ_iZ_j\mathcal{K}_{ik}\mathcal{K}_{jl}+\varepsilon_i\varepsilon_jZ_iZ_j\mathcal{K}_{jk}\mathcal{K}_{il}+\varepsilon_iZ_iZ_j\mathcal{K}_{ik}G_{jl}\\ &+\varepsilon_iZ_iZ_jG_{jk}\mathcal{K}_{il}+\varepsilon_jZ_iZ_j\mathcal{K}_{jk}G_{il}+\varepsilon_iZ_iZ_jG_{ik}\mathcal{K}_{jl}\\ &+Z_iZ_jG_{ik}G_{jl}+Z_iZ_jG_{jk}G_{il}\Big],\end{aligned} \tag{5.19}$$

and the result vector

$$\mathcal{V}_l = \mathcal{P}_{ij}\left[\varepsilon_i\varepsilon_j\mathcal{K}_{jl} + \varepsilon_i Z_j G_{jl} + \varepsilon_i\varepsilon_j Z_i\mathcal{K}_{il} + \varepsilon_j Z_i G_{il}\right]. \tag{5.20}$$

Flexion term

The same exercise is done for $x = F$. Again, we perform the calculation only for one component so f and F need to be substituted by either f_1 and F_1, f_2 and F_2, g_1 and G_1 or g_2 and G_2. Starting from Eqs. 4.41-4.42 the problem is linear. This might very well change if we include the so-called reduced flexion as stated in Schneider and Er (2008). For now we stick to the more simple case

$$\begin{aligned}\chi^2_F &= (f_i - Z_i F_i)\,\mathcal{C}^{-1}_{ij}\,(f_j - Z_j F_j)\\ &= \mathcal{C}^{-1}_{ij}\left[f_i f_j - f_i Z_j F_j - f_j Z_i F_i + Z_i Z_j F_i F_j\right],\end{aligned} \tag{5.21}$$

from which we take the derivative with respect to the lensing potential

$$\frac{\partial\chi^2_F}{\partial\psi_l} = \mathcal{C}^{-1}_{ij}\left[-f_i Z_j\frac{\partial}{\partial\psi_l}F_j - f_j Z_i\frac{\partial}{\partial\psi_l}F_j + Z_i Z_j F_j\frac{\partial}{\partial\psi_l}F_i + Z_i Z_j F_i\frac{\partial}{\partial\psi_l}F_j\right]. \tag{5.22}$$

Using $F_i = \mathbf{F}_{ik}\psi_k$ and $\frac{\partial}{\partial\psi_l}\mathbf{F}_{ik}\psi_k = \mathbf{F}_{ik}\delta_{kl}$, we find

$$\frac{\partial\chi^2_F}{\partial\psi_l} = \mathcal{C}^{-1}_{ij}\left[-f_i Z_j\mathbf{F}_{jk}\delta_{lk} - f_j Z_i\mathbf{F}_{ik}\delta_{lk} + Z_i Z_j\mathbf{F}_{jk}\mathbf{F}_{ik}\delta_{lk}\psi_k + Z_i Z_j\mathbf{F}_{ik}\mathbf{F}_{jk}\delta_{lk}\psi_k\right]. \tag{5.23}$$

This transforms Eq. 5.13 into the following LSE

$$\mathcal{B}_{lk} = \mathcal{C}^{-1}_{ij} Z_i Z_j\left[\mathbf{F}_{ik}\mathbf{F}_{jl} + \mathbf{F}_{jk}\mathbf{F}_{il}\right] \tag{5.24}$$

$$\mathcal{V}_l = \mathcal{C}^{-1}_{ij}\left[f_i Z_j\mathbf{F}_{jl} + f_j Z_i\mathbf{F}_{il}\right] \tag{5.25}$$

Multiple image term

The most difficult case is actually $x = m$. Because it would be very confusing otherwise, we drop for a moment the sum convention and consider one multiple-image system s with a total number of images M.

Using $\alpha_a = \mathcal{D}_{ak}\psi_k$ along with $\frac{\partial}{\partial\psi_l}\alpha_a = \mathcal{D}_{al}$, we minimise Eq. 4.50

$$\begin{aligned}\frac{\partial\chi^2_{\mathsf{m}}}{\partial\psi_l} &= \sum_a^M\frac{2}{\sigma^2}\left[\theta_a\mathcal{D}_{al} + \mathcal{D}_{ak}\mathcal{D}_{al}\psi_k - \theta_a\left(-\frac{1}{M}\sum_{n=1}^M\mathcal{D}_{nl}\right) + \mathcal{D}_{ak}\left(-\frac{1}{M}\sum_{n=1}^M\mathcal{D}_{nl}\right)\psi_k\right.\\ &\qquad\left. + \left(\frac{1}{M}\sum_n^M(\theta_n - \mathcal{D}_{nk}\psi_k)\right)\mathcal{D}_{al} - \frac{1}{M}\sum_{n,m}^M(\theta_m\mathcal{D}_{nl} - \mathcal{D}_{mk}\mathcal{D}_{nl}\psi_k)\right]\\ &= \sum_a^M\frac{2}{\sigma_2}\left(\left[-\theta_a\mathcal{D}_{al} + \frac{1}{M}\sum_{n,m}^M(\theta_a\mathcal{D}_{nl} + \theta_n\mathcal{D}_{nl} - \theta_m\mathcal{D}_{nl})\right]\right.\\ &\qquad\left. + \left[\mathcal{D}_{ak}\mathcal{D}_{al} + \frac{1}{M}\sum_{n,m}^M(\mathcal{D}_{nk}\mathcal{D}_{al} - \mathcal{D}_{ak}\mathcal{D}_{nl} + \mathcal{D}_{mk}\mathcal{D}_{nl})\right]\psi_k\right),\end{aligned} \tag{5.26}$$

and find another LSE

$$\mathcal{B}_{lk} = \sum_a^M\frac{2}{\sigma_2}\left[\mathcal{D}_{ak}\mathcal{D}_{al} + \frac{1}{M}\sum_{n,m}^M(\mathcal{D}_{nk}\mathcal{D}_{al} - \mathcal{D}_{ak}\mathcal{D}_{nl} + \mathcal{D}_{mk}\mathcal{D}_{nl})\right] \tag{5.27}$$

$$\mathcal{V}_l = \sum_a^M\frac{2}{\sigma_2}\left[-\theta_a\mathcal{D}_{al} + \frac{1}{M}\sum_{n,m}^M(\theta_a\mathcal{D}_{nl} + \theta_n\mathcal{D}_{nl} - \theta_m\mathcal{D}_{nl})\right]. \tag{5.28}$$

Critical curve term

In the case of $x = c$ we return to Einstein's sum convention in the χ^2-function

$$\chi^2_{\mathsf{c}} = \frac{(\det\mathcal{A})_i^2}{\sigma_i^2} = \frac{((1 - Z_i\kappa_i)^2 - |Z_i\gamma_i|^2)^2}{\sigma_i^2}, \tag{5.29}$$

where the non-linear terms are isolated and taken as a constant for now

$$\begin{aligned}
\frac{\partial\chi^2_{\mathsf{c}}(\psi_k)}{\partial\psi_l} &= \frac{2(\det\mathcal{A})_i}{\sigma_i^2}\frac{\partial}{\partial\psi_l}(\det\mathcal{A}(\psi_k))_i \\
&= \frac{2(\det\mathcal{A})_i}{\sigma_i^2}\left[\frac{\partial}{\partial\psi_l}(1 - Z_i\kappa_i(\psi_k)^2 - \frac{\partial}{\psi_l}|Z_i\gamma_i(\psi_k)|^2\right] \\
&= \frac{2(\det\mathcal{A})_i}{\sigma_i^2}\left[2(1 - Z_i\kappa_i(\psi_k))\left(-Z_i\frac{\partial}{\partial\psi_l}\kappa_i(\psi_k)\right) - \frac{\partial}{\partial\psi_l}\left(Z_i^2\gamma_{1i}^2(\psi_k) + Z_i^2\gamma_{2i}^2(\psi_k)\right)\right] \\
&= \frac{2(\det\mathcal{A})_i}{\sigma_i^2}\left[2(1 - Z_i\kappa_i(\psi_k))(-Z_i\mathcal{K}_{il}) - 2Z_i\gamma_{1i}(\psi_k)Z_iG^1_{il} - 2Z_i\gamma_{2i}(\psi_k)Z_iG^2_{il}\right] \\
&= \frac{4(\det\mathcal{A})_i}{\sigma_i^2}\left[Z_i^2\mathcal{K}_{ik}\mathcal{K}_{il}\psi_k - Z_i\mathcal{K}_{il} - Z_i^2G^1_{ik}G^1_{il}\psi_k - Z_i\mathcal{K}_{il} - Z_i^2G^2_{ik}G^2_{il}\psi_k\right].
\end{aligned} \tag{5.30}$$

The LSE appears as

$$\mathcal{B}_{lk} = \frac{4(\det\mathcal{A})_i}{\sigma_i^2}Z_i^2(\mathcal{K}_{ik}\mathcal{K}_{il} - G^1_{ik}G^1_{il} - G^2_{ik}G^2_{il}) \tag{5.31}$$

$$\mathcal{V}_l = \frac{4(\det\mathcal{A})_i}{\sigma_i^2}Z_i\mathcal{K}_{il}. \tag{5.32}$$

The complete system

When assuming that a reconstruction contains all possible contributions we find the following, total LSE

$$
\begin{aligned}
\mathcal{B}_{lk} = \mathcal{P}^1_{ij} &\left[\varepsilon^1_i\varepsilon^1_j Z_i Z_j \mathcal{K}_{ik}\mathcal{K}_{jl} + \varepsilon^1_i\varepsilon^1_j Z_i Z_j \mathcal{K}_{jk}\mathcal{K}_{il} + \varepsilon^1_i Z_i Z_j \mathcal{K}_{ik} G^1_{jl}\right.\\
&+ \varepsilon^1_i Z_i Z_j G^1_{jk}\mathcal{K}_{il} + \varepsilon^1_j Z_i Z_j \mathcal{K}_{jk} G^1_{il} + \varepsilon^1_i Z_i Z_j G^1_{ik}\mathcal{K}_{jl}\\
&\left.+ Z_i Z_j G^1_{ik} G^1_{jl} + Z_i Z_j G^1_{jk} G^1_{il}\right] \quad \text{(Reduced shear, first comp.)}\\
+\mathcal{P}^2_{ij} &\left[\varepsilon^2_i\varepsilon^2_j Z_i Z_j \mathcal{K}_{ik}\mathcal{K}_{jl} + \varepsilon^2_i\varepsilon^2_j Z_i Z_j \mathcal{K}_{jk}\mathcal{K}_{il} + \varepsilon^2_i Z_i Z_j \mathcal{K}_{ik} G^2_{jl}\right.\\
&+ \varepsilon^2_i Z_i Z_j G^2_{jk}\mathcal{K}_{il} + \varepsilon^2_j Z_i Z_j \mathcal{K}_{jk} G^2_{il} + \varepsilon^2_i Z_i Z_j G^2_{ik}\mathcal{K}_{jl}\\
&\left.+ Z_i Z_j G^2_{ik} G^2_{jl} + Z_i Z_j G^2_{jk} G^2_{il}\right] \quad \text{(Reduced shear, second comp.)}\\
+\mathcal{C}^1_{ij} &\left[\mathbf{F}^1_{ik}\mathbf{F}^1_{jl} + \mathbf{F}^1_{jk}\mathbf{F}^1_{il}\right] Z_i Z_j \quad (F-\text{flexion, first comp.})\\
+\mathcal{C}^2_{ij} &\left[\mathbf{F}^2_{ik}\mathbf{F}^2_{jl} + \mathbf{F}^2_{jk}\mathbf{F}^2_{il}\right] Z_i Z_j \quad (F-\text{flexion, second comp.})\\
+\mathcal{C}^1_{ij} &\left[\mathbf{G}^1_{ik}\mathbf{G}^1_{jl} + \mathbf{G}^1_{jk}\mathbf{G}^1_{il}\right] Z_i Z_j \quad (G-\text{flexion, first comp.})\\
+\mathcal{C}^2_{ij} &\left[\mathbf{G}^2_{ik}\mathbf{G}^2_{jl} + \mathbf{G}^2_{jk}\mathbf{G}^2_{il}\right] Z_i Z_j \quad (G-\text{flexion, second comp.})\\
+\frac{2}{\sigma_2} &\left[\mathcal{D}^1_{ak}\mathcal{D}^1_{al} + \frac{1}{M}\sum_{n,m}^{M}\left(\mathcal{D}^1_{nk}\mathcal{D}^1_{al} - \mathcal{D}^1_{ak}\mathcal{D}^1_{nl} + \mathcal{D}^1_{mk}\mathcal{D}^1_{nl}\right)\right] \quad \text{(M. systems, first comp.)}\\
+\frac{2}{\sigma_2} &\left[\mathcal{D}^2_{ak}\mathcal{D}^2_{al} + \frac{1}{M}\sum_{n,m}^{M}\left(\mathcal{D}^2_{nk}\mathcal{D}^2_{al} - \mathcal{D}^2_{ak}\mathcal{D}^2_{nl} + \mathcal{D}^2_{mk}\mathcal{D}^2_{nl}\right)\right] \quad \text{(M. systems, second comp.)}\\
+Z^2_b &\left[\mathcal{K}_{bk}\mathcal{K}_{bl} - G^1_{bk}G^1_{bl} - G^2_{bk}G^2_{bl}\right]\frac{4(\det\mathcal{A})_b}{\sigma^2_b} \quad \text{(Critical curve estimator)}
\end{aligned}
\tag{5.33}
$$

$$
\begin{aligned}
\mathcal{V}_l = \mathcal{P}^1_{ij} &\left[\varepsilon^1_i\varepsilon^1_j\mathcal{K}_{jl} + \varepsilon^1_i Z_j G^1_{jl} + \varepsilon^1_i\varepsilon^1_j Z_i\mathcal{K}_{il} + \varepsilon^1_j Z_i G^1_{il}\right] \quad \text{(Reduced shear, first comp.)}\\
+\mathcal{P}^2_{ij} &\left[\varepsilon^2_i\varepsilon^2_j\mathcal{K}_{jl} + \varepsilon^2_i Z_j G^2_{jl} + \varepsilon^2_i\varepsilon^2_j Z_i\mathcal{K}_{il} + \varepsilon^2_j Z_i G^2_{il}\right] \quad \text{(Reduced shear, second comp.)}\\
+\mathcal{C}^1_{ij} &\left[f^1_i Z_j \mathbf{F}^1_{jl} + f^1_j Z_i \mathbf{F}^1_{il}\right] \quad (F-\text{flexion, first comp.})\\
+\mathcal{C}^2_{ij} &\left[f^2_i Z_j \mathbf{F}^2_{jl} + f^2_j Z_i \mathbf{F}^2_{il}\right] \quad (F-\text{flexion, second comp.})\\
+\mathcal{C}^1_{ij} &\left[g^1_i Z_j \mathbf{G}^1_{jl} + g^1_j Z_i \mathbf{G}^1_{il}\right] \quad (G-\text{flexion, first comp.})\\
+\mathcal{C}^2_{ij} &\left[g^2_i Z_j \mathbf{G}^2_{jl} + g^2_j Z_i \mathbf{G}^2_{il}\right] \quad (G-\text{flexion, second comp.})\\
+\frac{2}{\sigma_2} &\left[-\theta^1_a\mathcal{D}^1_{al} + \frac{1}{M}\sum_{n,m}^{M}\left(\theta^1_a\mathcal{D}^1_{nl} + \theta^1_n\mathcal{D}^1_{nl} - \theta^1_m\mathcal{D}^1_{nl}\right)\right] \quad \text{(M. systems, first comp.)}\\
+\frac{2}{\sigma_2} &\left[-\theta^2_a\mathcal{D}^2_{al} + \frac{1}{M}\sum_{n,m}^{M}\left(\theta^2_a\mathcal{D}^2_{nl} + \theta^2_n\mathcal{D}^2_{nl} - \theta^2_m\mathcal{D}^2_{nl}\right)\right] \quad \text{(M. systems, second comp.)}\\
+Z_b\mathcal{K}_{bl}&\frac{4(\det\mathcal{A})_b}{\sigma^2_b} \quad \text{(Critical curve estimator).}
\end{aligned}
\tag{5.34}
$$

Please note some remarks. Superscripts are actually indices in the two equations above; $i, j \in [0, ..., N_{\text{pix}}]$; a runs over all images in a multiple image system, furthermore it has to be summed also over the total number of multiple-image systems in the reconstruction, which has been suppressed for convenience; and b runs over all pixels in the AMR grid , which are selected to be part of the critical curve.

5.4.2 Speeding things up

The by far most time consuming part of a reconstruction run is building up the LSE given by Eqs. 5.33, 5.34. This can be easily seen by a little back-of-the envelope calculation:

- We assume the AMR reconstruction grid to have 2000 pixels. This means that the indices $l, k, i, j \in [0, ..., 1999]$.
- All constraints will be used, but due to the fact that $\max(a, b, n, m) \sim 50$ we will ignore the strong lensing contributions.
- One element in $\mathcal{B}_{lk}$ is a sum over the indices i and j, one element in this sum consists of 24 terms, each of which contains is a product of around 6 numbers.
- This gives as a total number of arithmetic operations to build up $\mathcal{B}_{lk}$

$$2000 \cdot 2000 \cdot 2000 \cdot 2000 \cdot 24 \cdot 6 = 2.304 \cdot 10^{15} \tag{5.35}$$

- The same calculations for $\mathcal{V}_l$ adds no relevant contribution.
- In principle this value is still assumed low since we need to build up the LSE several times as we will see in the next section.
- By assuming a performance of $\sim$ 3 GFLOPS for a normal desktop computer, we calculate a runtime of $\sim$ 120 days to build up the LSE.

Such a runtime is of course not feasible so we present three different strategies, which actually reduce the runtime to the time scale of minutes. (See also Sec. 6.2).

Avoiding repetitions

The fact that the matrices $\mathcal{P}$ and $\mathcal{C}$ are symmetric and a close look at the terms of the reduced shear in Eqs. 5.33, 5.34 reveal, that several terms become indeed identical while summing over i and j. Thus, we perform only the sums for one term and multiply by the number of repetitions afterwards.

Avoiding multiplications with zero

One of the most expensive operations is the multiplication with the finite-differencing matrices like e.g. $\mathcal{K}$ or $\mathbf{F}^1$. As it is shown in Fig. 5.12, these matrices are, by construction, sparse band matrices and most of their components are zero. A sum over their full column-length is therefore pointless and our method applies specific summation schemes, which only sum over the non-zero entries of the finite-differencing matrices. This procedure actually results in a tremendous speed-up without which the whole method would not be feasible.

Dividing the work

Since the creation of an individual entry in $\mathcal{B}_{lk}$ and $\mathcal{V}_l$ is completely independent of the other entries, it is a perfect candidate for parallelisation. N^2_{pix} threads are created on the GPU device and each matrix and vector element is calculated on a separated core. Also this speed-up is tremendous. To get an idea of it one is refered to Sec. 6.2.

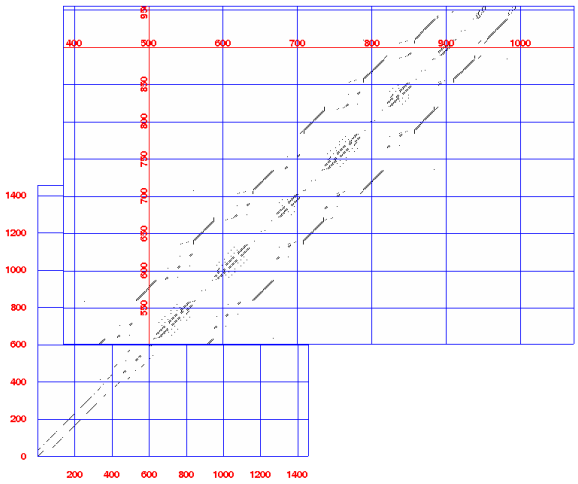


Figure 5.12: The finite-differencing matrix associated with the convergence operator. The dimensions are 1450×1450 pixels and one can see the sparse band structure, especially in the zoom on the diagonal.

5.4.3 Solving the LSE

After the full LSE has been created, the lensing potential is derived by solving it. In our case, this is done by using advanced Gauss-Jordan algorithms (see Rannacher, 2006). The computation time for solving a linear system of this dimension is negligible in comparison to the time needed to build up the linear system.

5.5 An iterative approach

The last section showed in great detail how to perform a reconstruction from the input data. Unfortunately, this is not enough yet to guarantee a satisfactory result. We will have to perform several of these single reconstructions in an iterative way. For this reason, it becomes very important again, that we implemented our method in a runtime-efficient way. Before we describe the two different types of iterations, that we use in our method, we will have to introduce the concept of regularisation.

5.5.1 Regularisation

In statistical methods incorporating a large amount of degrees-of-freedom (DOE) the problem of overfitting may appear. The model does not only fit the data but it tries to fit noise-patterns. It might even happen that overfitting becomes the dominant factor during the reconstruction, especially when using high spatial resolution. This highly contaminates the final result. Regularisation is a technique to avoid overfitting and to guarantee a smooth result. It introduces a penalty function, which punishes certain behaviours of the fitting process or deviations from pre-defined guidelines. In our method, this is achieved by introducing a regularisation function R in the χ^2-function, which also depends on the lensing potential and becomes large if the model tends to overfit

$$\chi^2 = \chi^2_{\mathsf{w}}(\psi) + \chi^2_{\mathsf{s}}(\psi) + \eta R(\psi). \tag{5.36}$$

The regularisation parameter η controls the strength of the penalty and should be chosen such, that the overall χ^2 per DOE is of order unity.
Several different approaches to define the regularisation function have been made, like maximum-entropy regularisation (Narayan and Nityananda, 1986; Bridle et al., 1998; Seitz et al., 1998) which seems well suited but we will follow the much simpler and easier-to-implement approach by Bradač et al. (2005), which has proven its ability to deliver excellent results in the past.
By comparing to a previous result, the reconstructed model is not allowed to deviate in its convergence by a too large amount. Therefore, we penalise these deviations with the following regularisation function

$$R = \eta_i \left(\kappa_i^{\text{previous}} - \kappa_i(\psi) \right)^2 . \tag{5.37}$$

Minimising Eq. 5.37 leads to

$$\begin{aligned} \frac{\partial R(\psi_k)}{\partial \psi_l} &= \eta_i \frac{\partial}{\partial \psi_l} \left(\kappa_i^{\mathrm{p}} - \kappa_i(\psi_k) \right)^2 \\ &= 2\eta_i (\kappa_i^{\mathrm{p}} - \kappa_i) \left(-\frac{\partial}{\partial \psi_l} \mathcal{K}_{ik} \psi_k \right) \\ &= 2\eta_i \left(\kappa_i^{\mathrm{p}} - \mathcal{K}_{ik}\psi_k \right) \left(-\mathcal{K}_{il} \right) \\ &= 2\eta_i \left(-\kappa_i^{\mathrm{p}} \mathcal{K}_{il} + \mathcal{K}_{ik}\mathcal{K}_{il}\psi_k \right) , \end{aligned} \tag{5.38}$$

which contributes one additional term to the coefficient matrix and the result vector.

$$\mathcal{B}_{lk}^{\text{reg}} = \eta_i \mathcal{K}_{ik} \mathcal{K}_{il} \tag{5.39}$$

$$\mathcal{V}_{l}^{\text{reg}} = \eta_i \kappa_i^{\mathrm{p}} \mathcal{K}_{il}. \tag{5.40}$$

Additional regularisation terms for the shear or the flexion are also possible.
The question how to obtain the previous result leads us to the first type of iteration.

5.5.2 Outer-level iteration

The whole point of regularisation is to avoid overfitting noise-patterns in the data. These noise patterns will become more and more pronounced as the grid resolution is increased, because the statistical sample, which was used to calculate the mean of shear or flexion in a pixel, becomes smaller and smaller. Furthermore, the correlations between the pixels as introduced in Eq. 5.2 will become more and more pronounced. This will not be the case for a very small resolution where the statistics for an individual pixel are good and correlations within the pixels are not present. Following this idea, we start our reconstruction on a very coarse resolution, by also assuming that the initial convergence, shear and flexion are flat and zero. We regularise on this flat prior and also insert it into all the non-linear terms of Eqs. 5.33, 5.34. After the reconstruction is finished on the coarse resolution, we interpolate the resulting lensing potential to a slightly higher resolution (e.g. increased by 2 pixels in both dimensions) by using a reliable bicubic-spline interpolation. We take this interpolation as a reference for insertion into the non-linear terms and as a regularisation template. We repeat this procedure until the final reconstruction is reached, which is given by the fact that pixel correlations become too strong and the covariance matrices (Eq. 5.2) become singular, thus non-invertible. Typical start and stop resolutions are visualised in Fig. 5.13.

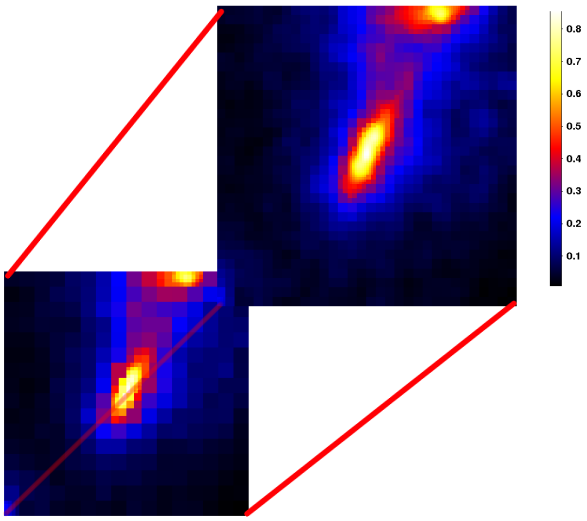


Figure 5.13: The grid resolution is continuously increased in the outer-level iteration. Starting from a low start resolution, which is too coarse for noise-patterns to appear, the high, final reconstruction resolution gives the result.

5.5.3 Inner-level iteration

We finally return to the problem of non-linear terms in Eqs. 5.15, 5.29, which are just held constant during the build-up of the LSE and fix it with a second type of iteration. As we have already mentioned, as a first step in each outer-level iteration the non-linearities are substituted with the values of the regularisation templates. This can serve as a first step but will not lead to an accurate result. By adopting the scheme as in the direct-inversion method of Seitz and Schneider (1995), we repeatedly insert the result of a previous reconstruction into the non-linear terms until convergence is achieved. We confirm this convergence by monitoring the change in the convergence map of the cluster *and explicitly apologise for the name confusion*. Usually, the reconstruction converges towards a unique solution within 2-5 steps. This inner-level iteration is embedded within each outer-level iteration step. A flowchart diagram, illustrating both iteration levels, can be seen in Fig. 5.14.

5.6 A complete reconstruction package

The last section finished our description of the different steps, which are applied in the course of a complete reconstruction. We will take the time to summarise and order these steps and to describe the actual result and how to process it further.

5.6.1 A summary: From the input to the result

Fig. 5.14 summarises the workflow of a complete reconstruction run. Several parameters have to be set in the beginning, including:

- The output field of the reconstruction,
- The start and stop resolutions (given in terms of refinement level 0 of the AMR grids),
- The minimum number of galaxies used for averaging in one pixel,

- Regularisation parameters, consisting of the quantities to be regularised on and the regularisation strength.

Of course, also the input catalogues have to be given, including:

- Ellipticities,
- Flexion,
- Multiple-image systems,
- Critical-curve estimators.

Not all of this catalogues have to be used, but it should be noted that a reconstruction, only based on strong-lensing constraints, is not possible. Either ellipticity or flexion catalogues have to be present. Pure weak lensing reconstructions based only on ellipticity or/and flexion are possible, but the advantages of combining weak and strong lensing are obviously not considered.
Once input and parameters are given, the reconstruction starts on the initial resolution, by assuming a flat convergence and shear profile. It would also be possible to assume an initial profile (Bradač et al., 2005) but this prior would spoil the nonparametric nature of our method. Another idea would be to use the outcome of a preceding direct-inversion technique. Both possibilities are also included in our method.
By sequentially increasing the reconstruction resolution and performing the necessary number of iterations to solve the non-linearities within the minimisation process, the final reconstruction resolution is reached. The result is the lensing potential on a highly-resolved, adaptively-refined grid.

5.6.2 Analysing the result

We obtain a physically meaningful quantity like the convergence by simply applying the Laplacian to the potential. To obtain the shear, we have to apply the other combinations of second derivatives to the potential. The convergence is a reasonable quantity, because it shows us in an intuitive way the mass distribution of the cluster through the surface-mass density.
Unfortunately, the convergence is again affected by the mass-sheet degeneracy . It can be transformed like (Gorenstein et al., 1988)

$$\kappa(\boldsymbol{\theta}, z) \rightarrow \kappa^{'}(\boldsymbol{\theta}, z) = (1 - \lambda) + \lambda\kappa(\boldsymbol{\theta}, z). \tag{5.41}$$

So how to fix the convergence? If our observed field is sufficiently large, we would assume that $\kappa \rightarrow 0$ for the outskirts of our field, so we can use Eq. 5.41, to satisfy this condition. A more elaborate method is the use of a quantity in the reconstruction, which is not invariant under the mass-sheet transformation and determines potential and convergence. One possibility would be the use of inverse magnification (Broadhurst et al., 1995)

$$R = \frac{1}{\mu} = (1 - \kappa)^2 - \gamma^2 \approx 1 - 2\kappa, \tag{5.42}$$

embedded in the maximum-likelihood approach

$$\chi^2_{\text{mag}}(\psi) = \frac{(R - R(\psi))^2_i}{\sigma^2_{Ri}}. \tag{5.43}$$

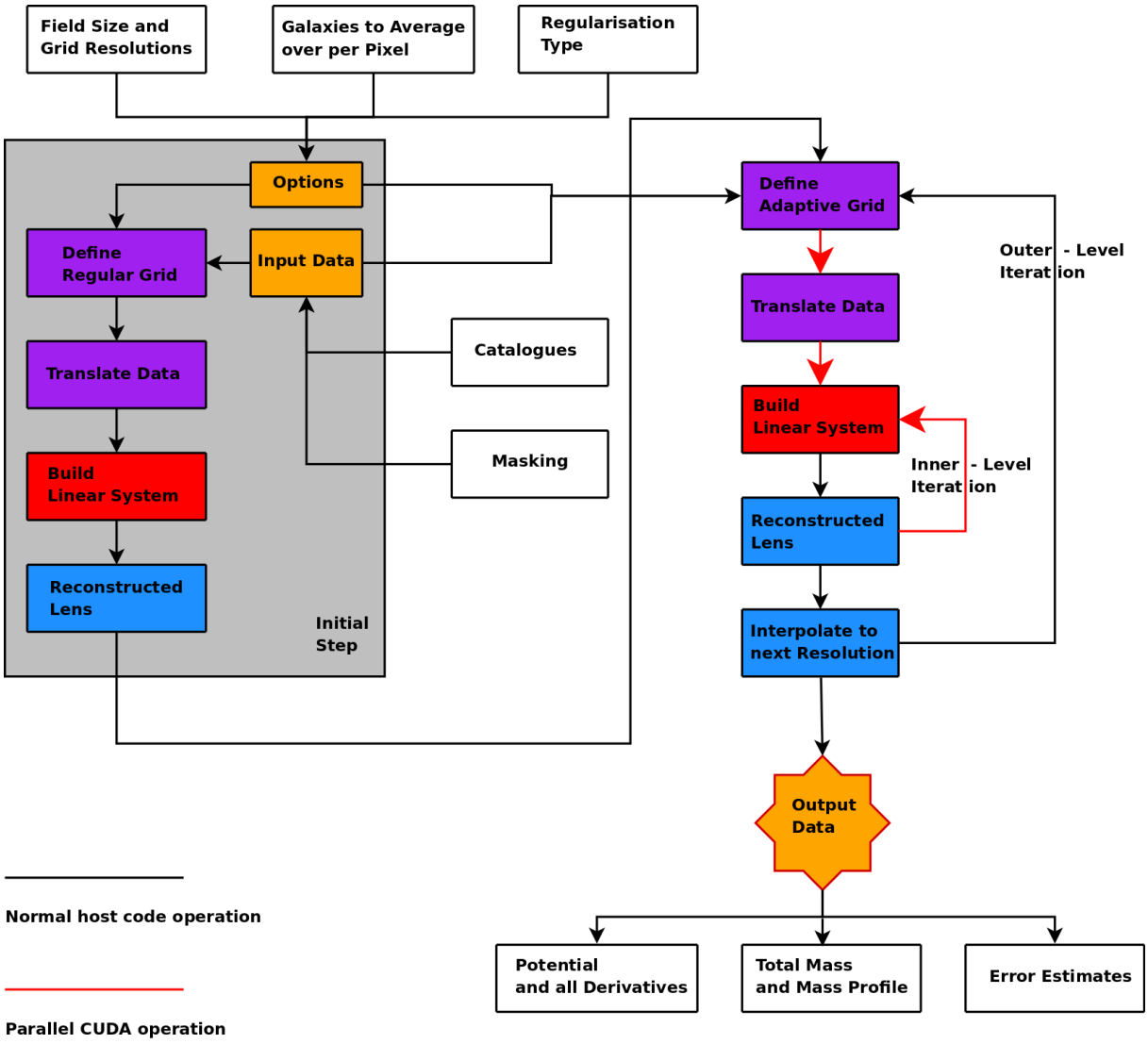


Figure 5.14: The workflow of our reconstruction method. After input catalogues and parameter definition follows an initial reconstruction step, assuming flat convergence and shear priors. The result is fed into the two-level iteration loop which leads to the final result.

Of course in order to use this method, one has to provide a good idea how to measure a magnification signal, but promising results including magnification have been made recently on cosmic scales (Hildebrandt et al., 2009; van Waerbeke, 2010)

$$\mu = \frac{d^2\Omega_{\text{lensed}}}{d^2\Omega_{\text{unlensed}}}. \tag{5.44}$$

Another approach to solve the transformation problem was proposed by Bradač et al. (2004, 2005), who used the knowledge of the redshift distribution of the sources.
Once we have the convergence possibly transformed with the methods shown above, it is quite easy to get a mass estimate for our lens. Convergence and surface-mass density are connected by

$$\Sigma(D_{\text{d}}\boldsymbol{\theta}) = \frac{c^2}{4\pi G}\frac{D_{\text{s}}}{D_{\text{d}}D_{\text{ds}}}\kappa(\boldsymbol{\theta}). \tag{5.45}$$

If we have the redshift of the lens, we also know the physical area of one pixel and if we have at least the mean redshift of the sources, we can calculate the surface mass density, which gives us the total mass of the cluster after summing over the whole grid. Radial profiles of the surface-mass density or the total mass within radius-bins might be produced, but we would like to mention that our method specifically tries to avoid the assumptions of radial symmetry.
Error bars are attached to all derived quantities by bootstrapping the input-catalogues or by exploring the parameter space through MCMC-methods as described in Sec. 4.5.

5.6.3 Concrete implementation

We decided to implement our method in C++, which is a powerful and highly flexible programming language. The use of object orientation makes the produced code easier to extend and more accessible to others. A large collection of publicly-available packages are available for C++ and we make use of the following:

- The GNU Scientific Library[9] (GSL), is a very substantial compilation of mathematical routines and tools, for scientific applications and is available for C and C++. It covers a whole bunch of different areas, like differential equations, polynomials, minimization methods, numerical integration and Fourier transformation. It is very well documented and easy to use. In our code we use GSL for vector and matrix handling, random number generation and statistics.

- The Linear Algebra Package[10] (LAPACK) was originally written in Fortran 77, but is also available in C and C++. It is highly optimized for linear algebra operations and in this sector still the best package available[11]. In our code we use the LAPACK routines, contained in the Automatically Tuned Linear Algebra Software[12] (ATLAS), for high-dimension matrix inversion and to solve linear systems of equations.

- The common standard in Astronomy for data transport is the Flexible Image Transport System (FITS). We also make wide use of this data format to store information on disk

[9] http://www.gnu.org/software/gsl
[10] http://www.netlib.org/lapack
[11] If the reader remembers us mentioning that most codes are implemented in an inefficient way with respect to memory access, this is certainly not the case for these routines.
[12] http://www.netlib.org/atlas

during the reconstruction process. The library which is needed by C++ to read and write files in the FITS-format is CFITSIO[13] and the C++ wrapper CCfits[14].

- For CPU-based parallelisation we use OpenMP[15] and the Message Passing Interface[16] (MPI).
- For GPU-based parallelisation we use CUDA[17] and the CUDA Software Development Kit (NVIDIA, 2009c,b,a).

The complete code package is of medium size ($\sim$ 12000 lines of code) and the host code makes extensive use of object orientation. Classes to represent the general parameter-structure, gravitational lenses, AMR grids and input constraints are implemented with comprehensive functionality. This makes it possible to use those classes as a general library also in other projects. The CPU-demanding routines are implemented in MPI and CUDA and are highly optimised for numerical performance.
A point which cannot be highlighted enough is documentation. All header-files of the code are fully documented and class overviews can be created by the use of e.g. Doxygen[18]. A comprehensive manual to the complete package is available (Merten, 2009) and the distribution of the code is done via the version-management software Subversion[19]. All libraries and packages, mentioned in this section, are publicly available.

[13] http://heasarc.gsfc.nasa.gov/fitsio
[14] http://heasarc.gsfc.nasa.gov/fitsio/CCfits/
[15] http://openmp.org/wp/
[16] http://www.mcs.anl.gov/research/projects/mpi/
[17] http://www.nvidia.com/object/cuda_home_new.html
[18] http://www.stack.nl/ dimitri/doxygen/
[19] http://subversion.tigris.org/

6 Proofs of concept

Complex numerical algorithms must not be applied to real data, if they have not been tested carefully with synthetic data. In this chapter, we will test the most important parts of our implementation regarding two crucial aspects.

First, we test the numerical performance of our parallel `CUDA` implementation, compared to simpler implementations without sophisticated numerical schemes and without parallelisation. In Sec. 6.1 we test the performance of the second-most time consuming step in our method, which is the process of finding pixel correlations in the reconstruction grid. In Sec. 6.2 we focus on the most time consuming step in the algorithm, which is the construction of the coefficient matrix, shown in Eq. 5.33. Sec. 6.3 concludes the proofs of concept with the reconstruction of an idealised, synthetic cluster to test the accuracy of the whole package. Such a test is crucial, since it probes the full implementation and, if passed, allows the application of the code to more realistic data, where non-perfect results might not be due to the method, but due to the noisy input data.

6.1 Galaxy associations

While reconstructing cluster fields on highly refined grids, neighbouring cells become heavily correlated as we pointed in Sec. 5.3.2. This need not be a problem, using the techniques that we have described before. The problem arises from the fact that weak-lensing catalogues from observations with a wide field-of-view contain a large number of background galaxies, which renders it time consuming to derive the correlations between the individual reconstruction pixels. In our implementation, every background galaxy with ellipticity measurement is associated through the adaptive-shape-averaging technique described in Sec. 5.3.1, to particular grid pixels. After this association, the covariances between grid pixels are calculated, by using the number of galaxies that two different grid pixels have in common (see Eq. 5.3). This means, that for every pixel pair within the reconstruction grid, the full input catalogue has to be searched and shared background galaxies have to counted. This time consuming procedure is ideal for parallelisation since the correlations for the different grid pixels can be treated separately.

6.1.1 Producing a synthetic catalogue

To create a catalogue with a realistic number of background galaxies, we used the lensing potential of a simulated galaxy cluster with a resolution of 512×512 pixels, as shown in Fig. 6.1. We randomly sampled 10000 positions in this simulated cluster field and analytically calculated the reduced shear at these positions. This delivered a synthetic weak-lensing catalogue with 10000 reduced-shear measurements, as it is also shown in Fig. 6.1.

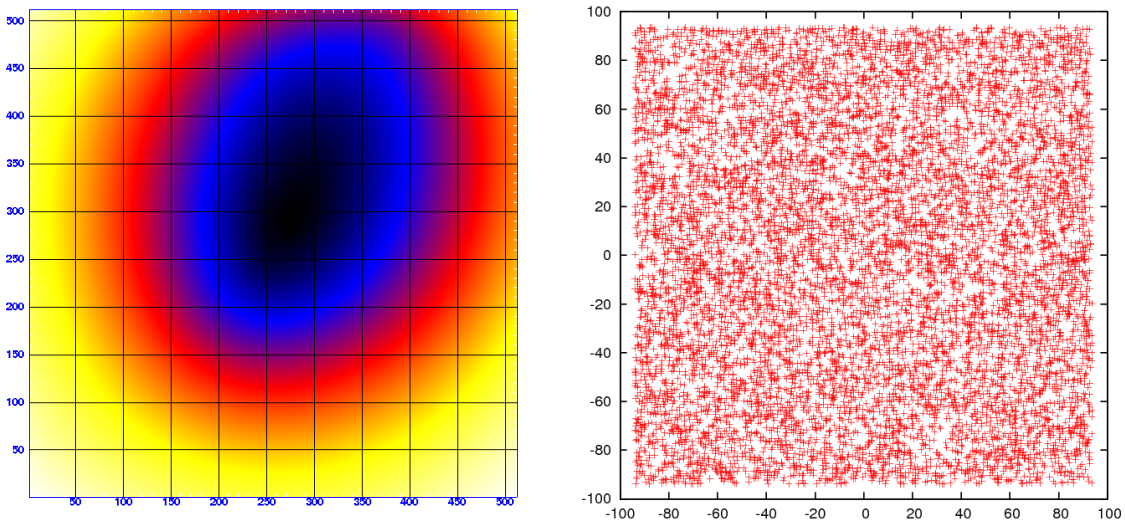


Figure 6.1: The *left panel* shows the lensing potential of a simulated galaxy cluster, discretised on a regular grid of 512×512 pixels. In the *right panel* 10000 of these grid positions are randomly sampled to create a synthetic weak-lensing catalogue. The axes ticks in the *right panel* give the distance from the cluster centre in arcsec.

method	runtime [sec]
single-thread	759
GPU	13

Table 6.1: The runtime comparison between a single-thread routine, associating galaxies and covariances to a grid pixel and a multi-thread GPU routine. The grid size of the AMR grid is 50×50 pixels. The test was performed on a GPU server with two INTEL XEON E5420 quad-core processors, 8GB of main memory and equipped with a NVIDIA TESLA C1060 GPU board with 240 streaming cores and 4GB of graphics memory.

6.1.2 Runtime results and conclusion

We performed the galaxy association and covariance detection with two different routines. First, using a single-thread routine and afterwards, a massively-parallel routine, implemented in `CUDA`. The runtime comparison is shown in Tab. 6.1. Obviously, the combined numerical power of the 240 available cores on the GPU device helped dramatically to speed up the calculation. This speed-up is especially important for large weak-lensing catalogues, containing even more shear measurements than the 10000, used in this run. The result of the correlation detection can be seen in Fig. 6.2, where the number of shared galaxies between a grid pixel with index i and a grid pixel with index j is shown as a matrix element $\mathcal{C}_{ij}$. The matrix $\mathcal{C}$ is of dimension N_{pix}^2, with N_{pix} being the number of pixels in the reconstruction grid. Obviously, $\mathcal{C}$ is symmetric and that should be taken into account during its construction.

6.2 Building up the coefficient matrix

We have already mentioned in Sec. 5.4.2, that the by far most time-consuming step during the reconstruction process is the build-up of the linear system of equations and its coefficient

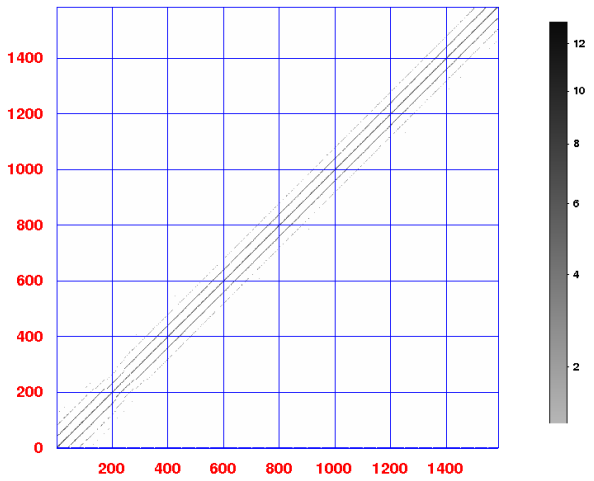


Figure 6.2: Shown here is the correlation in a 40×40 pixel reconstruction grid in matrix representation. We express this correlation by the number of galaxies that two pixels i and j have in common. The main diagonal shows the autocorrelation of the grid pixels and the distance to the upper and lower neighbours in the grid can be read off from the distance of the sub-diagonals to the main diagonal. One can also see, that the correlation is decreasing with increasing distance from the main diagonal, so with increasing distance between the pixels with index i and j, for which the correlation is evaluated.

matrix in particular. We suggested three recipes to speed this process up, where the first, being the fact that certain terms in the coefficient matrix are identical, is trivial to realise. The other two ideas, the usage of specialised summation schemes and the parallelisationof the whole process, will be tested in the following.

6.2.1 A toy problem

To simplify the test, whose only intention is to show the gain in numerical performance, we do not create the complicated coefficient matrix from Eq. 5.33, but we will simulate with random numbers the typical structure of a term in this matrix. The runtime test consists of creating the square matrix $\mathcal{B}$ whose elements are given by

$$\mathcal{B}_{lk} = a_i b_j \mathcal{C}_{ij} \mathcal{D}_{il} \mathcal{E}_{jk}. \tag{6.1}$$

The vectors a and b have the same length as the dimension of the square matrix $\mathcal{B}$. $\mathcal{C}$ is the equivalent of a covariance matrix in a real reconstruction and has the same dimension as $\mathcal{B}$. $\mathcal{D}$ and $\mathcal{E}$ represent the sparse finite-differencing matrices, introduced in Sec. 4.1.3. While applying no numerical optimisation, which makes use of the special structure of $\mathcal{D}$ and $\mathcal{E}$, the summation over the indices i and j in Eq. 6.1 runs over the full dimensionality of $\mathcal{B}$, in other words

$$l,\ k,\ i,\ j \in [0, ..., N_{\text{pix}}^2], \tag{6.2}$$

where N_{pix} is the number of pixels in the reconstruction grid that we consider.

method	runtime [sec]
full summation	2282
advanced summation	2

Table 6.2: The speed-up of a single-thread routine, while making use of special summation schemes and exploiting the simple structure of the finite differencing matrices. The test was run on a INTEL Core2 Duo P8600 with 4GB of main memory. The number of pixels in this test was fixed to $N_{\text{pix}} = 400$ (see text). The low number of pixels was chosen low to keep the runtime of the full-summation routine at an acceptable level.

method	runtime [sec]
single-thread	82
GPU	1.03

Table 6.3: The same test as in Tab. 6.2 using the advanced summation scheme and run on the machine of Tab. 6.1, using also the capabilities of the GPU device . The number of pixels in this test was fixed to $N_{\text{pix}} = 2500$.

6.2.2 Runtime results and conclusion

We have already mentioned that the finite-differencing matrices are sparse and have band structure, which makes it possible to speed up the summation in Eq. 6.1 significantly. Depending on which element of $\mathcal{B}$ is considered, only a few entries in $\mathcal{D}$ and $\mathcal{E}$ are different from zero. As we see from Eq. 6.1, the index l defines the non-zero elements of $\mathcal{D}$ in the summation over i and k defines the non-zero elements of $\mathcal{E}$ in the summation over j. We obtain the following transformation

$$i \rightarrow i^*(\mathcal{D}, l), \tag{6.3}$$

$$j \rightarrow j^*(\mathcal{E}, k). \tag{6.4}$$

Depending on the finite differencing matrix that we consider, the number of summation with respect to i or j reduces from N_{pix}^2 to between three (Eqs. 4.12, 4.13) and nine (Eq. 4.14). We explore the speed-up of using this advanced summation scheme, compared to the full summation scheme in Tab. 6.2.

The speed-up, when comparing the advanced summation scheme running as a single thread, to the scheme running on a massively parallel GPU is shown in Tab. 6.3.

It seems obvious, how heavily the reconstruction algorithms gain from these special implementation techniques, especially when considering high resolution reconstructions with a large number of iterations, which makes it necessary to build up the linear system of equations up to 200 times. We would like to point out, that all the techniques to accelerate the calculation are exact and perform not a single approximation, they just exploit the ideas, presented in Sec. 5.4.2.

6.3 Reconstruction of idealised input data

In order to test the complete reconstruction package, we considered again the cluster potential shown in Fig. 6.1 and tried to reconstruct the underlying cluster-mass distribution.

6.3.1 Synthetic catalogues

We derived the weak-lensing catalogue in the same way as in Sec. 6.1.1. To simulate a somewhat realistic scenario for such a small cluster field, we sampled only 2000 reduced-shear values. In addition, we calculated the position of the critical curve for this cluster. The resulting catalogues are visualised in Fig. 6.3.

6.3.2 Results and conclusion

The result of the reconstruction is shown as a comparison between the original density profile on high resolution and the reconstructed convergence map in Fig. 6.4. One can already see by eye, that the reconstruction is able to recover the density profile of the simulated cluster almost perfectly. A more quantitative analysis is presented in Fig. 6.5 by a density profile along the main diagonal of the field. In that figure, also the results of reconstructions without the additional strong-lensing constraints and the results without the usage of the adaptively refined resolution in the cluster centre are shown. The following conclusions shall be noticed:

1. A pure weak-lensing reconstruction underestimates the steepness of the density profile. This is not surprising, since we sampled only a very limited number of points to constrain the cluster field, thereby possibly neglecting localised peaks and our algorithm averages over a number of reduced-shear values per reconstruction pixel. This is a general trend of weak-lensing reconstructions, especially when applied to real observations.

2. By adding strong-lensing constraints, the trend described above is significantly suppressed.

3. By also using the advantages of an enhanced central resolution with the help of our adaptive grid, we can almost perfectly recover the original profile even if we have sampled the cluster only with a very limited number of constraints.

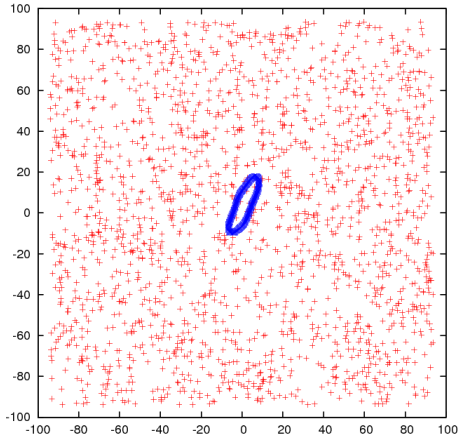


Figure 6.3: The input catalogues for the synthetic cluster reconstruction. Shown in red are the 2000 positions of the reduced-shear sample and in blue one can see the position of the cluster's critical line. The axes give the distance from the cluster centre in arcsec.

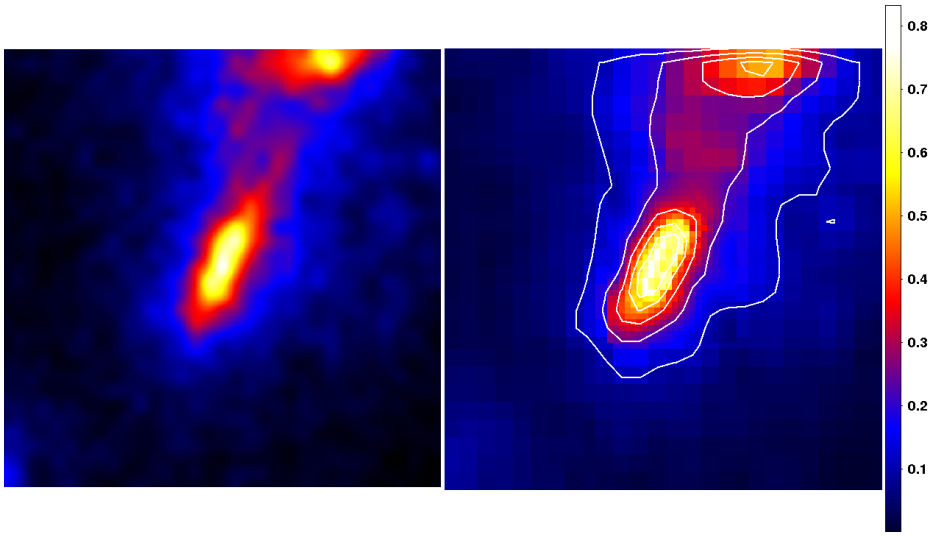


Figure 6.4: The *left panel* shows the smoothly resolved, real convergence map of the simulated cluster on a grid of 512×512 pixels. The *right panel* shows our reconstruction with a pixel resolution refering to 32×32 pixels in the outskirts of the cluster and to 75×75 pixels in its innermost core. The contours start at $\kappa = 0.1$ with a linear spacing of $\Delta\kappa = 0.12$.

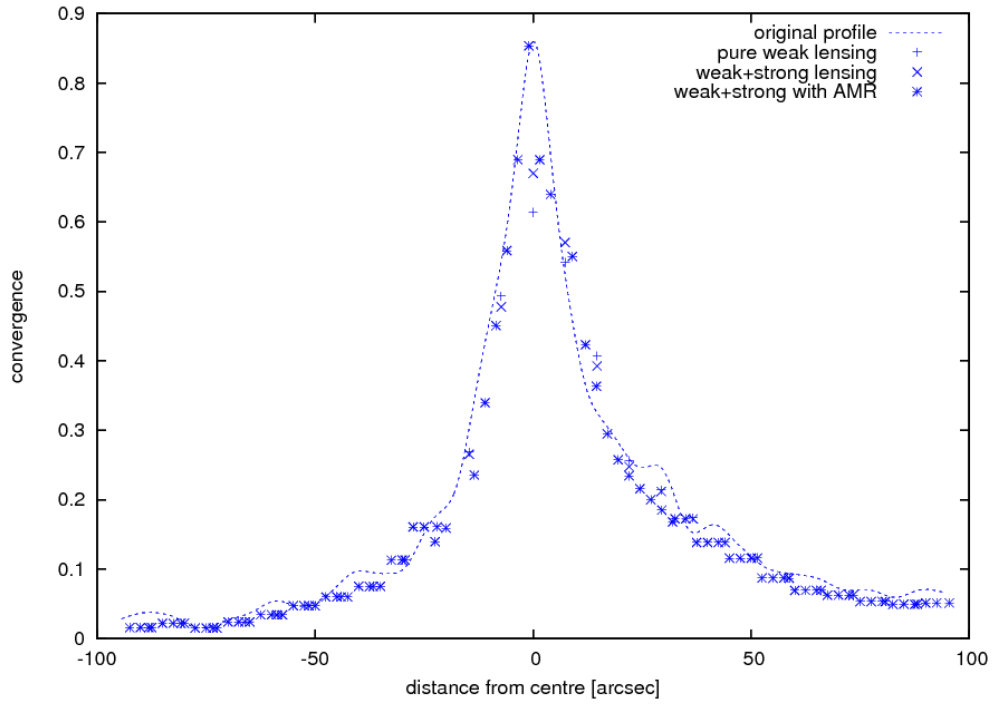


Figure 6.5: The convergence profile along the main diagonal of the reconstructed field. The dotted line shows the original profile of the cluster. The (+)-like data points refer to the reconstruction without the usage of any strong-lensing constraints. The (x)-like data points make use of strong-lensing constraints, but not of the AMR grids. The (*)-like data points use the full potential of our method, while combining weak and strong lensing on adaptively refined grids.

7 Weighing simulated galaxy clusters

As the title of the last chapter already indicates, even the reconstruction of the simulated cluster in Sec. 6.3 can at best be seen as a proof of concept for several reasons. First, the numerical simulation from which this cluster was extracted, was a pure dark-matter N-body simulation, neglecting the effects of baryonic physics. Such a simulation cannot compete any more with modern simulations of galaxy clusters, since the interesting and important effects within the ICM are not considered and cannot be explored. So even if a pure dark-matter cluster is reconstructed properly with our method, the conclusions from such a reconstruction cannot necessarily be compared with real observations, as we have pointed out in Sec. 3.1.3.

A much more severe problem with the approach that we have chosen in Sec. 6.3 arises from the fact that we have not simulated the actual lensing process, being the effects of a lens on the photons, emitted from a distant background source, probably seen through the Earth's atmosphere, with a real telescope. A realistic lensing scenario introduces a number of systematic effects, most of which we have already described in Sec. 2.2.3, making it much more difficult to recover the properties of the deflector reliably.

Both issues are addressed in the realistic lensing scenarios that we use in this chapter to test and calibrate our reconstruction algorithm. The state-of-the-art numerical hydro-simulations, from which we extracted the lensing clusters, are described in Sec. 7.1. A comprehensive software package named `SkyLens`, simulating extremely realistic lensing situations, is described Sec. 7.2. We use `SkyLens` to compare different lensing-reconstruction methods, each of which is presented in Sec. 7.3. In addition, an X-ray analysis that was performed on the simulations is briefly mentioned. The results of the different analysis are compared and evaluated in Sec. 7.4.

7.1 The cluster simulations

For our analysis we use three different, numerically simulated clusters, each in three projections along different lines-of-sight. The clusters are called `g1`, `g51` and `g72` and we denote the projection along the different coordinate-frame axes with the addition -z, -y or -x to the cluster's name, respectively. We obtain nine cluster projections denoted as `g1-z`, `g1-y`, `g1-x`, `g51-z`, `g51-y`, `g51-x`, `g72-z`, `g72-y` and `g72-x`.

These clusters are the objects of interest in our subsequent analysis and they have already been used in several other studies (Dolag et al., 2005; Puchwein et al., 2005; Meneghetti et al., 2007a, 2008; Rasia et al., 2006, 2008). A detailed description of the underlying hydrodynamical simulations is provided in Saro et al. (2006). In the following, we shall just give an overview on the most important aspects of these simulations.

7.1.1 Initial conditions

For the simulation of our clusters, the technique of re-simulation was used. Starting from a cosmological parent simulation with a box size of $479\ h^{-1}$ Mpc and containing only dark matter particles, areas in which clusters formed, were re-simulated with a much higher mass

resolution and with an additional baryonic component. The parent simulation is described in Yoshida et al. (2001) and assumed a flat ΛCDM cosmological model with the following parameters: $\Omega_\mathrm{m} = 0.3$, $h = 0.7$, $\sigma_8 = 0.9$ and $\Omega_\mathrm{b} = 0.04$. The initial conditions for the re-simulations were created with the Zoomed Initial Condition technique (Tormen et al., 1997). The masses of the dark-matter and gas particles were set to $m_\mathrm{DM} = 1.13 \times 10^9\ h^{-1} M_\odot$ and $m_\mathrm{b} = 1.7 \times 10^8\ h^{-1} M_\odot$, respectively. The number of gas particles was derived from the cosmic baryon fraction and the initial particle positions were displaced according to the Zel'dovich approximation (see Sec. 3.1.3).

7.1.2 Physics in the simulations

The re-simulations of Saro et al. (2006) were performed with the tree-SPH code `Gadget-2` (Springel et al., 2005) and contained a number of additional physical processes. Apart from the standard implementations in `Gadget-2`, including tree N-Body summation, artificial viscosity treatment, (Dolag et al., 2005) radiative cooling and galactic winds (Springel and Hernquist, 2003), a more elaborate treatment of supernova feedback and chemical enrichment was implemented. Not only the energy feedback and ICM enrichment due to supernovae of type SNII, but also by SNIa was considered and the according sub-grid models assumed an initial stellar mass function following the logarithmic slope of Salpeter (1955). The simulations assumed a galactic wind speed, driven by the supernova feedback, of $v_\mathrm{w} = 500$ km s^{-1}.

7.1.3 Description of the clusters

Those clusters of the simulations, that we used in our analysis, have quite different properties, which raises the interesting question how the different morphologies will be reproduced in the following reconstructions. Tab. 7.1 shows the main properties of the three clusters, where also the parameters of the respective NFW fits are shown. A common reparametrisation of the scale radius in the NFW profile is given by the so-called concentration parameter c, which computes the ratio between some given radius and the scale radius $c(r) = r/r_\mathrm{s}$. A typical radius, used to describe clusters, is r_{200}, which is the radius from the cluster centre within which the average cluster density is more than 200 times higher then the critical density of the Universe. The according concentration parameter is given by $c_{200} = r_{200}/r_\mathrm{s}$.

The cluster `g1` is the most massive cluster in our analysis and it also shows the most regular shape. The density maps of the three projections can be seen in Figs. 7.7 - 7.15. `g51` is less massive and shows a slightly larger amount of substructure. The most interesting case is `g72`, who has a massive companion of $5 \times 10^{13}\ h^{-1} M_\odot$ relatively close to the cluster centre ($3\ h^{-1}$ Mpc). This companion is less visible in the `g72-z` projection in Fig. 7.13, compared to the other two projections in Figs. 7.14, 7.15. The mass profiles of all three cluster can be found in Fig. 7.1.

7.2 The `SkyLens` lensing simulator

Now that we are in the possession of adequate numerical simulations of galaxy clusters, we have to make sure that also the lensing process is simulated accurately. For this purpose we make use of a ray-tracing code named `SkyLens`, which was first introduced in Meneghetti et al. (2008) and has been continuously developed. `SkyLens` simulates the propagation of light rays between different layers, identical to the scenario shown in Fig. 2.1. In this work, we will just sketch the content of the different layers, the actual ray-tracing algorithm, which

	z	r_{200}	M_{200}	b/a	c/a	θ_x	θ_y	θ_z	c_{200}	r_s
g1	0.297	1.54	1.14×10^{15}	0.64	0.57	33.3	57.4	96.1	4.62	0.310
g51	0.2335	1.39	7.85×10^{14}	0.78	0.65	81.5	75.59	16.8	5.37	0.241
g72	0.297	1.30	6.83×10^{14}	0.31	0.29	98.9	92.8	9.4	3.99	0.299

Table 7.1: Main properties of the simulated clusters. Column 1: cluster name; Column 2: redshift; Column 3: r_{200} [h^{-1} Mpc]; Column 4: M_{200} [$h^{-1}M_{\odot}$]; Columns 5-6: principal axes ratios: b/a, c/a, where $a > b > c$; Columns 7-9: angles between the main principal axis and the $x-$, $y-$, and $z-$axes of the simulation box [deg]; Column 10: best-fit 3D-concentration; Column 11: best-fit 3D-scale radius

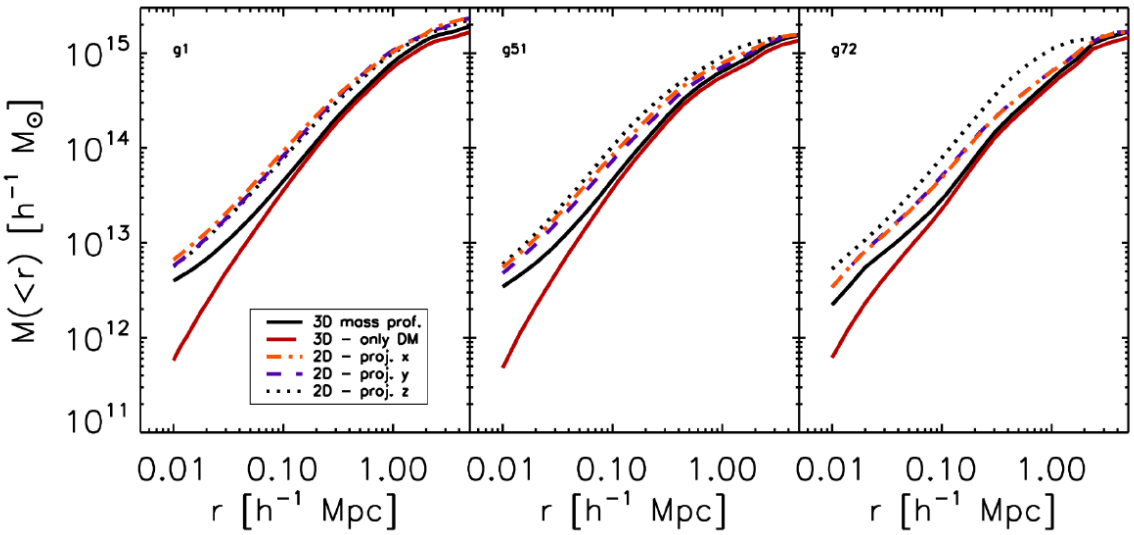


Figure 7.1: Mass profiles of the clusters g1, g51 and g72. The solid black and red lines indicate the total and dark matter only 3D-mass profiles, respectively. The total 2D-mass profiles corresponding to the x, y, and z projections of each cluster are given by the dotted, dashed, and dash-dotted lines.

follows the pathway of photons and evaluates transformations in position and intensity of light rays through the different layers, is presented in more detail in Meneghetti et al. (2008) and Melchior (2010).

The most distant layer consists of a population of realistic background galaxies, as we will describe in Sec. 7.2.1. The second layer is the deflector itself, bending the light rays according to Eq. 2.7 and we will explain in Sec. 7.2.2 how to derive the necessary deflection-angle field from the numerical simulations of Sec. 7.1. The last and presumably most interesting layer, at least while developing a lensing simulator, is the plane of the observer, involving the characteristics of the telescope and its detector, in addition to seeing and background light in the sky. The influence of these effects on the actual image, from which we try to derive the properties of the deflector, are explained in Sec. 7.2.3.

7.2.1 The background population

To ensure a realistic simulation, `SkyLens` offers two different possibilities to populate the source layer with galaxies. The analytic Sérsic profile provides a well-fitting description of the surface-brightness distribution of observed galaxies. This profile, together with a realistic spectral energy distribution (SED) can be used to create synthetic galaxies.

To provide a selection of galaxies based on real observations, `SkyLens` is connected to a large database, containing different information of the galaxies in the `GOODS`[1] (Giavalisco et al., 2004) and the `HUDF`[2] (Beckwith et al., 2006) surveys. The original images of these galaxies are decomposed into `shapelets` (see Sec. 2.2.3) to reproduce their morphology within the simulation and their SED's are derived from photometric and spectroscopic analysis, available for the surveys (Coe et al., 2006; Grazian et al., 2006; Vanzella et al., 2008; Popesso et al., 2009). In total, this provides `SkyLens` $\sim$ 10000 galaxies, which can be varied and reused by rotations and parity flips, to guarantee realistic sources in the lensing simulation, observable in different wave bands. An example of a `shapelet` decomposition of an `HUDF` galaxy in different bands is shown in Fig. 7.2. For the future, it is planned to extend this database by incorporating galaxies from the much wider surveys like `GEMS`[3] (Rix et al., 2004) or from the `COSMOS` field (Koekemoer et al., 2007). Another way to obtain a larger number of realistic galaxies is provided by slightly varying the `shapelet` coefficients of existing galaxy models, as shown by Massey et al. (2004). Ideas to perform this variation in a meaningful manner, involve the help of existing `shapelet`-based galaxy-morphology classifiers as e.g. presented in Andrae et al. (2010).

7.2.2 The deflector

The deflector in the ray-tracing process of `SkyLens` can be given by an analytic profile (e.g. SIS, NFW or Einasto), but this does not coincide with our intention to create realistic lensing scenarios. More complicated is our case, where we want to use a numerical simulation of a cluster as gravitational lens. The snapshot of a numerical simulation at a given redshift provides particle positions and velocities within the simulated box. From this data, we have to derive deflection-angle maps for each line-of-sight projection of the simulated clusters. We perform this task in the following way:

[1] `Great Observatories Origins Deep Survey`
[2] `Hubble Ultra Deep Field`
[3] `Galaxy Evolution from Morphologies and SEDs`

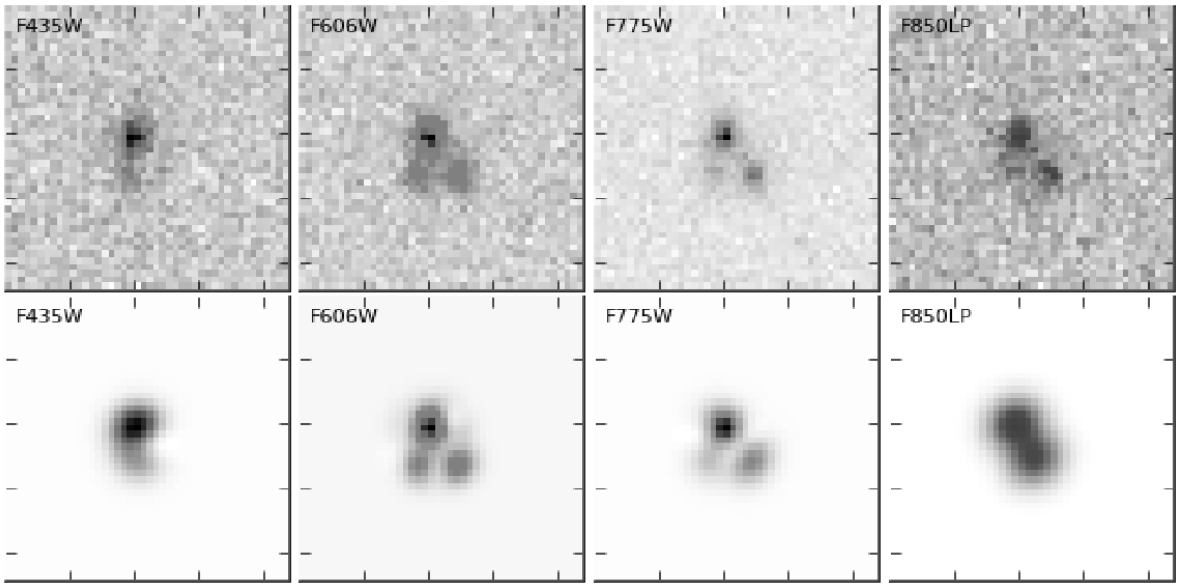


Figure 7.2: From Melchior (2010). Multi-colour images (*top row*) and `shapelet` models (*bottom row*) of a galaxy in the HUDF. These images are taken in the ACS filters F435W, F606W, F775W, and F850LP (from left to right).

The particles projected on each lens plane are used to calculate the deflection angles of light ray bundles. The light rays are traced from the observer position towards the background sources through two regular grids with different spatial resolutions. The inner $1.5 \times 1.5\,h^{-2}$ Mpc^2 region around the cluster centre is sampled with 2048×2048 light rays. This guarantees sufficient spatial resolution for reproducing accurately the positions of multiple images in the strong-lensing regime (Meneghetti et al., 2007a). For the weak-lensing regime, we need to sample a much wider area, while the spatial resolution is less important. Thus, we cover the whole lens plane with a grid of 4096×4096 light rays. The deflection angles are computed using a tree-based code, which works as follows. First, it ranks the particles based on their distances from the light ray positions, building a Barnes & Hut tree in 2D (Barnes and Hut, 1986). The contributions to the deflection angles from nearby and distant particles are calculated separately using direct summation or higher-order Taylor expansions of the deflection potential around the light ray positions. Precisely, given a light ray at position $\boldsymbol{R}$ in physical units, which corresponds to an angular position $\boldsymbol{\theta} = \boldsymbol{R}/D_\mathrm{l}$, the contribution to its deflection angle by a system of mass elements, m_a at positions $\boldsymbol{R}_a$ ($a = 1, 2, ..., N-1, N$), with centre of mass $\boldsymbol{R}_\mathrm{CM}$, and with $|\boldsymbol{R} - \boldsymbol{R}_\mathrm{CM}| \ll |\boldsymbol{R}_a - \boldsymbol{R}_\mathrm{CM}|$ for all the mass elements a, is

$$\begin{aligned} \alpha_i(\boldsymbol{R}) &= \frac{4GM}{c^2}\left[F_1(R')\delta^{ij} + F_2(R')Q^{ij} + \frac{1}{2}F_3(R')(R'^k Q_{kn} R'^n)\delta^{i,j}\right. \\ &+ \left.\frac{1}{2}F_4(R')P^{ij}\right] R'_j \end{aligned} \tag{7.1}$$

where M is the total mass of the system, $\boldsymbol{R}' = \boldsymbol{R} - \boldsymbol{R}_\mathrm{CM}$, δ^{ij} is the Kroneker function, and

the tensors P and Q are defined as

$$Q_{ij} = \frac{1}{M}\sum_{a=1}^{a=N} m_a R_i'^a R_j'^a, \tag{7.2}$$

$$P_{ij} = \frac{1}{M}\sum_{a=1}^{a=N} m_a |R'^a|^2 \delta_{ij}. \tag{7.3}$$

Assuming a Plummer softening to avoid the deflection angles diverging, the $F_k(R')$ functions are defined as

$$F_1(R') = \frac{1}{(R'^2+s^2)} \tag{7.4}$$

$$F_2(R') = \frac{-2}{(R'^2+s^2)^2} \tag{7.5}$$

$$F_3(R') = \frac{8}{(R'^2+s^2)^3} \tag{7.6}$$

$$F_4(R') = \frac{-2}{(R'^2+s^2)^2}\ . \tag{7.7}$$

Nearby particles are treated as point lenses and Eq. 7.1 reduces to

$$\alpha_i(\boldsymbol{R}) = \frac{4GM}{c^2} R_i' F_1(R')\ . \tag{7.8}$$

The fraction of particles that are evaluated with Eqs. 7.1 or 7.8 is set by the Barnes-Hut opening criterion, θ_{BH} (see e.g Springel, 2005), which we fix at $\theta_{\mathrm{BH}} = 0.4$. As shown by Aubert et al. (2007), the optimal softening length s depends on the resolution of the simulation. We performed several tests to determine which values to use. Doing ray-tracing through NFW halos, sampled with a similar number of particles as our simulated clusters, we verified that a softening scale of $5\,h^{-1}$ kpc is appropriate for reliably reproducing the deflection angle field of the input models over the range of scales relevant for both strong and weak lensing.

7.2.3 Observational effects

The main motivation for the development of `SkyLens` was to produce realistic, but simulated astronomical CCD images in several wave bands. To those images established lensing-analysis pipelines can be directly applied. So far, we described the extraterrestrial ingredients for such a simulation, but not the observational process. In order to do so, we start at the very end of this process, being the CCD camera of a telescope. Following the derivation in Grazian et al. (2004), we can calculate the actual CCD measurement value (ADU) for one specific pixel with index i as

$$\mathrm{ADU}_i = \frac{1}{g}\int_{A_{\mathrm{pix}}} d^2x\ n_\gamma(\boldsymbol{x}), \tag{7.9}$$

where $n_\gamma(\boldsymbol{x})$ is the number of photons passing through the observational plane at position $\boldsymbol{x}$, A_{pix} the area of one pixel and g is the detector gain. The last two numbers define the first parameters that we have to fix in order to simulate the observation.

The number of photons, reaching the detector, depends mainly on the SEDs of our sources, given by the galaxy models of Sec. 7.2.1, and several other parameters in

$$n_\gamma(\boldsymbol{x})\ d^2x = \frac{\pi D^2 t_{\mathrm{exp}}}{4h}\int d\lambda \frac{T(\lambda)\mathrm{SED}(\lambda,\boldsymbol{x}_{\mathrm{s}})}{\lambda}\ d^2x_{\mathrm{s}}, \tag{7.10}$$

where $\boldsymbol{x}_s$ denotes coordinates in the source plane, D is the telescope's aperture diameter, t_{exp} the exposure time, h Planck's constant and $T(\lambda)$ the transmissivity of the telescope for a certain wavelength λ. We have neglected extinction in our considerations, but an according term is easily added to the equation above.

The total transmissivity of the telescope is given by

$$T(\lambda) = 10^{-0.4 m_a A(\lambda)} M(\lambda) O(\lambda) F(\lambda) C(\lambda). \quad (7.11)$$

In this equation, m_a is the optical path length through the atmosphere called airmass and depends on the zenith angle of the observation. $A(\lambda)$ is the atmosphere's extinction at the telescope site and M, O, F and C are optical quantities, given by the telescope design, describing the transmissivity of the mirrors, the optics, the filter and the CCD, respectively.

One should not forget the photons coming from unwanted sources in the sky like scattered light from the ground or from the moon. Ideally, a full spectrum of the sky is taken before an observation and the derived SED_{sky} is directly added to Eq. 7.10, but the following approximation from Grazian et al. (2004), using only the measured sky magnitude M_{sky} (in the AB system) of some empty patch in the sky, may also be useful

$$ADU_{sky} = \frac{\pi D^2 A_{pix}^2 t_{exp}}{4\pi g} 10^{-0.4(M_{sky}+48.6)} \int d\lambda \frac{T(\lambda)}{\lambda}.$$

Noise shall be added to the ADU counts, considering the read-out noise of the CCD, hereafter RON and an imperfect flat-fielding, described by the flat field accuracy a and the residual-flat field error f . The total variance in one pixel with index i is then given by (Grazian et al., 2004)

Name	Description
D	aperture diameter
g	detector gain
A_{pix}	pixel area
$F(\lambda)$	used filter
$M(\lambda)$	mirror filter curve
$O(\lambda)$	optics filter curve
$C(\lambda)$	CCD filter curve
FoV	total field-of-view
RON	detector readout-noise
f	flat-field accuracy
a	residual flat-field error
PSF	PSF model
t_{exp}	exposure time
$A(\lambda)$	atmospheric extinction
m_a	airmass
SED_{sky}	sky-background emission
SED_{gal}	background population
α	deflection angle map

Table 7.2: SkyLens parameters. The first group defines the telescope characteristics, the second group environmental conditions and the last group the extraterrestrial components.

$$\sigma_i^2 = \frac{ADU_i + ADU_{sky}}{g} + n\left(\frac{RON}{g}\right)^2 + \left(f + \frac{a^2}{n^2}\right)(ADU_i + ADU_{sky})^2, \quad (7.12)$$

where n denotes the number of exposures. Additionally, SkyLens is also able to deal with shifts in the coordinate postions, due to dithering patterns of the observations and to mask areas of the CCD, due to gaps in the detector and CCD boundaries.

The last thing, that we have to take into account for a realistic observation, is the PSF . We already explained this point in Sec. 2.2.3 and our simulator is able to convolve the lensed images of the background sources with any given PSF model. In order to obtain the PSF model from the observation in a realistic way, stars can be added to the simulated observation.

All the parameters, defining a simulated observation with SkyLens are summarised in Tab. 7.2 and a simulated multi band observation, resulting in realistic CCD images is shown in Fig. 7.3.

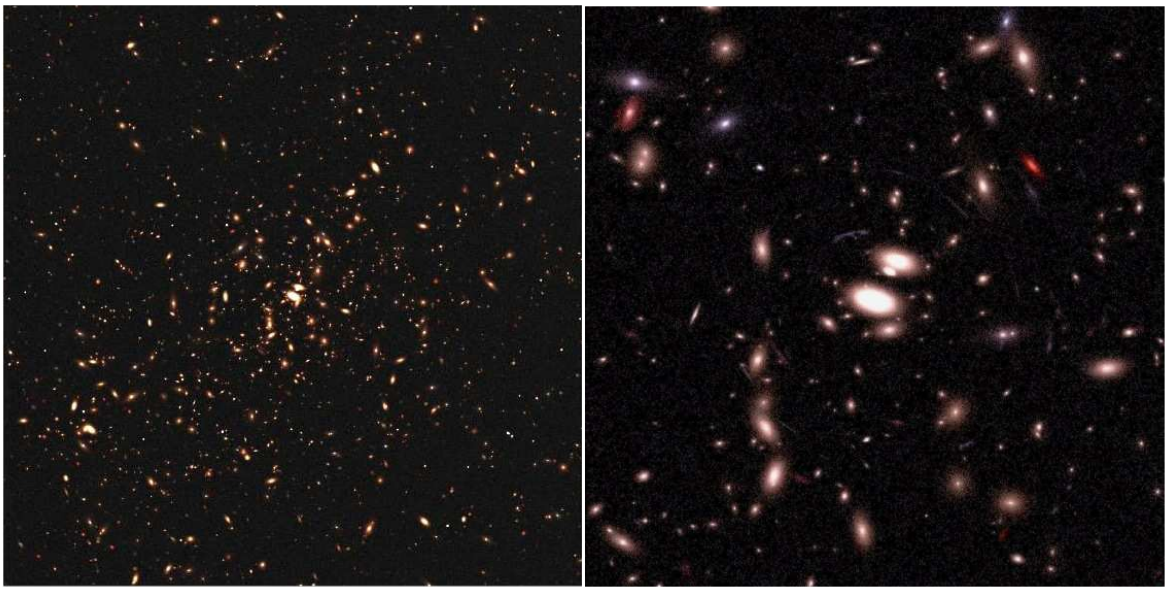


Figure 7.3: The *left panel* shows a colour-composite image of a simulated galaxy cluster, obtained by combining three SUBARU exposures of 2500s each in the B, V, I bands. The field-of-view corresponds to $\sim 450'' \times 450''$. The *right panel* shows an HST/ACS composite image of the central $\sim 100'' \times 100''$ area in the same cluster. The image has been produced by combining mock observations with the filters F475W, F555W and F775W.

7.3 Different analysis techniques

We tried to find out, how well different reconstruction techniques perform on the realistic lensing scenarios, especially with respect to their ability to recover the mass profiles of the different cluster projections. We applied strong and weak-lensing reconstruction techniques and simulated two types of CCD images for each cluster projection. The first type is adapted to the need of a strong-lensing analysis and mimicked the characteristics of the Advanced camera for surveys (ACS) on board the HST. Sky background and detector noise were added, the exposure time was set to 7500 sec and the field-of-view for the F775W filter was limited to $120'' \times 120''$. This is smaller than the original ACS field-of-view, but sufficient to cover the full Einstein radii of the lenses.

To simulate a typical instrument used for a weak-lensing analysis, we chose the Suprime-CAM mounted on the SUBARU telescope. An exposure of 6000 sec in the I band was convolved with an isotropic, Gaussian PSF, reflecting a seeing of $0\farcs6$. The fields-of-view were chosen such, that significant substructure in all projections was included in the simulation. This led to slightly larger or smaller fields-of-view than in the case of a real Subaru exposure.

7.3.1 Strong-lensing analysis

The strong-lensing analysis is performed with LensTool, as described in Sec. 4.4. We used an NFW profile to model the main cluster halo. Moreover, we added several subcomponents, representing the contribution from the most massive galaxies in the cluster, since, as shown in some previous studies, it is important to include the cluster members in the model because they can affect the positions and the magnifications of the strong lensing features

(Meneghetti et al., 2003a). These are considered using pseudo-isothermal-elliptical-mass-distributions (PIEMD) , described by the following density profile

$$\rho_{\rm PIEMD}(r) = \frac{\rho_0}{(1+r^2/r^2_{\rm core})(1+r^2/r^2_{\rm cut})} \ . \tag{7.13}$$

We chose the PIEMD model because this is widely used for modelling the lensing properties of cluster galaxies in observations (see e.g. Limousin et al., 2007; Riemer-Sørensen et al., 2009; Donnarumma et al., 2009, for some recent references).

Observationally, the galaxies to be included in the model should be selected as those lying in the cluster red sequence and being brighter than a given apparent luminosity (e.g. Limousin et al., 2007). Of course what matters for lensing is not the luminosity but the mass, which is assumed to be traced by the light. Indeed, the minimal luminosity should be interpreted as a minimal mass. Working with simulations, we identify the cluster galaxies using the `SUBFIND` code (Springel et al., 2001) and then apply a selection based directly on the stellar mass. `SUBFIND` decomposes the cluster halo into a set of disjoint substructures and then identifies each of them as a locally overdense region in the density field of the background halo. In our reconstructions, we include those galaxies that have stellar mass $M_{\rm stars} \geq 10^9\,h^{-1}\,M_\odot$ and that are contained in a region of $500\ h^{-1}$ kpc around the cluster centre. This is typically more than three times the size of the Einstein rings of the clusters in our sample. The orientation and the ellipticity of each galaxy are measured from the distribution of the star particles belonging to it. Following this procedure, we typically end up with catalogues of several tenth of cluster members.

The Brightest-Central-Galaxy (BCG) is included in the lens model by optimising its parameters individually, rather than scaling them with the luminosity/mass. Since the BCG forms in the simulations in a strong cooling region, we assumed it might have significantly different properties compared to the other cluster members. Thus, we prefer to treat it individually. Analogously, we use individual optimisation with some other cluster members which lay particularly close to some multiple image systems. Indeed, their influence on the local lensing properties of the cluster requires to be modelled carefully.

The total number of free parameters in the model depends on the complexity of the lens. Usually, we consider a cluster-scale mass component, a galaxy-scale component to describe the BCG, and other galaxy-scale terms to incorporate the relevant cluster members.

We distribute the sources behind the clusters such as to have $\sim 3-7$ strong lensing systems available for the optimization. For this condition to be satisfied, we randomly distribute few sources in a shell surrounding the lens caustics, enhancing the chances that they are strongly lensed. Then, we visually check whether the multiple images belonging to each source are detectable in the simulation and then retain those systems that are useful for the strong lensing analysis. The optimisation is done, using the Bayesian method implemented in `LensTool` with an optimisation rate of $\delta\lambda = 0.1$. We assume the uncertainty in the lensed image positions to be $\sigma_{\rm I} = 0.3''$.

To give a specific example, we describe in the following the strong-lensing reconstruction of `g1-y`.

Among the sources, which were distributed along the caustics of the input cluster lens, seven produced multiple-image systems detectable in this deep exposure (7500s) in the F775W filter. More precisely, two sources produced five images, while the other sources are imaged into triplets. Using these observables, the reconstruction converges, finding a good fit to the lensing features ($\chi^2 = 18$ for 21 degrees of freedom). The best-fit model consists of an NFW halo with concentration $c = 10.57^{+2.82}_{-1.81}$ and scale radius $r_s = 29.07^{+20.68}_{-1.85}$ arcsec (correspond-

ing to $90^{+64}_{-6}\ h^{-1}$ kpc). Two separately modelled galaxies, including the BCG, have velocity dispersions of 340^{+21}_{-30} km s^{-1} and 269^{+12}_{-17} km s^{-1}, respectively.

The inner projected mass profile, derived from the model, $M(< R)$, is shown in Fig. 7.4, where we also show the true profile of the cluster acting as lens in this simulation. Since the model reproduces the lens' tangential critical line well, it is not surprising that the model is very reliable at estimating the mass enclosed in the strong lensing region. The shaded area in the figure indicates the radial range of the multiple images, excluding the central images, which are located at $R \lesssim 10\ h^{-1}$ kpc. The reconstruction reproduces the true mass profile well up to $\sim 150\ h^{-1}$ kpc from the centre, where the deviation from the true mass profile is $\lesssim 10\%$. At larger radii, the differences become significant. Thus, extrapolating the strong lensing model to distances where no strong lensing features are observed may result in very incorrect mass estimates. This issue is discussed in more detail in Sec. 7.4.1.

To evaluate how the reliability of the model degrades by reducing the number of constraints, we performed another reconstruction using only one system with five images and two triplets. The final reconstruction did not differ significantly from the previous one. The projected mass profile for this new lens model is also given in Fig. 7.4. This result shows that reliable reconstructions can be achieved even with a limited number of lensing constraints, if they are widely spread across the cluster. This finding coincides with the results of Merten et al. (2009) and supports the approach of incorporating the strong-lensing constraints in the combined lensing analysis of Sec. 7.3.3. We also attempted a reconstruction by neglecting the central images (and using all the seven lensed systems). This is likely to be a realistic situation, since the central images are generally demagnified and hidden behind the BCG, hence difficult to detect. In this case, the mass enclosed by the strong lensing region is again correctly estimated, but the reconstructed profile deviates more from the true one at small radii.

7.3.2 Weak-lensing analysis

The weak-lensing ellipticity measurements were obtained, using the `KSB+` method presented in Sec. 2.2.3. This method is internally implemented in `SkyLens`[4]. Even though our lensing scenarios are very realistic, we employ some simplification that should be noted.

By selecting the galaxies with $S/N > 10$, we end up with catalogues of galaxy ellipticities with a source density of ~ 30 arcmin^{-2}. The median redshift of these sources is $z_{\mathrm{s,true}} \sim 1.05$. In the following analysis we assume that all sources have the same redshift of $z_{\mathrm{s}} = 1$. Furthermore, we assume that we can separate the population of background galaxies perfectly from the foreground cluster members. This is intentionally very optimistic, since we aim at verifying the capabilities of several lensing methods to retrieve the cluster mass in the best possible conditions. The misidentification of cluster members as background galaxies leads to a dilution of the lensing signal, which leads to erroneous mass estimates (see e.g. Medezinski et al., 2007, 2009). When increasing the distance from the cluster centre, the probability that nearby substructures or additional mass clumps affect the mass estimates becomes higher. In this work, we have not included the effects of uncorrelated large-scale-structures (LSS) on the weak-lensing signal. The effects of the LSS on the weak lensing mass estimates have been discussed in detail in several other works (Cen, 1997; Reblinsky and Bartelmann, 1999; Metzler et al., 1999; Hoekstra, 2001, 2003; White and Vale, 2004; Clowe et al., 2004a).

[4] Following the approach of Hoekstra et al. (1998)

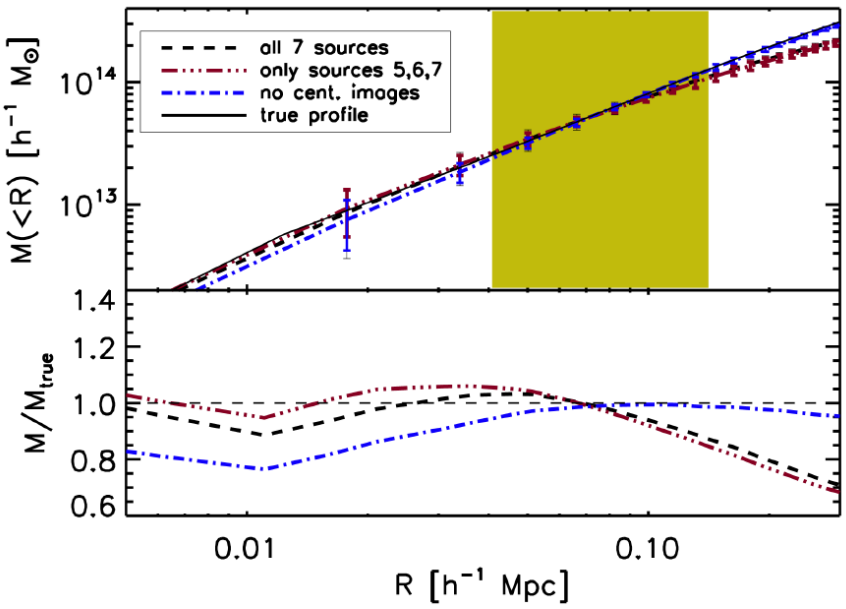


Figure 7.4: The total projected mass profile of the inner region of cluster `g1-y`, as recovered from the strong-lensing mass reconstruction using `LensTool`. The dashed line shows the result obtained by using seven multiple-image systems. The red three-dot-dashed line shows the mass profile if only three multiple-image systems are used. The blue dot-dashed line indicates the mass profile recovered by fitting all the 7 multiple-image systems, but assuming that all the central images are not detectable. Finally, the true mass profile, as drawn from the particle distribution in the input cluster, is given by the solid line. The shaded region shows the radial range of the tangential strong lensing constraints. The *bottom panel* shows the ratios between the recovered mass profiles and the true mass profile.

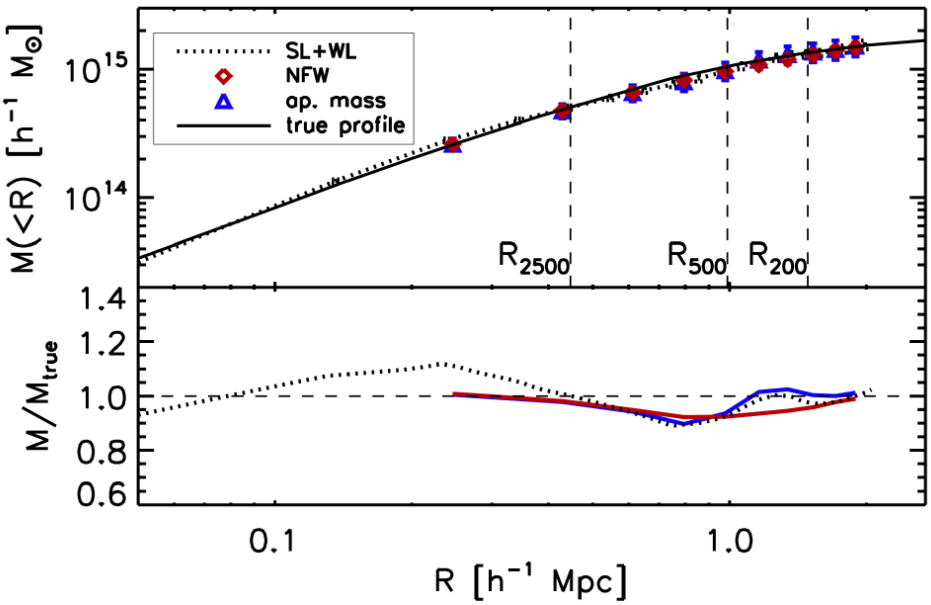


Figure 7.5: Radial 2D-mass profiles of the cluster `g1-y`, as obtained from three different methods, namely the NFW fit to the shear profile (diamonds), the aperture-mass densitometry (triangles) (Fahlman et al., 1994; Clowe et al., 1998; Bartelmann and Schneider, 2001), and the 2D mass reconstruction combining weak and strong lensing (dotted line). The solid line shows the true mass profile. The vertical lines indicate the positions of R_{2500}, R_{500}, and R_{2500} as derived from the NFW fit. The *bottom panel* shows the ratios between the mass profiles recovered from the lensing analysis and the true mass profile. The dotted line refers again to the joint method, while the red and blue solid lines indicate the results for the NFW fit and for the aperture mass, respectively.

The cluster masses are derived with the following approach:
Assuming that the cluster is described well by an NFW density profile, we use the corresponding formula for the reduced shear to fit the azimuthally averaged profile of the tangential component of the reduced shear. For the NFW profile, the formulas for the radial profiles of the shear and of the convergence can be found in Bartelmann (1996) and in Meneghetti et al. (2003b). The tangential component of the reduced shear is given by

$$g_+ = -\Re[g\, e^{-2\mathrm{i}\phi}] \,, \tag{7.14}$$

where the angle ϕ specifies the direction from the galaxy centroid towards the centre of the cluster, which we identify with the position of the BCG. The cross component of the reduced shear is given by

$$g_\times = -\Im[g\, e^{-2\mathrm{i}\phi}] \,. \tag{7.15}$$

If the distortion is caused by lensing, this component of the shear should be zero.

To give an example, the tangential shear component of `g1-y` is well-fitted by an NFW profile with $c = 4.82 \pm 0.64$ and $r_\mathrm{s} = 0.307 \pm 0.048\ h^{-1}$ Mpc, where the errors are provided by the model-selection technique. As expected in the absence of systematics, the cross component of the shear is consistent with zero. The according mass profile of the same cluster, as obtained with weak-lensing methods and compared with the method, that will be presented in the next section is shown in Fig. 7.5.

7.3.3 Combining weak and strong lensing

For the final lensing analysis, we did a joint reconstruction, combining weak and strong lensing, with the method that we have described in the preceding three chapters. We performed the reconstruction on an initial grid of 10×10 pixels and steadily increased the resolution, until we reached a final grid of 60×60 pixels, according to refinement level 0. The innermost core of the cluster was resolved with a finer grid, refering to 120×120[5] pixels on refinement level 0. We used the weak-lensing catalogues, described in Sec. 7.3.2 and set the parameters of the adaptive-averaging procedure such, that at least twelve galaxies were contained in the sample to estimate the reduced shear value within one cell of the reconstruction grid. We used the multiple-image systems, described in Sec. 7.3.1 to estimate the position of the critical line, as it visualised in Fig. 7.6 for one cluster projection. By estimating the geometry of the lens system, the central coordinate of the connecting line between two multiple images is a good prior of the critical line of a lens. We showed in Merten et al. (2009) that such an estimate of the critical curve leads to practically identical reconstructions, compared to results using the full information on the position of the critical line. We decided not to incorporate any flexion constraints, since we have not a reliable pipeline available yet, which guarantees an accurate measurement of the flexion signal. This pipeline is in the process of development and will be used for further analysis, but at the present stage we did not want to compromise the reconstruction with a not completely trustable input constraint, even if the reconstruction method itself is fully capable of incorporating flexion.

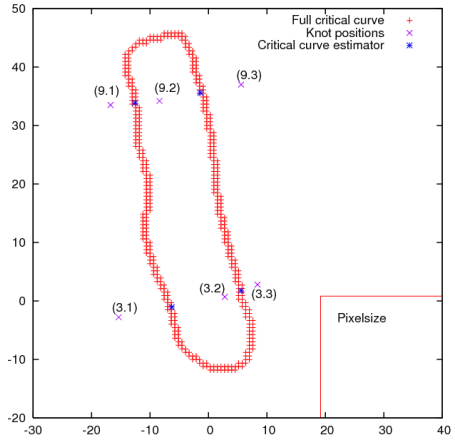


Figure 7.6: This figure illustrates how the critical curve estimators are obtained from given multiple lensed images of the simulation. In the bottom-right corner of the plot we show the pixel size, resolving the cluster core.

As the final result, we obtained the discretised lensing potential on a grid with different refinement levels. From the lensing potential, we derived a convergence map, by applying the Laplacian and the results, compared with the real density maps of the simulations, are shown in Figs. 7.7 - 7.15. Those figures already indicate, that the individual morphologies of the clusters are reproduced very accurately with the nonparametric method and that substructure is recovered in some detail. This feature cannot be achieved with the other two methods, since they rely on a parametrised mass model.

We accounted for the mass-sheet degeneracy by normalising the average convergence at the boundaries of the reconstructed field, defined as the outmost two pixel rows respectively columns), to zero and translated the convergence maps into physical surface-mass density maps by using the average source redshift of the background sources and of the strong-lensing features.

[5] At this stage of the code development only two refinement levels were implemented.

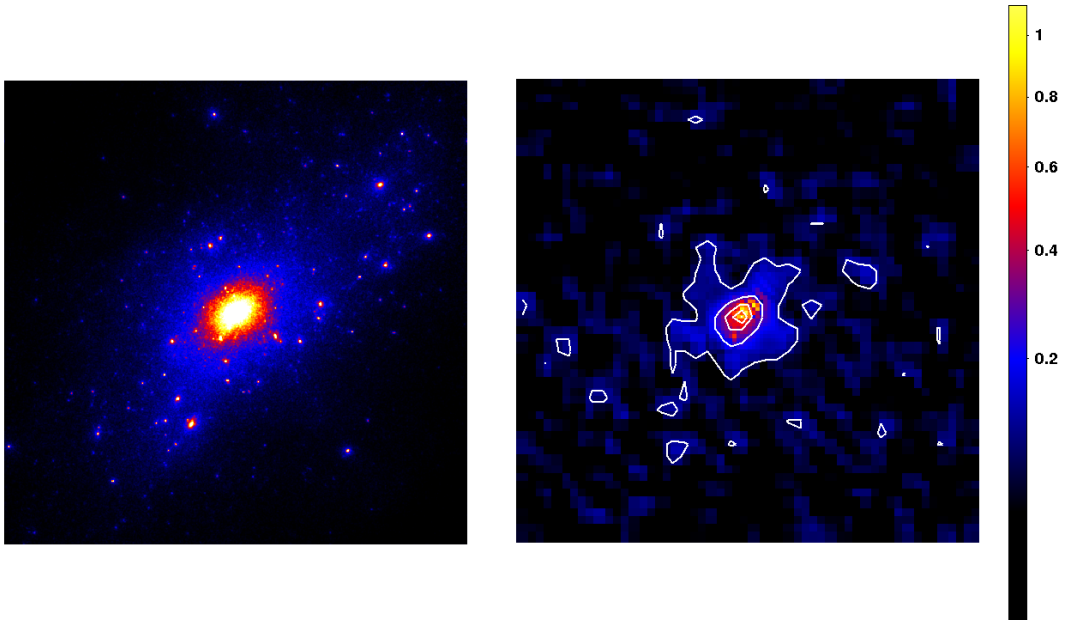


Figure 7.7: The projection along the z-axis of the cluster g1. The side length of the square field is $1280''$, refering to $\sim 4\ h^{-1}$ Mpc at the cluster's redshift. The *left panel* shows the original density map of the simulation. The *right panel* shows the according convergence map as it is obtained by the nonparametric reconstruction method, combining weak and strong lensing. The contours start at $\kappa = 0.09$ with a linear spacing of $\Delta\kappa = 0.17$.

Error bars were attached to the measurements by sampling the weak-lensing catalogues with 24 bootstrap realisations. Since the reconstructions have to be rerun for each bootstrap, we performed the reconstructions on a lower resolutions of 32×32 pixels, refering to refinement level 0. The errors were obtained from the scatter in the resulting convergence maps.

7.4 Comparing the results

In the last section, we introduced three different lensing methods to recover the mass profiles from the simulations and we saw how the different approaches make use of the input data. As an example, the recovered mass profiles for one cluster projection were given in Figs. 7.4, 7.5. What we are now interested in, is how the different methods perform on average, while reconstructing the complete cluster sample.

One should never forget, that lensing probes the projected 2D surface-mass density along the line-of-sight. In order to obtain the full 3D mass-profile, like e.g. X-ray techniques do, one has to make e.g. the assumption of a spherically symmetric lens and deproject the lensing profiles accordingly. To compare the lensing results with realistic X-ray simulations, performed in Meneghetti et al. (2009), we follow this idea. We used the `X-ray MAp Simulator` (`XMAS`) (Gardini et al., 2004; Rasia et al., 2008) to produce for all clusters realistic X-ray images as they would be seen with the `Chandra` X-ray satellite. This provides the basis to compare deprojected lensing and X-ray results on the grounds of realistic simulations .

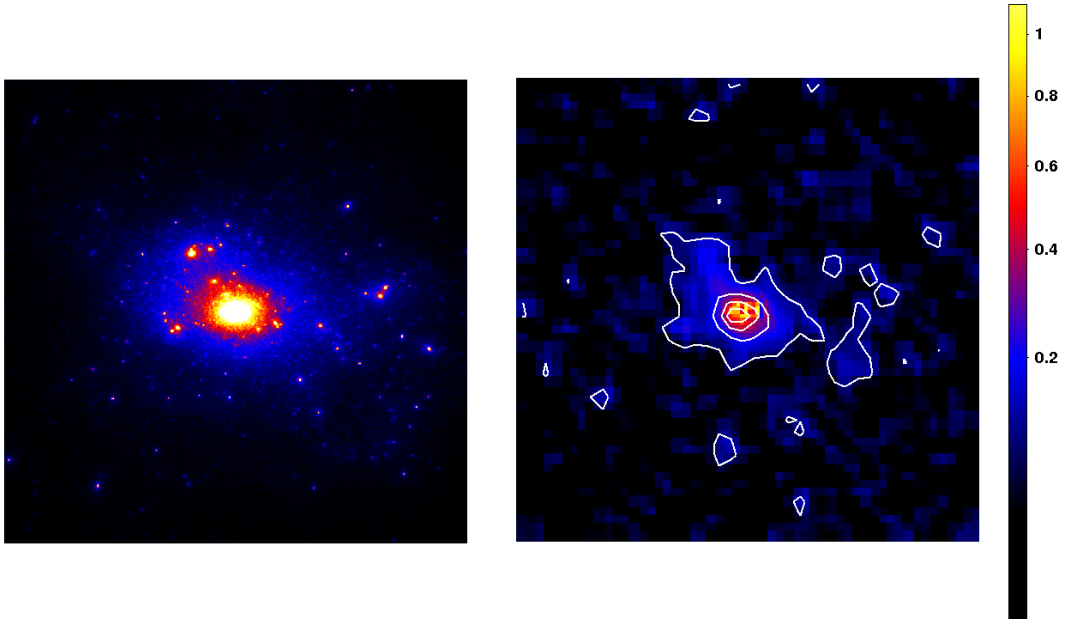


Figure 7.8: As Fig. 7.7 but for the y-axis projection of `g1`.

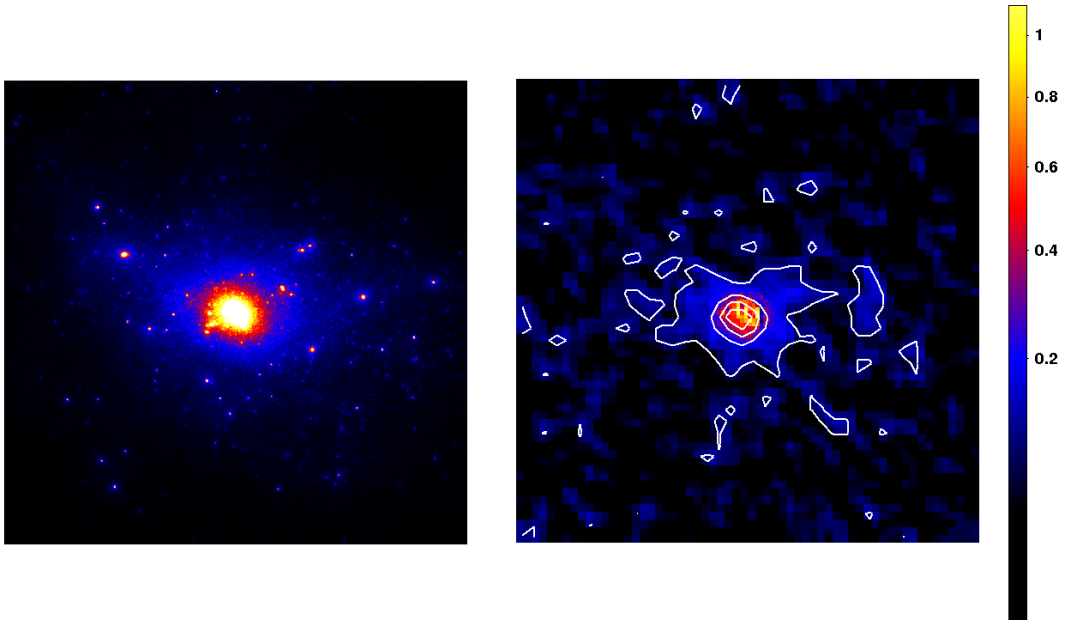


Figure 7.9: As Fig. 7.7 but for the x-axis projection of `g1`.

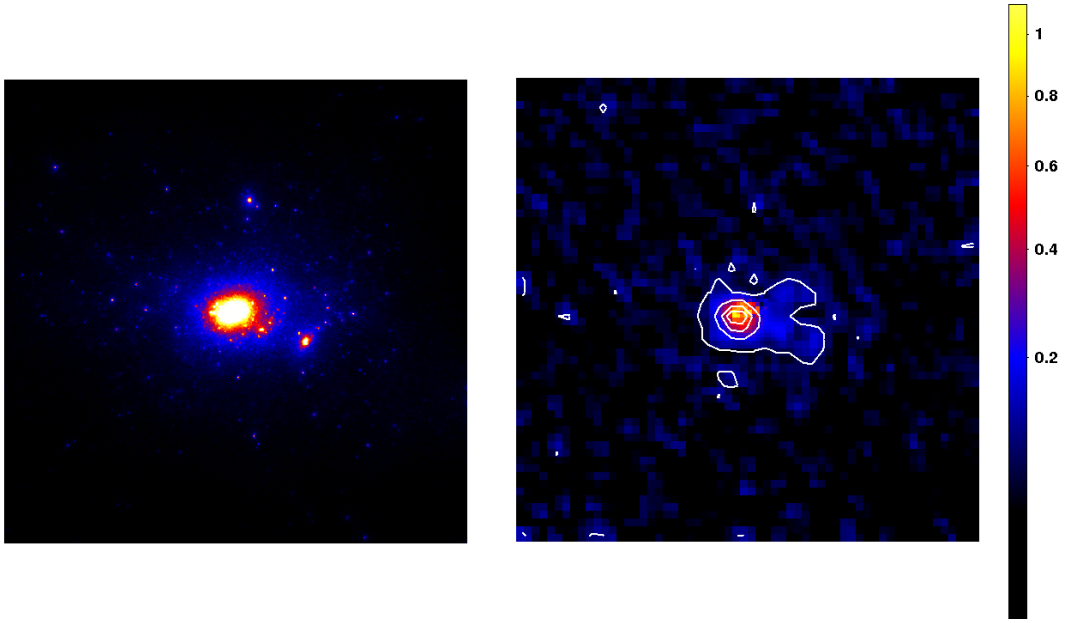


Figure 7.10: The projection along the z-axis of the cluster g51. The side length of the square field is $1900''$, refering to $\sim 4.9\ h^{-1}$ Mpc at the cluster's redshift. The *left panel* shows the original density map of the simulation. The *right panel* shows the according convergence map as it is obtained by the nonparametric reconstruction method, combining weak and strong lensing. The contours start at $\kappa = 0.09$ with a linear spacing of $\Delta\kappa = 0.17$.

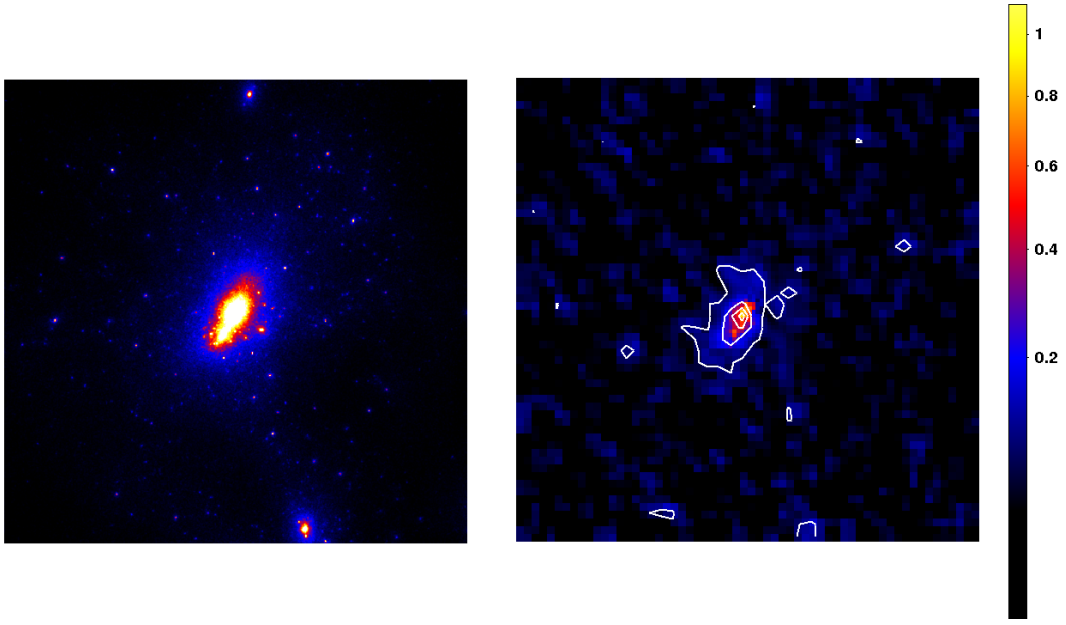


Figure 7.11: As Fig. 7.10 but for the y-axis projection of g51.

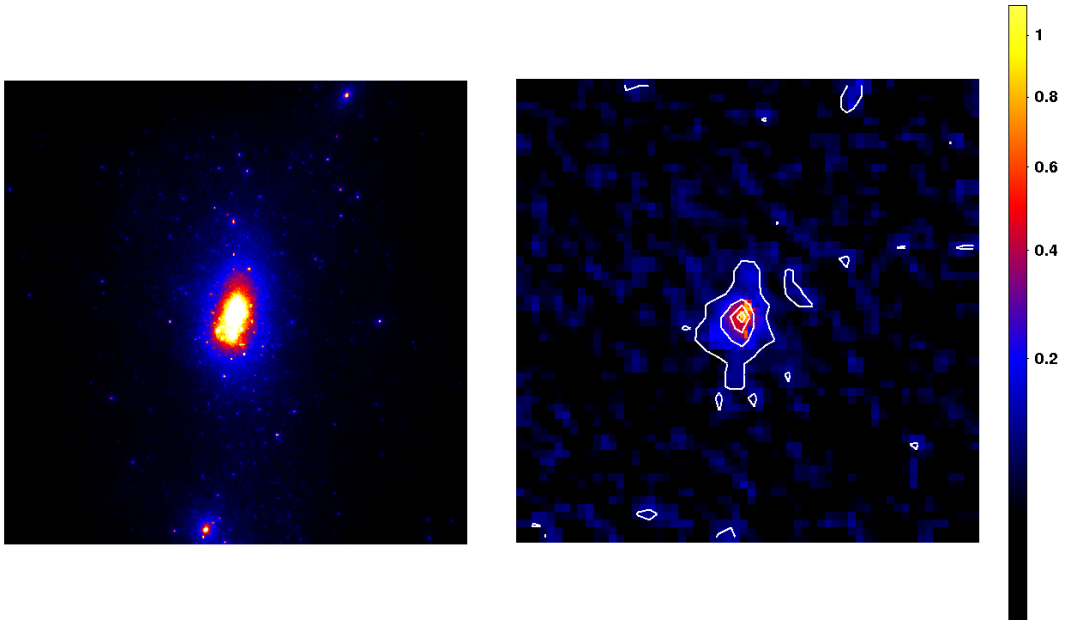


Figure 7.12: As Fig. 7.10 but for the x-axis projection of g51.

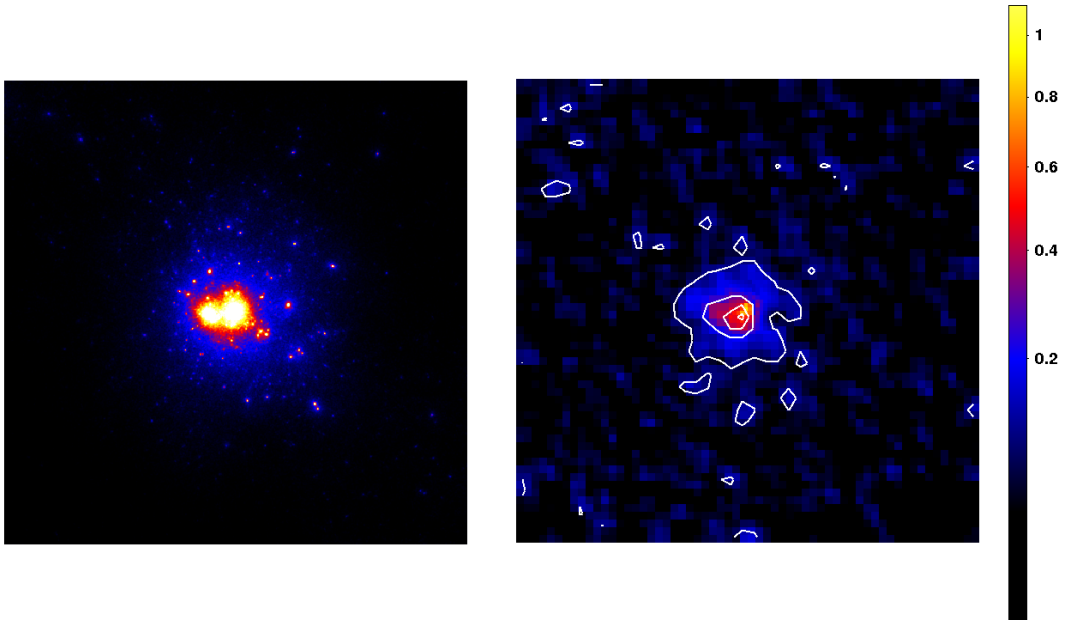


Figure 7.13: The projection along the z-axis of the cluster g72. The side length of the square field is $1520''$, refering to $\sim 4.65\ h^{-1}$ Mpc at the cluster's redshift. The *left panel* shows the original density map of the simulation. The *right panel* shows the according convergence map as it is obtained by the nonparametric reconstruction method, combining weak and strong lensing. The contours start at $\kappa = 0.09$ with a linear spacing of $\Delta\kappa = 0.17$.

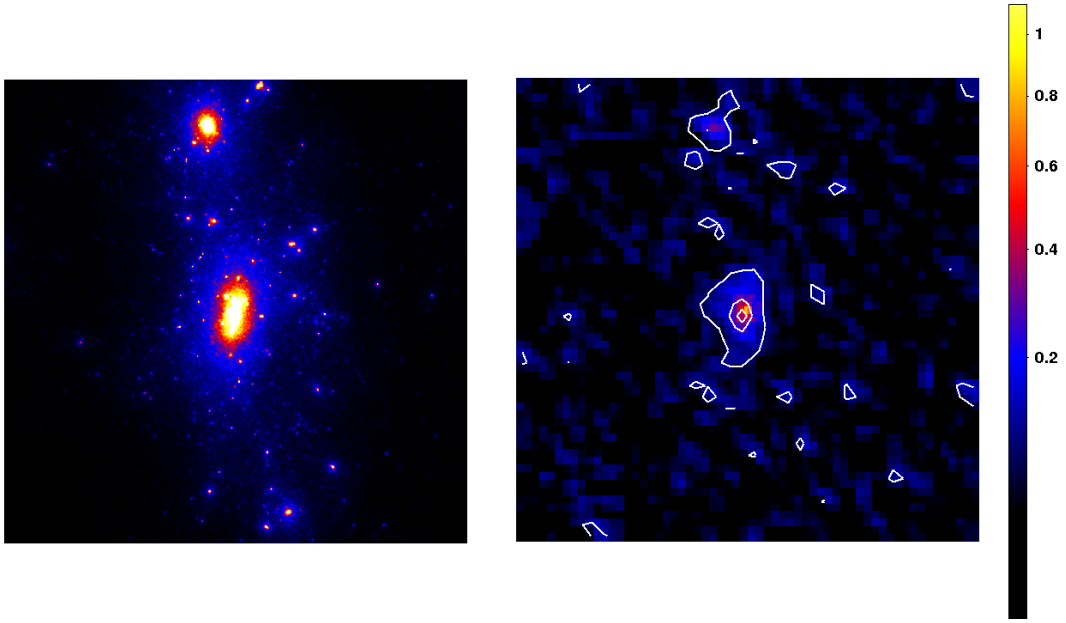


Figure 7.14: As Fig. 7.13 but for the y-axis projection of g72.

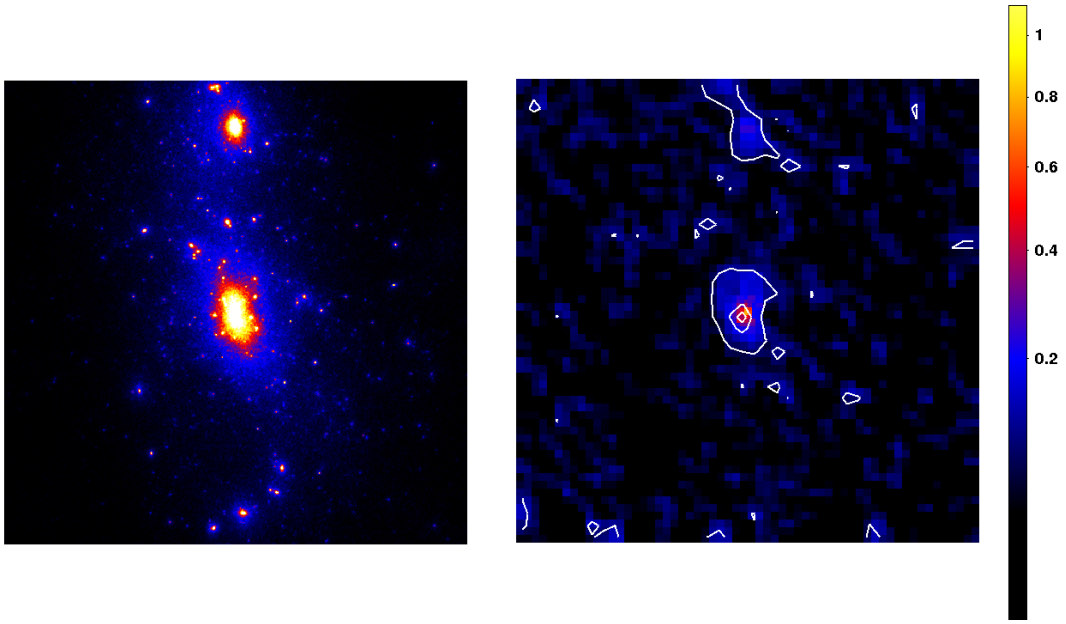


Figure 7.15: As Fig. 7.13 but for the x-axis projection of g72.

7.4.1 Strong lensing: good news and a big problem

In Fig. 7.16 we report the strong-lensing results as obtained by `LensTool` . It can be seen, that the reconstructions perform very well and recover the projected real masses for all clusters with an accuracy of $\sim$ 10% or better.

The big problem with the strong lensing results is the fact, that they are only accurate within a radius around the cluster centre, where the strong-lensing constraints are actually observed. While going to larger distances than the Einstein radius of the cluster, the differences between real, projected mass and the reconstructed mass become more and more prominent, leading to an underestimation of the mass up to 30% even when averaging over the whole cluster sample. This is a clear bias in the strong-lensing reconstructions, as stated in Fig. 7.17.

To investigate the reason for this behaviour, we produced a new set of simulations, which only considered the dark-matter component of the cluster. As can be seen from the flat curve in Fig. 7.17 the bias disappears, indicating that the problem arises from the stellar component of the lens-modelling in `LensTool`. Indeed, it has been shown in previous works on real clusters, like `MS2137` or `Abell611` (Donnarumma et al., 2009, 2010; Comerford et al., 2006), that different models for the cluster's BCG lead to significant differences in the recovered mass profile, especially when extrapolated to radii larger than the distance of the strong-lensing constraints from the cluster centre.

Nevertheless, the good results of the strong-lensing modelling in the innermost region of the clusters are remarkable. The only cases where the cluster core's mass distribution is not modelled accurately are `g51-z` (Fig. 7.10) and `g72-z` (Fig. 7.13). Those two cases are characterised by a complicated double-core structure where, in this particular projection, the companion is not well-separated from the main halo. As a consequence, the multicomponent mass modelling becomes more complicated.

7.4.2 Comparing weak lensing results

In Fig. 7.18 we show the recovered projected mass by the NFW-fit method, where M_x denotes the mass within a radius, within which the average density is x-times the critical density. The errors reflect the uncertainty in the best-fit parameters of the NFW-profile. Three trends are noticeable:

1. The deviation between real and recovered mass increases with distance from the cluster centre, as it is expected and in agreement with Okabe et al. (2009).

2. While considering the whole sample of all nine projections, the NFW-fit method underestimates the mass.

3. The largest errors appear for the projections with asymmetrical profiles and complicated substructure like `g51-y` (Fig. 7.11), `g72-y` (Fig. 7.14) and `g72-x` (Fig. 7.15).

Before we explain those findings, we shall look at the same plot for the nonparametric method in Fig. 7.19. The result is remarkable for several reasons and highlights the performance of our nonparametric approach:

1. The bias has almost disappeared completely[6].

2. Also the masses of the highly substructured lenses are recovered well.

[6] The slight trend to underestimate the mass was already discussed in Sec. 6.3

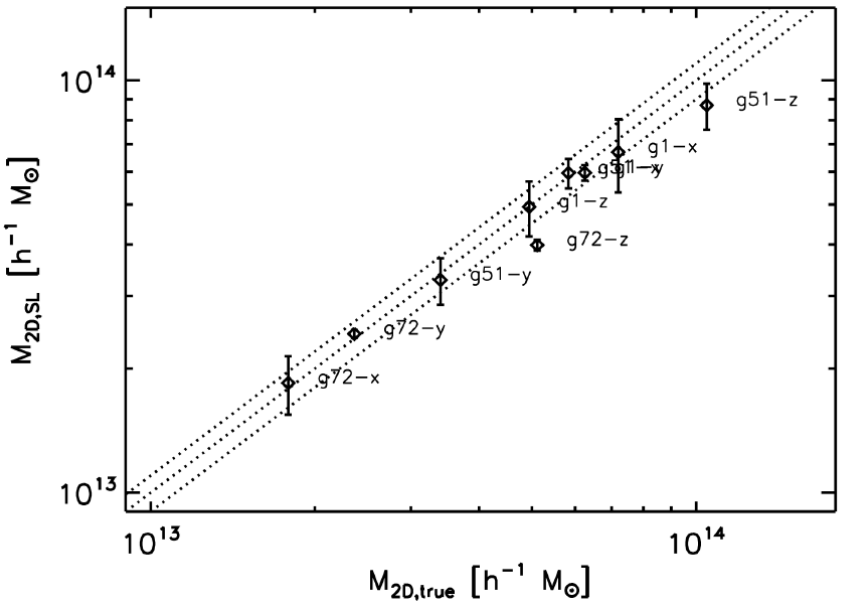


Figure 7.16: The projected masses estimated through the strong-lensing analysis vs. the corresponding true masses of the lenses. The dotted lines correspond to $M_{\rm 2D,SL} = M_{\rm true}$ and to $M_{\rm 2D,SL} = M_{\rm true} \pm 10\%$. The masses are measured within a circle centred on the BCG and having a radius equal to the mean distance of the lensing constraints from the cluster centre.

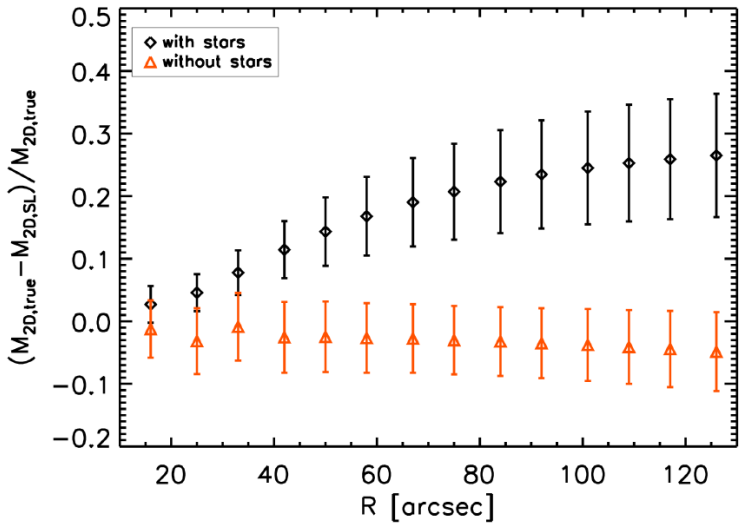


Figure 7.17: The relative difference between true and estimated projected masses as a function of the distance from the cluster centres. The results are obtained by averaging over all the clusters in the sample. The diamonds and the triangles refer to the simulations including and excluding the contribution of the BCG to the lensing signal (see text for more details). The error bars show the scatter among all the reconstructions. The regions probed by strong lensing are typically smaller than $\sim$ 30–40 arcsec, thus the mass estimates at larger radii are extrapolations of the strong-lensing mass model to distances that are unconstrained by the data.

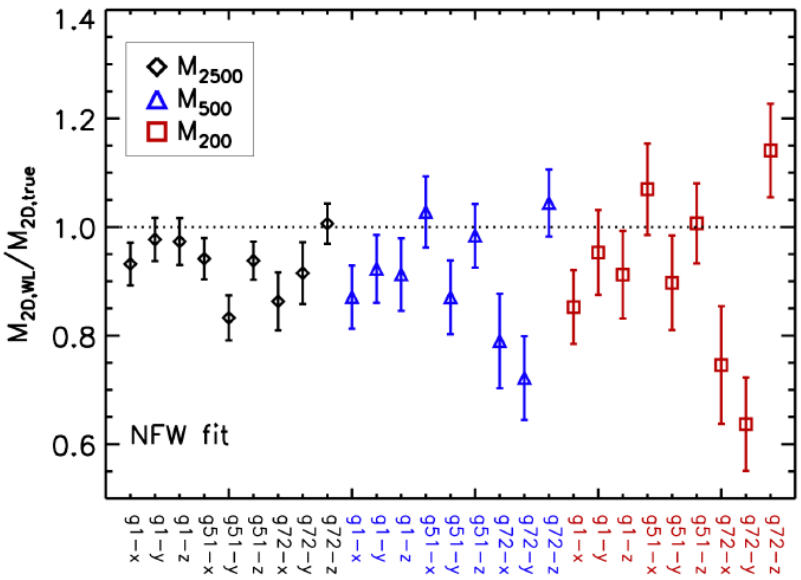


Figure 7.18: Comparison between the NFW-fit and the true 2D-masses of all the simulated clusters. Shown are the ratios between the estimated and the true masses measured at three characteristic radii, namely r_{2500} (diamonds), r_{500} (triangles), and r_{200} (squares), versus the cluster names.

Regarding the basic principles of the two methods, the differences in the mass estimates are not surprising. The NFW-fit method assumes a specific, regular form of the dark-matter halo, which might be a good description for dark-matter halos on average, but not necessarily for individual systems. If a system contains significant amounts of substructure, it is likely that the mass will be underestimated by a profile, which is not able to resolve these additional sub-halos. The nonparametric, joint method does not suffer from this problem by construction and can additionally constrain the profile to much smaller radii, where usually no weak-lensing data are available (compare Sec. 4.5). The radii, within which the different methods deliver reliable mass-profiles when compared to the real profiles, is shown in Tab. 7.3.

7.4.3 Deprojecting the lensing profiles

For a comparison with X-ray results, we need to deproject the lensing profiles. Therefore, we assume spherical symmetry and an underlying NFW-profile to describe the dark-matter halo. The considered fitting radii are given in Tab. 7.3 and their range is based on the ability of the different methods to provide reliable results, as discussed above. One should notice the restriction to the innermost core for strong lensing and the wide range of available scales while combining weak and strong lensing.

Deprojection is done differently for the three methods investigated here. For the strong lensing profiles and the weak-lensing NFW-fit method, we use the fit parameters to calculate the 3D-mass profile of the lenses. For the nonparametric method, we fit the 2D-mass profiles with a projected NFW profile and use the best fit parameters to derive the 3D-mass profiles. Fig. 7.20 shows the 3D profiles for several methods, including two X-ray techniques as described in Meneghetti et al. (2009).

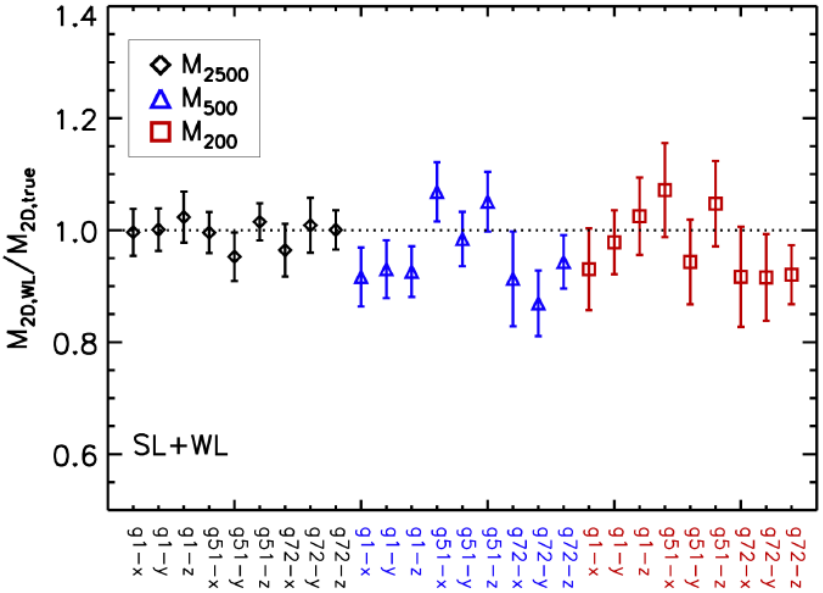


Figure 7.19: As Fig. 7.18 but for the nonparametric method, combining weak and strong lensing.

cluster	min-max SL	min-max WL	min-max SL+WL	min-max X-ray
g1 - x	0.002-0.08	0.16-1.00	0.03-1.00	0.45-0.90
g1 - y	0.002-0.09	0.16-1.00	0.03-1.00	0.45-0.90
g1 - z	0.002-0.06	0.16-1.00	0.03-1.00	0.45-0.90
g51 - x	0.003-0.07	0.20-1.00	0.05-1.00	0.50-1.00
g51 - y	0.002-0.05	0.20-1.00	0.05-1.00	0.50-1.00
g51 - z	0.002-0.12	0.20-1.00	0.05-1.00	0.50-1.00
g72 - x	0.006-0.05	0.20-1.00	0.05-1.00	0.54-1.08
g72 - y	0.006-0.05	0.20-1.00	0.05-1.00	0.54-1.08
g72 - z	0.002-0.05	0.20-1.00	0.05-1.00	0.54-1.08

Table 7.3: Radial ranges, used for fitting NFW profiles to the lensing and X-ray data. All radii are expressed in units of r_{200} reported in Tab. 7.1

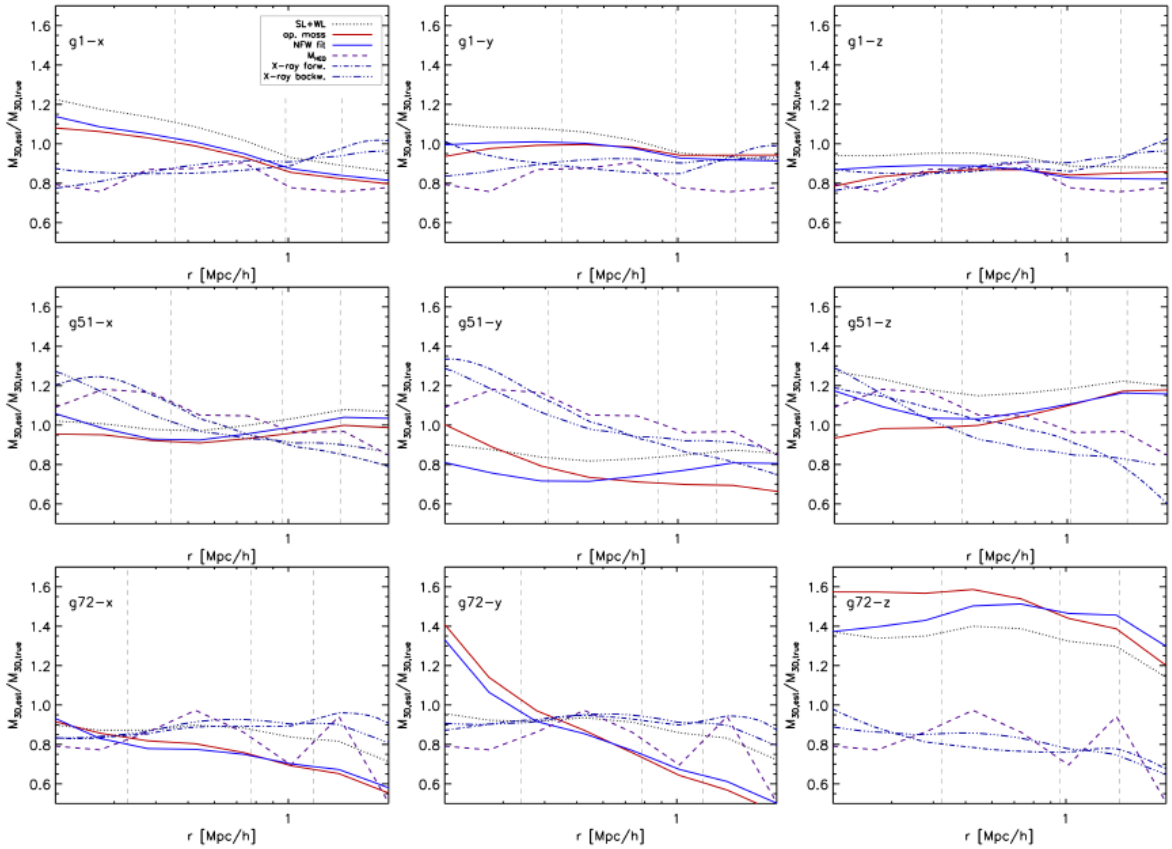


Figure 7.20: Ratios between the mass profiles recovered from the lensing and the X-ray analyses and the true 3D mass profiles of each cluster. The vertical dashed lines in each panel mark the positions of r_{2500}, r_{500}, and r_{200}, as derived from the lensing analysis.

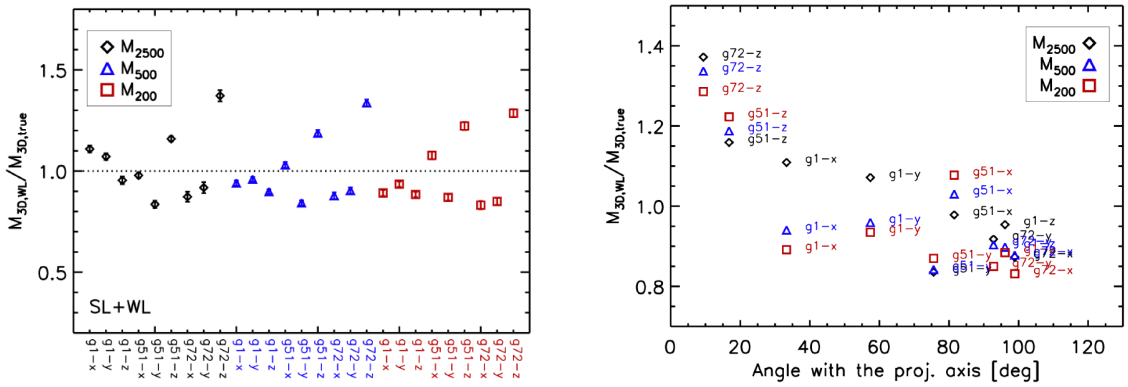


Figure 7.21: The *left panel* shows the same plot as Fig. 7.19 but for the 3D masses. The *right panel* shows the ratio between estimated and true lensing masses as a function of the angle between the major axis of the cluster inertia ellipsoid and the axis along which the mass distribution is projected.

Similarly to Fig. 7.19, we show in Fig. 7.21 the ratios between estimated and true 3D-masses at three over-density radii. It is clear that in 3D the scatter between the estimates and the input masses is significantly larger than in 2D for two main reasons. First, lenses are triaxial, while we are assuming spherical symmetry during the deprojection. The impact of triaxiality on lensing mass estimates was already highlighted by several authors in the past (e.g Gavazzi, 2005; Oguri et al., 2005). Depending on the degree of triaxiality and on the orientation of the clusters with respect to the line-of-sight, 3D-masses may result to be over- or under-estimated. In particular, we find that, in the cases of good alignment (i.e. small angles) between the major axis of the cluster and the projection axis, the lensing masses tend to be systematically larger than the true masses, while the opposite occurs in those cases where the major axis is nearly perpendicular to the line of sight. This is also shown in Fig. 7.21, where the lensing masses are derived with the nonparametric method, which even in 3D seems to provide the most accurate mass estimates. Given that the masses of ellipsoids and spheres with the same azimuthal density profile tend to converge at large distances from their centres, the effect is strongest for M_{2500} and for M_{500}, and mildest for M_{200}. However, even at r_{200}, the analysis of our sample shows that the scatter due to triaxiality is of the order of $\sim 20\%$. Similar results are found by Corless and King (2007).

The second factor, which makes the 3D lensing mass estimates so noisy, is the presence of substructures along the line of sight. Since their distance from the lens plane is unknown, the 3D-mass estimates can be severely affected by these mass clumps, especially if they are located close to the cluster core in projection. The high ratio between the estimated and the true mass of g72-z is in large part due to the presence of the massive sub-clump. This accounts for $\sim 15\%$ of the total cluster mass, but its erroneous inclusion in the central $300\,h^{-1}$ kpc significantly affects the mass estimates at small radii. This peculiar feature of the g72-z projection has been mentioned already several times and can be followed by comparing the three different projections of g72 in Figs. 7.13 - 7.15, while focussing on the appearance of the massive companion.

7.4.4 Comparison between X-ray and lensing

We have deprojected all reconstructions by assuming an NFW-profile. Therefore, we can now compare the best-fits for the concentration and the scale radius between the different

methods. This includes a forward and a backward X-ray method (see Sec. 3.2.2) an the results are shown in Tab. 7.4.

In column 4 of the table, we report the best fit concentrations and scale radii obtained for the main cluster halo using the strong-lensing constraints. These results are obtained by including also the central images in the strong lensing modelling. As discussed in Sec. 7.4.1, this leads to systematically overestimate the concentration and underestimate the scale radius, being the PIEMD model used in the strong-lensing modelling inadequate to describe the distribution of the stars in the simulations. In particular we note that, in some cases, the concentrations are off by a factor of ~ 3 with respect to the true values reported in the second column. On average, the concentrations derived from strong lensing alone are almost 85% larger than the true concentrations of the dark-matter-only profiles. Conversely, the scale radii are almost $\sim 50\%$ smaller. We remind that the strong-lensing fits refer to the cluster halos, while the stars are modelled apart. Such a bias is not present in the simulations without stars, where the average ratio between estimated and true concentrations is around unity.

The weak lensing best fit parameters, obtained from the fit of the shear profiles, are given in column 5. The concentrations tend to be slightly larger than those of the dark-matter distributions in the input models ($\langle c^{\rm WL}/c^{\rm DM\ only}_{\rm true}\rangle = 1.17$) and the scale radii are on average $\sim 10\%$ smaller that the input values ($\langle r_s^{\rm WL}/r^{\rm DM\ only}_{\rm s,true}\rangle = 0.93$). Similar results are found combining strong and weak lensing (columns 6). In this case the average ratio between estimated and true concentrations is $\langle c^{\rm SL+WL}/c^{\rm DM\ only}_{\rm true}\rangle = 1.15$ (for the scale radius we find $\langle r_s^{\rm SL+WL}/r^{\rm DM\ only}_{\rm s,true}\rangle = 0.93$). The tendency to over-estimate more the concentration when adding the strong-lensing constraints is caused by the large contribution of the stellar and gas masses within the inner $100\ h^{-1}$ kpc, as shown in Fig. 7.1. We remind that in the cases of the weak-lensing and combined lensing methods the fits are done over the total projected mass profiles, thus without distinguishing the dark-matter component from the stellar and gas masses. To support this interpretation, we note that the largest discrepancies with the true concentrations arise for the cluster $g51$, which is characterised by an extended strong over-cooling region in the centre, where the total density profile steepens, compared to the dark-matter-only profile. This mimicks a larger concentration, being the mass profile still compatible with an NFW model. Fitting the total density profiles of the inputs clusters indeed leads to higher concentrations and smaller scale radii, as reported in column 3, which are in a much better agreement with the results of the lensing fits. As highlighted in the previous sections, the weak-lensing fits of some systems, like the three projections of $g72$, can be strongly biased assuming a single mass component. Indeed, their substructures need to be properly modelled when deriving the cluster mass from the shear signal. The nonparametric method provides a better chance to measure the density profile.

Comparable estimates of c and $r_{\rm s}$ are obtained with the two X–ray techniques. Even the X-ray analysis probes the total mass distribution, including the gas and the stars. Overall, the deviations from the true estimates go in the same direction: in `g1`, c is underestimated consistently by a factor $0.7-0.8$, with a corresponding overestimate of $r_{\rm s}$ up to a factor 1.45, if compared with the true concentrations, obtained by fitting the input total mass distributions of the clusters; the same considerations apply to `g72`, apart from `g72-z`, where the concentration (scale radius) is estimated higher (lower) than $c^{\rm total}_{\rm true}$ ($r^{\rm total}_{\rm s,true}$) by about 15 % as a consequence of the alignment along the line of sight of the two main clumps. More critical is the case of `g51`, where both the X–ray methods provide a measure of the concentration that is twice $c^{\rm total}_{\rm true}$, due again to the large contribution of the cool substructures in the central regions.

An interesting aspect, while comparing lensing masses with masses derived from X-ray

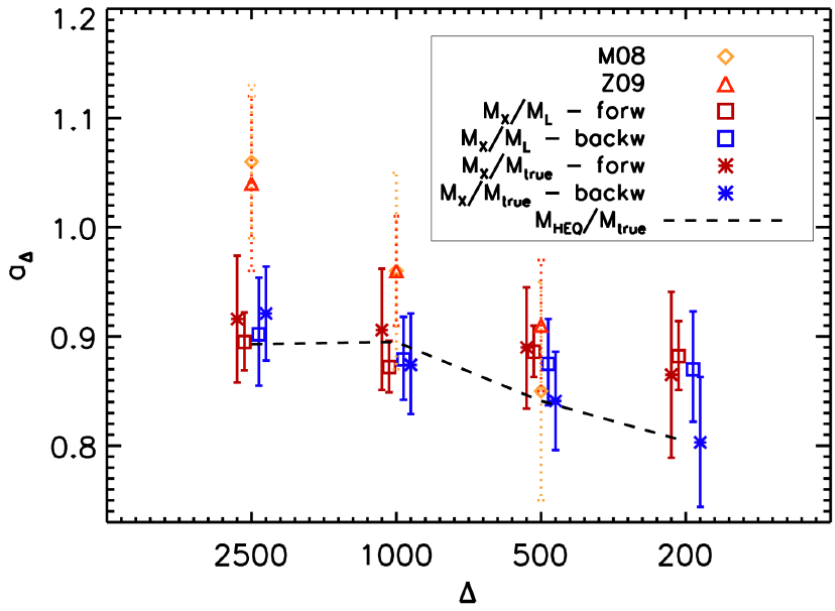


Figure 7.22: Ratio between X-ray and lensing masses as a function of the overdensity Δ (squares). The results are shown for the X-ray masses obtained with the forward (red) and with the backward (blue) methods and for the lensing masses obtained with the nonparametric method. For comparison, we also show the ratios between the X-ray masses and the true masses of the clusters (asterisks) and the ratios between the masses determined via the hydrostatic equilibrium equation using the true gas density and temperature profiles and the true masses (dashed line). The diamonds and the triangles show the results published by Mahdavi et al. (2008) (M08) and Zhang et al. (2010) (Z09), based on the analysis of a sample of 18 and 12 galaxy clusters, respectively. The data-points are slightly shifted along the Δ-axis at each over-density, in order to avoid overlapping and facilitating the comparisons.

methods, is the possibility to trace the lack of hydrostatic equilibrium in a cluster. This has also been discussed from an observational point of view in Mahdavi et al. (2008) and Zhang et al. (2010). Since we deal with simulations, where particle positions and velocities are exactly known in every snapshot, we can analytically calculate the total mass of the cluster under the assumption of hydrostatic equilibrium (Eq. 3.43). We denote this mass with $M_{\rm HEQ}$. Deviations from hydrostatic equilibrium are therefore traced by the ratio $M_{\rm HEQ}/M_{\rm true}$. We compare this ratio with the ratios between X-ray and lensing masses in Fig. 7.22 and indeed identify some weak correlations between the two quantities. A careful analysis is performed in Meneghetti et al. (2009), indicating that lensing masses may indeed be used to calibrate the X-ray mass-observable relations, by correcting for the lack of hydrostatic equilibrium, as it has to be assumed in X-ray mass estimates.

cluster	$c_{\rm true}^{\rm DM\ only}$	$c_{\rm true}^{\rm total}$	$c^{\rm SL}$	$c^{\rm WL}$	$c^{\rm SL+WL}$	$c^{\rm X,forw}$	$c^{\rm X,back}$
g1 - x	4.62	5.38	$11.31^{0.19}_{-0.59}$	6.21 ± 0.87	6.59 ± 0.32	$3.58^{0.06}_{-0.06}$	$3.59^{0.32}_{-0.28}$
g1 - y	4.62	5.38	$10.57^{2.82}_{-1.81}$	4.82 ± 0.64	5.39 ± 0.23	$4.51^{0.08}_{-0.08}$	$4.05^{0.40}_{-0.25}$
g1 - z	4.62	5.38	$6.92^{2.25}_{-1.12}$	4.44 ± 0.62	4.54 ± 0.23	$4.06^{0.07}_{-0.07}$	$3.62^{0.22}_{-0.31}$
g51 - x	5.37	7.20	$3.36^{0.46}_{-0.11}$	5.83 ± 0.96	6.10 ± 0.29	$11.15^{0.19}_{-0.20}$	$10.76^{0.47}_{-0.65}$
g51 - y	5.37	7.20	$9.72^{0.68}_{-0.61}$	5.13 ± 1.05	6.35 ± 0.42	$12.39^{0.22}_{-0.22}$	$10.80^{0.56}_{-0.56}$
g51 - z	5.37	7.20	$8.78^{0.19}_{-0.46}$	6.06 ± 0.93	7.41 ± 0.32	$11.04^{0.21}_{-0.22}$	$11.30^{0.65}_{-0.62}$
g72 - x	3.99	4.22	$6.46^{0.99}_{-2.54}$	4.17 ± 1.54	3.88 ± 0.30	$3.31^{0.05}_{-0.05}$	$3.22^{0.06}_{-0.02}$
g72 - y	3.99	4.22	$7.26^{1.47}_{-0.14}$	7.91 ± 2.86	4.17 ± 0.31	$3.48^{0.05}_{-0.05}$	$3.29^{0.06}_{-0.02}$
g72 - z	3.99	4.22	$11.39^{0.42}_{-0.85}$	4.19 ± 0.62	4.50 ± 0.24	$4.92^{0.08}_{-0.08}$	$4.49^{0.38}_{-0.29}$
$c/c_{\rm true}^{\rm DM\ only}$			1.84 (0.60)	1.17 (0.30)	1.15 (0.15)	1.33 (0.59)	1.25 (0.56)
$c/c_{\rm true}^{\rm total}$			1.59 (0.59)	1.01 (0.33)	0.98 (0.11)	1.09 (0.38)	1.03 (0.36)
$c/c_{\rm true}^{\rm DM\ only}$	no g51		2.09 (0.48)	1.23 (0.35)	1.12 (0.15)	0.92 (0.15)	0.86 (0.12)
$c/c_{\rm true}^{\rm total}$	no g51		1.88 (0.45)	1.12 (0.35)	1.00 (0.12)	0.83 (0.15)	0.78 (0.13)
	$r_{\rm s,true}^{\rm DM\ only}$	$r_{\rm s,true}^{\rm total}$	$r_{\rm s}^{\rm SL}$	$r_{\rm s}^{\rm WL}$	$r_{\rm s}^{\rm SL+WL}$	$r_{\rm s}^{\rm X,forw}$	$r_{\rm s}^{\rm X,back}$
g1 - x	0.310	0.278	$0.097^{0.007}_{-0.002}$	0.229 ± 0.037	0.221 ± 0.012	$0.408^{0.007}_{-0.007}$	$0.409^{0.038}_{-0.037}$
g1 - y	0.310	0.278	$0.090^{0.064}_{-0.006}$	0.307 ± 0.048	0.276 ± 0.013	$0.315^{0.006}_{-0.006}$	$0.360^{0.026}_{-0.036}$
g1 - z	0.310	0.278	$0.152^{0.029}_{-0.031}$	0.317 ± 0.053	0.320 ± 0.017	$0.351^{0.007}_{-0.007}$	$0.403^{0.043}_{-0.027}$
g51 - x	0.241	0.189	$0.385^{0.072}_{-0.053}$	0.242 ± 0.045	0.235 ± 0.012	$0.120^{0.002}_{-0.002}$	$0.124^{0.009}_{-0.006}$
g51 - y	0.241	0.189	$0.099^{0.008}_{-0.012}$	0.246 ± 0.057	0.206 ± 0.014	$0.107^{0.002}_{-0.002}$	$0.124^{0.008}_{-0.007}$
g51 - z	0.241	0.189	$0.141^{0.002}_{-0.003}$	0.244 ± 0.043	0.204 ± 0.010	$0.120^{0.003}_{-0.003}$	$0.116^{0.008}_{-0.007}$
g72 - x	0.299	0.299	$0.087^{0.032}_{-0.009}$	0.262 ± 0.105	0.306 ± 0.025	$0.366^{0.006}_{-0.006}$	$0.386^{0.001}_{-0.007}$
g72 - y	0.299	0.299	$0.094^{0.029}_{-0.027}$	0.135 ± 0.053	0.287 ± 0.023	$0.355^{0.005}_{-0.006}$	$0.386^{0.002}_{-0.008}$
g72 - z	0.299	0.299	$0.059^{0.008}_{-0.004}$	0.367 ± 0.062	0.324 ± 0.018	$0.223^{0.004}_{-0.004}$	$0.251^{0.019}_{-0.022}$
$r_{\rm s}/r_{\rm s,true}^{\rm DM\ only}$			0.49 (0.40)	0.93 (0.20)	0.93 (0.11)	0.89 (0.33)	0.97 (0.36)
$r_{\rm s}/r_{\rm s,true}^{\rm total}$			0.59 (0.53)	1.05 (0.27)	1.04 (0.12)	0.98 (0.32)	1.06 (0.34)
$r_{\rm s}/r_{\rm s,true}^{\rm DM\ only}$	no g51		0.31 (0.09)	0.88 (0.24)	0.95 (0.12)	1.10 (0.18)	1.20 (0.17)
$r_{\rm s}/r_{\rm s,true}^{\rm total}$	no g51		0.34 (0.11)	0.94 (0.26)	1.001 (0.11)	1.17 (0.22)	1.27 (0.21)

Table 7.4: The NFW concentrations and scale radii (upper and lower part of the Table, respectively) resulting from the strong-lensing (column 4), weak-lensing (column 5), strong+weak lensing (column 6), and X-ray analyses (column 7-8 for the forward and for the backward methods) of the clusters in our sample. The weak-lensing estimates are obtained by fitting the shear profile with an NFW model. In columns 2 and 3, we quote the true concentrations obtained by fitting the DM-only and the total density profiles of the three clusters in the radial range between $10\,h^{-1}$ kpc and r_{200}.

8 Weighing real galaxy clusters

After the extensive tests with simulated, but realistic lensing scenarios, we prove the abilities of our method with real observations of galaxy clusters. The two clusters that we consider in our subsequent analysis are very different in their dynamical state and should be considered very interesting test cases for our reconstruction algorithm.

MS2137.3-2353 is a rather relaxed system and is therefore an ideal laboratory for a comparison between X-ray and lensing-mass estimates. This picture changes in the case of CL0024+1654, which is very interesting from a dynamical point of view, but also a mysterious object as we will see in the course of this chapter.

Our main intention in the analysis of those two objects is not primarily to obtain new insights into the physics of these clusters, since they were already in the focus of several decent analyses, but more to show that our method is not only able to recover the results of previous analyses, but moreover to recover the mass distribution on an unequalled range of length scales, where other methods need to artificially combine separate analysis techniques. Furthermore, we will reveal some interesting trends in our reconstructions of real clusters, that were already present in the reconstructions of the simulated clusters in the last chapter. Since in the reconstructions of the simulated data our method performed best, those trends should be considered substantial.

In Sec. 8.1 we will give an overview over the cluster MS2137 (short for MS2137.3-2353), describe the input data for our method and compare our reconstruction with previous studies. In Sec. 8.2 we will follow the same outline for CL0024 (short for CL0024+1654).

8.1 MS2137.3-2353

MS2137 is a rich cluster, dominated by a bright, central cD galaxy at a redshift of $z_c = 0.313$ Stocke et al. (1991). It was the subject of several X-ray studies and appears as a rather spherically symmetric and relaxed system (Gioia et al., 1990; Schmidt and Allen, 2007; Donnarumma et al., 2009), indicating that X-ray masses should be quite reliable in the case of this system.

Apart from its dynamical state, MS2137 shows a particular interesting strong-lensing configuration, consisting of two multiple-image systems. The first system contains a giant, tangential arc, composed of two merging images with opposite parity. Following Donnarumma et al. (2009), we denote these images as A01 and A02 and show their position in the cluster's field in Fig. 8.1. The underlying source produces two additional images, which are confirmed observationally and denoted as A3 and A4. The existence of a demagnified fifth image is of course expected, but its identification is rather difficult due to the luminous foreground light of the BCG. The possible detection of the fifth image is discussed in Gavazzi et al. (2003). The second image system contains a radial arc, in fact the first ever-detected (Fort et al., 1992), denoted as B1 and an additional counter image B2. This configuration marks MS2137 as a textbook example of a strong-lensing system, which becomes pretty obvious if one compares the described configuration with the example in the upper panel of Fig. 2.6.

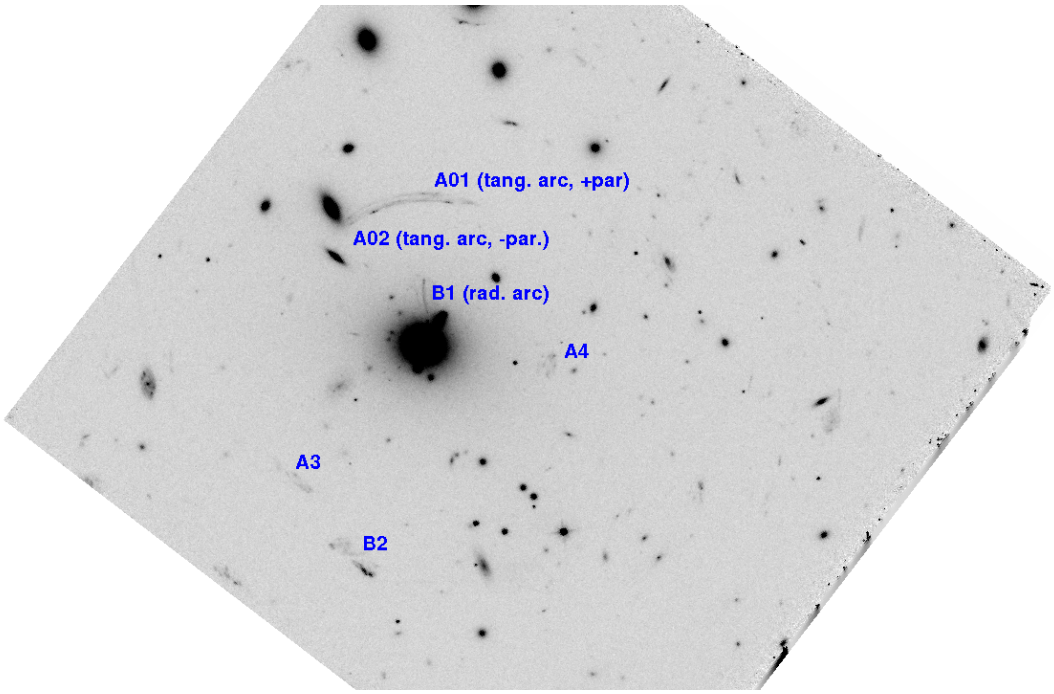


Figure 8.1: 22.2 ksec exposure of the central part of MS2137.3-2353. The image is taken in the F702W filter band of the WFPC2 on board HST. The strong lensing features are accompanied by labels, following the notation of Donnarumma et al. (2009). The side length refers to $\sim 113''$ or $360\ h^{-1}$ kpc at the cluster's redshift of $z_c = 0.313$.

The beautiful lensing configuration was capitalised in several, early reconstructions of the underlying lens model (e.g. Mellier et al., 1993; Miralda-Escude, 1995), which incorporated deep HST imaging later on (Hammer et al., 1997), thereby resolving nicely the structure of the giant arc with its two-fold image configuration. Spectroscopic redshifts of the two multiple-image systems were provided by Sand et al. (2002), finding that the tangential system arises from a source at $z_{\rm tang} = 1.501$ and the radial system represents the images of a distant background galaxy at $z_{\rm rad} = 1.502$.

The HST images, of which we show an example in Fig. 8.1, also gave rise to more recent strong and weak-lensing studies (Gavazzi et al., 2003; Gavazzi, 2005; Comerford et al., 2006) (hereafter G03, G05 and C06), where some also incorporated dynamical constraints on the BCG (Sand et al., 2002, 2008) or X-ray observations (Donnarumma et al., 2009) (hereafter D09). In our analysis, we apply the nonparametric method to study the mass profile of MS2137, while combining weak and strong lensing.

8.1.1 Input data

We used the same ellipticity catalogue as in G03 and G05, which was kindly provided by Raphael Gavazzi (IAP Paris). The catalogue is based on an observation of MS2137 with the FORS and ISAAC instruments at the Very Large Telescope (VLT). With exposure times between 5.28 and 12.0 ksec, the cluster was observed in the optical U, V and I bands and in the infrared J and K bands. The number of available optical and infrared wave bands allowed an accurate photometric analysis, including redshift estimations for the background galaxies using the hyperz package (Bolzonella et al., 2000). For the photo-z estimation, additional B and R

band observations were used and a complete description of the observational data and its reduction can be found in Gavazzi et al. (2003).

The shear estimation was performed with a KSB+ variant, described in Gavazzi et al. (2004) and Gavazzi (2005). In total, the shape analysis provided a catalogue of 1500 ellipticity measurements, in a field of $\sim 400'' \times 400''$ centred on the BCG of MS2137. The distribution of galaxies with shear estimate is shown in Fig. 8.2, corresponding to an average weak-lensing galaxy density of 33.75 arcmin^{-2}.

The strong-lensing constraints were obtained by a deep HST/WFPC2 programme, consisting of 10 exposures in the F702W filter and a stacked exposure time of 22.0 ksec. We used the images A01, A02, A3 and A4, indicated in Fig. 8.1 to obtain constraints on the critical curve of the system at a redshift of $z_{\text{tang}} = 1.501$. We decided to exclude the assumed position of the fifth image, which cannot be clearly identified. The opposite-parity images A01 and A02 allow already for reliable position estimate of the critical line, especially while taking into account the fact, that the redshift of the radial system $z_{\text{rad}} = 1.502$ is practically identical to the tangential system and including the images B1 and B2 in the critical curve estimate. The final position estimators of the radial and tangential critical line, based on a few points near the position of the multiple images, were obtained as described in Sec. 7.3.3 and Merten et al. (2009). The exact positions of these estimates with respect to the weak-lensing galaxies can be seen in Fig. 8.2. The exact coordinates of the strong-lensing features are given by Tab. 3 in D09 or Tab. A.1 in G03.

8.1.2 Reconstruction

We performed the combined analysis with the two catalogues, described above. It should be noted, that the field-of-view of our catalogues is relatively small, compared to the reconstructions of Sec. 7.3.3. This is the reason why the following numbers appear somewhat smaller than in the reconstructions that we have described before and will describe later on.

The outer-level iteration was started on an initial resolution of 10×10 reconstruction pixels and steadily increased until a final resolution of 25×25 pixels was reached, refering to refinement level 0. The adaptively refined resolution in the innermost core of our reconstruction refered to 40×40 pixels on a refinement level of 0. For the adaptive-averaging scheme, we used a minimum of 15 backgrounds galaxies, to be included in the shear estimate of a single reconstruction cell. The final convergence map, as derived from the reconstructed lensing potential, is shown in Fig. 8.2.

A peculiar problem with the small field-of-view is that the assumption of a convergence tending to zero at the boundaries of the observed field is not necessarily valid. We therefore mass-sheet normalised our convergence map with the help of the former reconstruction by D09, which agrees nicely with G03, G05 and C06 at the field boundaries.

An eye-catching feature of the convergence map in Fig. 8.2 is the flattening of the convergence contours in the upper-left area near the cluster core. While comparing the galaxy density, shown in the left panel of the same figure, the reason for this feature becomes clear. Due to a lack of background galaxies, the algorithm was not able to put any reliable constraints on those areas, dominated by extended voids in the background-galaxy distribution. An exception is of course the cluster core itself, since it is reliably constrained by a large number of strong-lensing features.

To obtain a physical surface-mass density out of the convergence map, we used the median redshift of the background sources, being $\langle z_{\text{s}} \rangle = 0.9$, as derived from the photo-z analysis, and the spectroscopic redshifts for the two arcs, together with the redshift of the cluster.

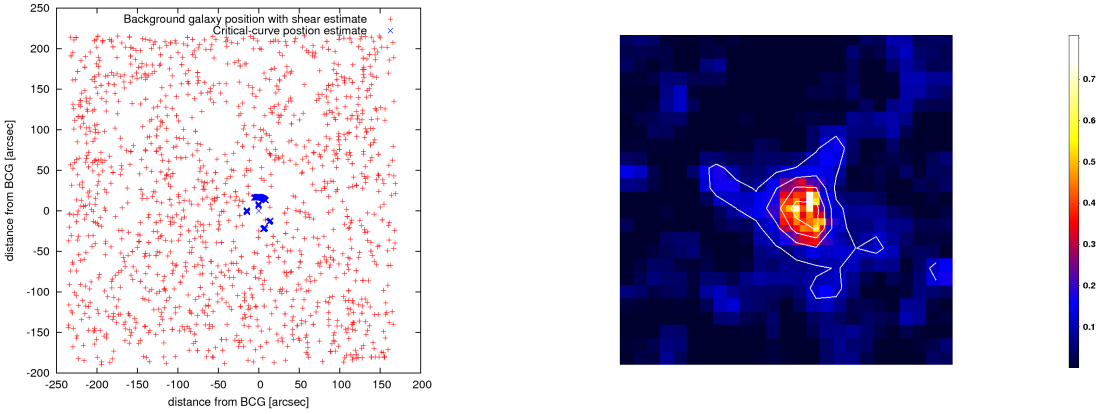


Figure 8.2: The *left panel* shows the constraints used to obtain the convergence map of MS2137, shown in the *right panel*. Red data points indicate the positions of background galaxies and blue data points show the positions of the strong-lensing constraints. It can be seen, that an under-dense region in the upper-left part of the reconstruction coincides with an empty patch in the background-galaxy field. The contours of the convergence map start at $\kappa = 0.1$ with a linear spacing of $\Delta\kappa = 0.1$. The field size in both panels is $\sim 400'' \times 400''$, refering to a physical side length of $\sim 1.3\ h^{-1}$ Mpc at the cluster's redshift of $z_c = 0.313$.

8.1.3 Comparison with previous results and conclusions

In Fig. 8.3, we show the projected mass profile that we obtained from the convergence map in Fig. 8.2. The fact that we mass-sheet normalised our reconstruction according to the profile of D09, is reflected by the perfect match between the two reconstructions at the outer-most radius. This obviously does not hamper the possibility for our reconstruction to differ from this result towards the centre of the cluster.

We compare our mass profile with the weak and strong-lensing results of three former studies. In G05, the ζ-statistics approach (Fahlman et al., 1994; Clowe et al., 1998, 2000; Bartelmann and Schneider, 2001) was used to convert the measured, tangential shear profile into a mass profile, marked by the blue area in Fig. 8.3. As one can see, our result is in perfect agreement with the weak-lensing study of G05, which is reassuring since we use the same ellipticity catalogue and at radii far from the centre, our reconstruction should not be affected by the additional strong-lensing constraints. Strong deviations appear, while comparing the parametric strong-lensing results of G05 with the other parametric strong-lensing reconstructions of C06 and D09. The same is true for the comparison with our nonparametric result at small radii. This issue is also addressed in D09, but no convincing explanation was found. Nevertheless, agreement exists between the other methods in the strong-lensing area of the cluster, including our result. The trend of a shallower mass profile towards the centre of D09 compared to C06, is pursued by our result, while compared to D09. An explanation for the difference in the two parametric strong-lensing methods was given in D09. In the reconstruction by C06, the stellar component was not modelled separately, but was included in the general NFW-profile, modelling the whole dark-matter halo. While choosing the same approach and neglecting their stellar PIEMD-component, D09 found the same model parameters as C06. Interestingly, our method seems to predict an even lower central mass, especially when regarding that we do not have to rely on any model assumptions. We have already seen in Sec. 7.4.1, that modelling the stellar component is difficult and can lead to

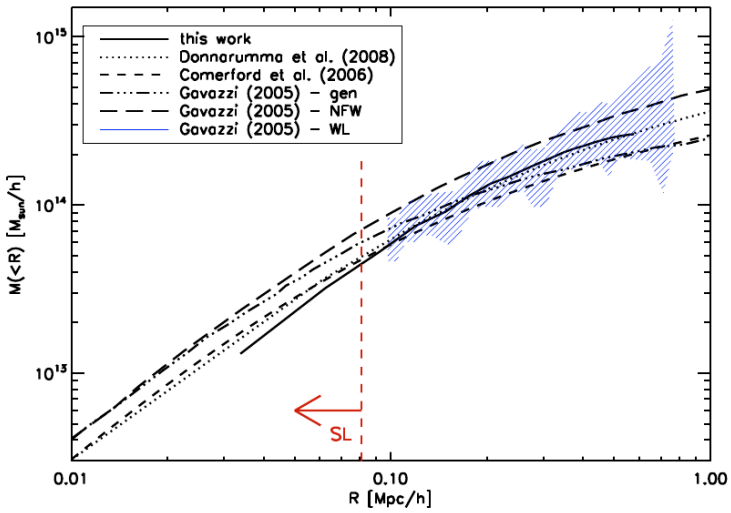


Figure 8.3: The figure shows our reconstructed mass profile of MS2137, as it is published in Merten et al. (2009), in comparison with other reconstructed mass profiles from Gavazzi (2005); Comerford et al. (2006) and Donnarumma et al. (2009). The red line shows the the position of the most distant strong-lensing feature with respect to the zero point of the radial coordinate frame, being the BCG of the cluster.

erroneous results in the strong-lensing analysis. This might explain the slight difference between our profile and the other strong-lensing reconstructions. This explanation is supported by the mass profile of the X-ray reconstruction in D09, based on Chandra data. Since MS2137 is regarded as a relaxed system, X-ray mass estimates should be particularly accurate and indeed, we see exactly the same trends when comparing the lensing result of D09 with our result in Fig. 8.3 and when comparing the X-ray with the lensing result in Fig. 8.4.

8.2 CL0024+1654

The last reconstructed cluster in this work is CL0024, a rich cluster at redshift $z_c = 0.395$. CL0024 is one of the most distant clusters discovered by Zwicky (1959) and, unlike MS2137, it shows a very interesting and irregular dynamical state. A wide-field spectroscopic analysis by Czoske et al. (2001, 2002) unravelled a bimodal redshift distribution in the cluster member galaxies peaked at $z_1 = 0.395$ and $z_2 = 0.381$. Furthermore, the redshift distribution in the central region of the cluster is strongly skewed towards negative velocities. A not less interesting feature is the unusually low X-ray luminosity and gas temperature in the centre of the cluster (Ota et al., 2004; Zhang et al., 2005), regarding the large projected mass of $M_{\rm proj} \approx 10^{15}\ M_\odot$, derived from gravitational lensing (Hoekstra, 2007; Jee et al., 2007; Zitrin et al., 2009; Umetsu et al., 2009b).

A possible explanation for these observations can be found in Ricker and Sarazin (2001), where numerical simulations indicated, that an incomplete merging event between two massive subcomponents, leads to a substantial expansion of the gas component and therefore to the low X-ray emissivity, 1–3 Gyrs after the first encounter. This post-merging scenario was also consulted by Jee et al. (2007), to interpret a ring-like dark matter structure in their weak-lensing analysis. This led to new hydrodynamical simulations, mimicking the suggested line-of-sight merger of CL0024 1–3 Gyrs ago (Zu Hone et al., 2009b,a), without any evidence

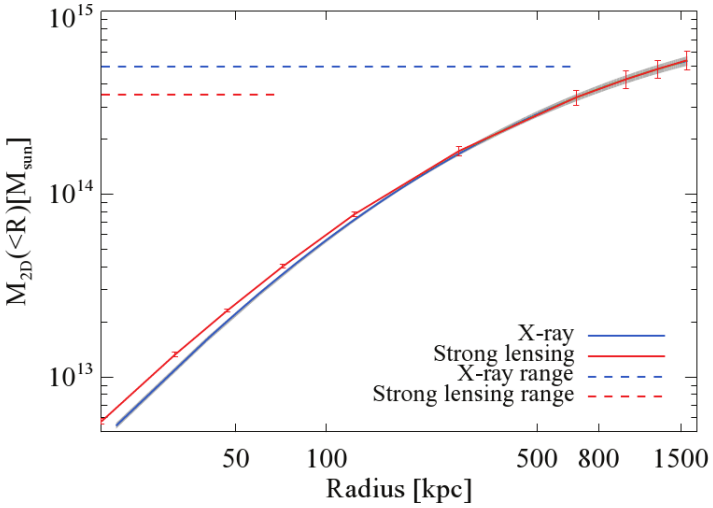


Figure 8.4: From Donnarumma et al. (2009). Projected total mass enclosed in cylinders of radius R. The solid blue line represents the projected total mass profile derived through the X-ray analysis, the red line shows the result of the strong lensing analysis. The blue (red) dashed horizontal line indicates the spatial range over which the X-ray (strong-lensing) analysis has been performed.

for the appearance of such a dark matter ring. An alternative explanation for a sighting like in Jee et al. (2007) was given by Liesenborgs et al. (2008), who showed that a generalisation of the mass-sheet degeneracy in lensing reconstructions, called the monopole degeneracy, can produce reconstruction artefacts of this kind. The new hydro-simulations of this particular candidate merging event in CL0024 predicted quite accurately the observed gas temperature and suggested a head-on collision along the line-of-sight between two mass clumps according to a mass ratio of 2:1.

The interesting dynamical state of CL0024 marks it well-suited and interesting for reconstructions based on gravitational lensing, since, unlike X-ray mass estimates, lensing does not rely on the assumption of hydrostatic equilibrium, which is obviously not guaranteed in a system like CL0024. Indeed, the cluster shows a complete multiple-image system, consisting of a full set of five images. A giant arc, identified by Koo (1988), splits into a triplet of three images when observed with the high angular resolution of the HST. The redshift of this system was determined spectroscopically by Broadhurst et al. (2000) to be $z_5 = 1.625$. Many strong-lensing studies of CL0024 made use of only this single multiple-image system (Kassiola et al., 1992; Smail et al., 1996; Broadhurst et al., 2000), until a recent study by Zitrin et al. (2009) revealed a total number of eleven multiple-image systems, producing a total number of 33 images, as it is expected for a massive lens like CL0024. This work made use of the superior image quality, provided by HST/ACS multi-colour observations and used an iterative modelling technique. Starting from the original system of five images, a lens model is created and additional systems are predicted by this model. If such a system is confirmed photometrically, it is included as constraint to model the lens and the procedure starts again.

Recent weak-lensing studies were performed by Kneib et al. (2003) and Umetsu et al. (2009b) based on HST/WFPC2 and SUBARU imaging, respectively.

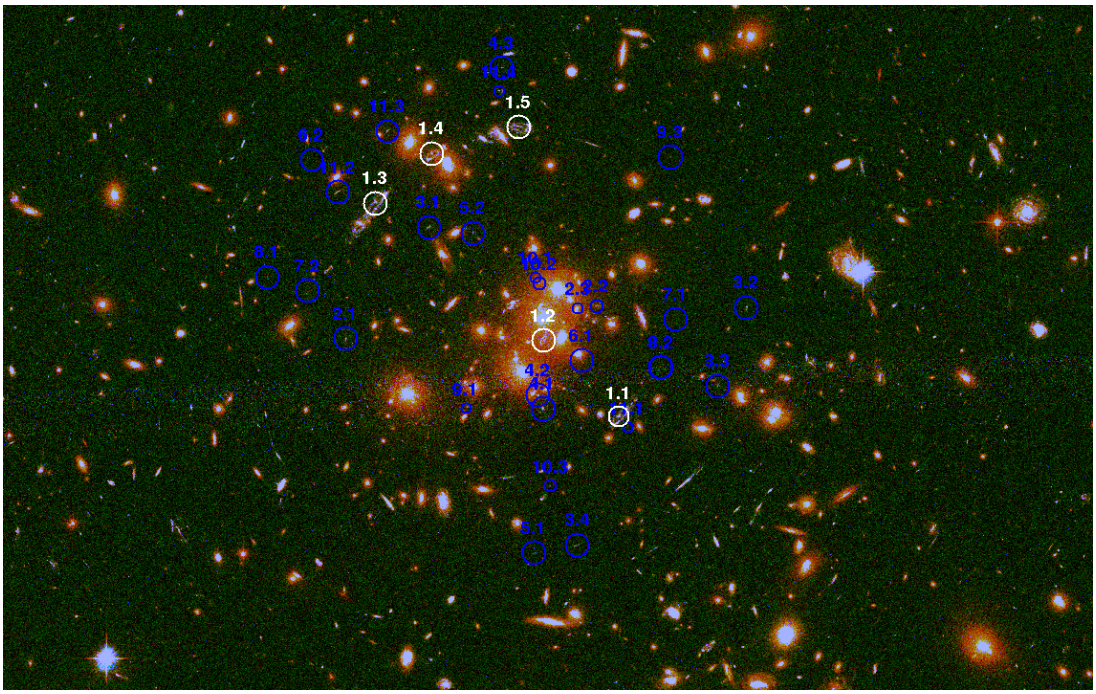


Figure 8.5: This RGB false-colour image shows the central region of CL0024+1654. The R(G,B)-channel refers to the F775W (F555W,F475W) filter band of the HST/ACS. The field size is $185'' \times 116''$, meaning $687 \times 431\ h^{-2}$ kpc^2 at the cluster's redshift of $z_c = 0.395$. Marked in white is the original multiple-image system at a redshift of $z_5 = 1.695$ and in blue the ten additional systems, predicted by the model of Zitrin et al. (2009).

8.2.1 Input data

For our reconstruction we used the weak-lensing catalogue, kindly provided by Keiichi Umetsu (IAA, Taipei). Based on B, R_c and z′-band observations with the SuprimeCam at the SUBARU telescope, a field of $34' \times 26'$, centred on the core of CL0024 was analysed. Using the three-band photometry, the selection technique of Medezinski et al. (2007, 2009) was applied to avoid dilution of the shear signal by foreground and cluster member galaxies. The shear measurement was performed with a KSB+ variant, described in Umetsu et al. (2009a) and Umetsu and Broadhurst (2008). In total, we received a sample of 13680 background galaxies with ellipticity measurement , summing to a total weak-lensing galaxy density of 17.2 arcmin^{-2}. The mean redshift of the background population was found to be $\langle z_s \rangle = 1.31 \pm 0.06$ by the photo-z analysis.

As strong-lensing input, we used the 33 multiple images found by Zitrin et al. (2009) and the respective redshift prediction of their model. All image positions are marked in Fig. 8.5 and, as usual, we used the approach of Merten et al. (2009) to estimate the position of the critical line for different source redshifts from the positions of the multiple images. The exact positions of the multiple-image systems are provided in Tab. 1 of Zitrin et al. (2009) (hereafter Z09).

8.2.2 Reconstruction

We started our nonparametric reconstruction on an initial resolution of 10×9 pixels and refined it in the outer-level iterations until a final resolution of 75×61 pixels, refering to

refinement level 0, was reached. The resolution of the refined grid, covering the central area of the cluster, refered to a resolution of 150×121 pixels on refinement level 0. We used a minimum number of at least 10 galaxies to be included in the grid cell samples of the adaptive-averaging process. The convergence map, as it is obtained from the reconstructed lensing potential, is shown in Fig. 8.6, together with the convergence maps obtained by the analysis of Umetsu et al. (2009b) (hereafter U09) and Kneib et al. (2003) (hereafter K03).

8.2.3 Comparison with previous results and conclusions

The contours of our convergence map (Fig. 8.6) within the innermost region of the cluster are shown in Fig. 8.7, together with the critical curve, predicted by our nonparametric method. As one can see, the contours are nicely centred on the light-distribution as it is expected. More interesting is the critical line, especially when compared to the reconstructed line of Z09, shown in Fig. 4.10 and obtained from a parametric strong-lensing model. The agreement on the general shape of the critical line is remarkable. Even without any model assumptions, our nonparametric approach is able to resolve not only the tangential outer critical line, but also the radial inner one. Strong deviations are noticeable only in those areas, where no input constraints are available and our method has no chance to constrain the critical curve. This can be seen in the central right region of Fig. 8.7, where we also indicate the pixel size of our refined reconstruction grid in the cluster centre. Despite these unconstrained regions, the reconstructed critical line follows the positions of the strong-lensing features. We would like to highlight, that Fig. 8.7 just shows the zoom into the reconstruction of the innermost cluster core. Indeed, we reconstructed CL0024, in one single reconstruction step, on the field shown by Fig. 8.6. As a result, our reconstructed convergence profile is directly comparable with the weak-lensing analysis of U09, determining reliably the cluster in a radius range between $\sim 1'$ distance from the cluster centre until its outskirts and with the strong-lensing analysis of Z09, determining the density profile on a scale $< 1'$. We show this comparison in Fig. 8.8, finding that our reconstructed profile is in good agreement with the other two methods in their respective regimes, but interestingly deviations appear within the error bars exactly at the nexus, where strong-lensing constraints become important and in the innermost core of the cluster. The latter trend is well-known, regarding the findings of the former chapters. The former trend should be investigated carefully, since it might indicate, that combining separate weak and strong-lensing analysis to deliver a combined mass profile (e.g. Oguri et al., 2009; Umetsu et al., 2009b) is not as unproblematic as it is usually assumed.

Since our result seems to be comparable with former weak and strong-lensing reconstructions, we think that it will be interesting in the future to compare our analysis also with mass reconstructions based on dynamical constraints (Diaferio, 1999) and compare the mass profile of CL0024, based on cluster member dynamics (Diaferio et al., 2005) with our analysis. Such a comparison is not trivial since the lensing masses have to be deprojected for comparison. Interestingly, the comparison of the combined weak and strong-lensing analysis in U09 shows deviations between lensing and dynamical mass profiles. Therefore, we would like to deproject our profiles for a comparison, to shed light on the question if CL0024, in its actual appearance, really emerged from a head-on collision along the line-of-sight.

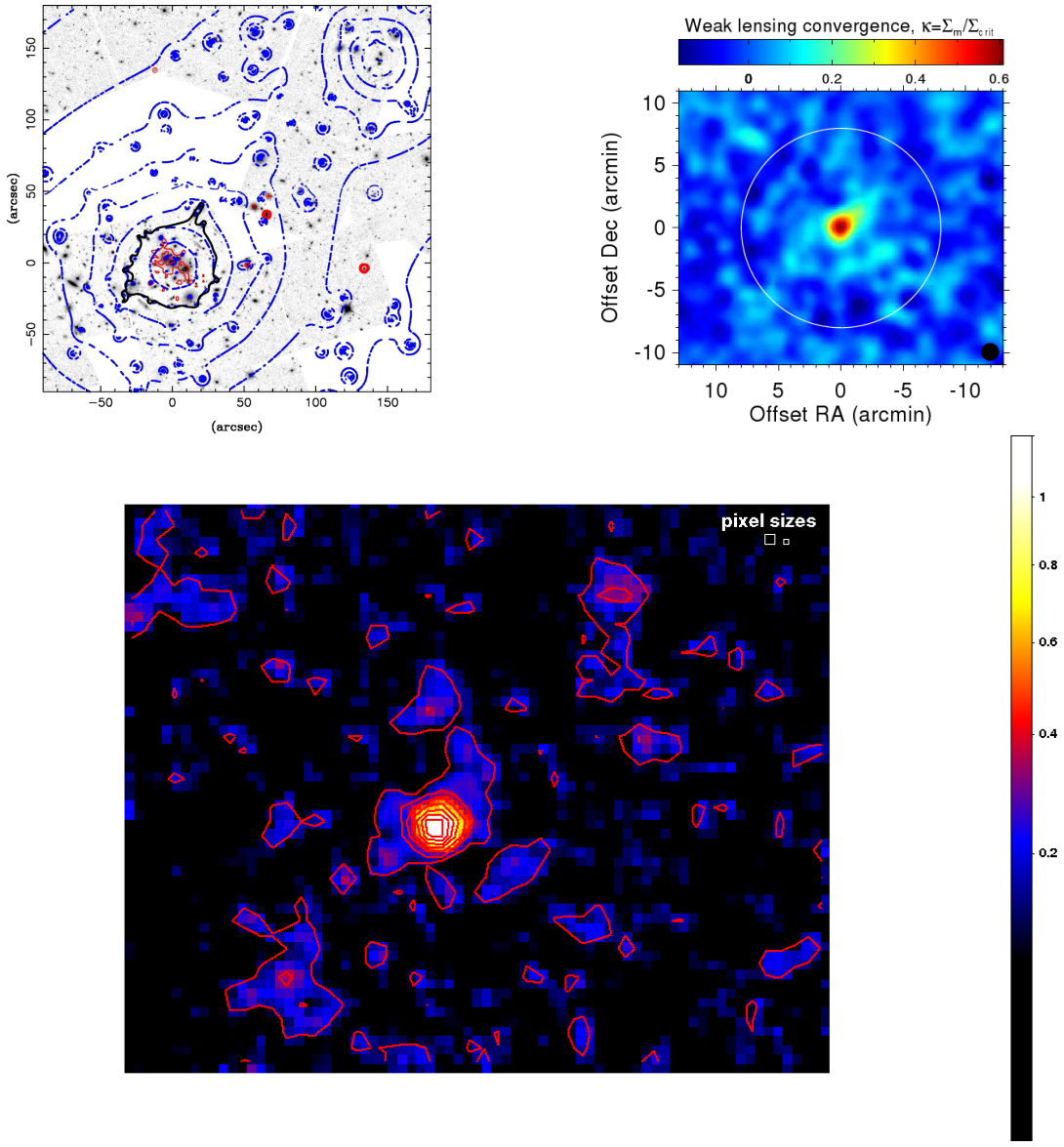


Figure 8.6: This figure compares the reconstructed convergence maps of Kneib et al. (2003) (*top left panel*) and Umetsu et al. (2009b) (*top right panel*) with our reconstruction (*bottom panel*). The general shapes of the maps agree very well, but one should note the different field sizes, directly indicated in the upper two panels and being $1860'' \times 1500''$ for our reconstruction. At a redshift of $z_c = 0.395$ our reconstructed field refers to $\sim\ 6.9 \times 5.6\ h^{-2}$ Mpc^2. The contours in the *bottom panel* start at $\kappa = 0.13$ with a linear spacing of $\Delta\kappa = 0.15$. The largest and the smallest pixel size used in the AMR grid are shown in the top right corner of the convergence map.

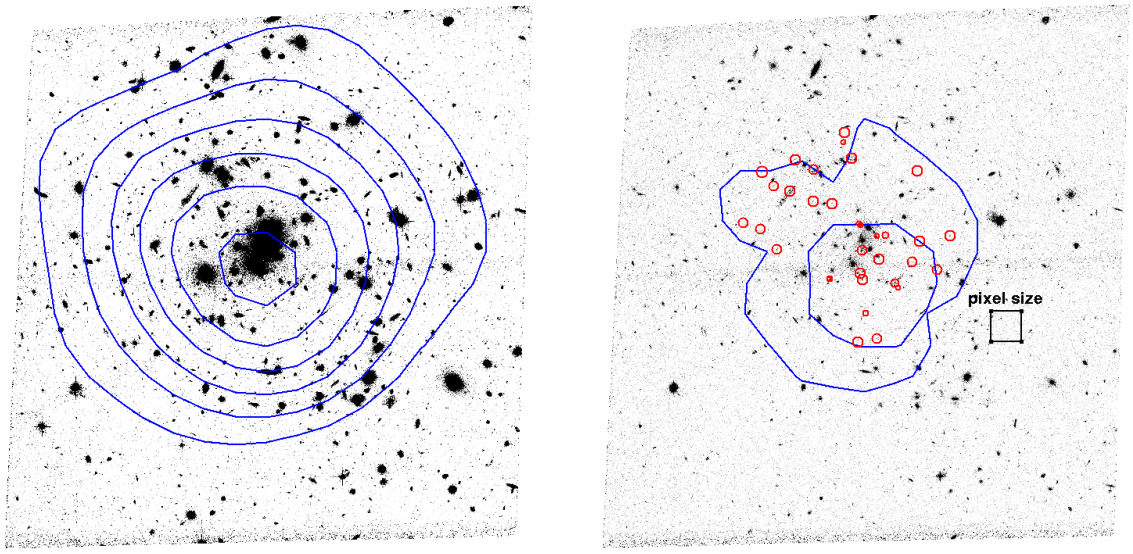


Figure 8.7: HST exposures, showing a slightly larger field than Fig. 8.5. The *left panel* shows the inner contours from the bottom panel of Fig. 8.6 in the F775W filter of the ACS to highlight the cluster members. The central contours coincides nicely with the centre of light. Overlaid on the F475W exposure to highlight the multiple-image systems, the *right panel* shows the reconstructed critical line in blue and multiple-image positions in red. The cuspy shape of the critical line is caused by the still relatively large, refined pixel size, as also indicated by the figure in black.

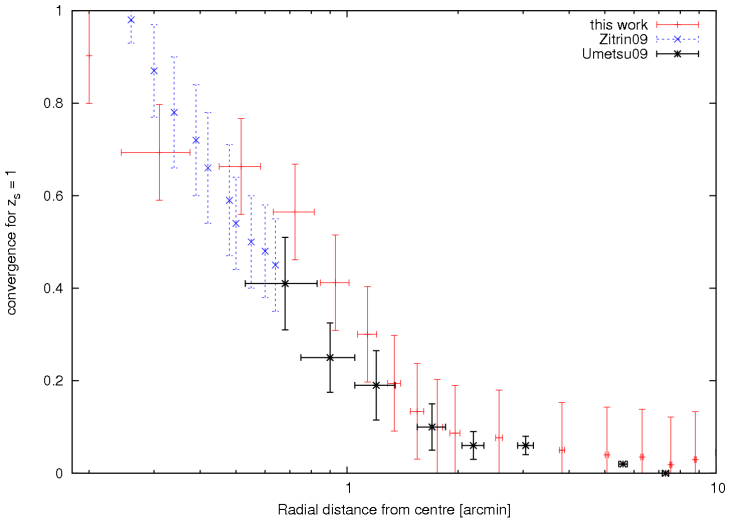


Figure 8.8: The radial convergence profile of CL0024, scaled to a source redshift of $z_s = 1.0$. The red data points show our reconstruction, where the horizontal error bars are derived from the pixel size of the reconstruction grid and the vertical error bars represent the scatter in each radial bin. The blue data points are derived from the parametric reconstruction of Zitrin et al. (2009) and the black data points show the ζ-statistics profile of Umetsu et al. (2009b).

9 Conclusions and outlook

We tried to summarise in this work, why the technique of gravitational lensing is particularly powerful to determine the main properties of galaxy clusters. While in other techniques, like e.g. in X-ray observations, it might be easier to establish a mass-observable relation, they also rely on particular assumptions on the dynamical state of the system. In this sense, gravitational lensing can be assumed to be the more direct method for the determination of essential cluster properties like e.g. the mass profile.

Clusters of galaxies turn out to be interesting objects in a cosmological context not only because they represent the high-mass tail of the mass function, but surprisingly because they still provide a number of unsolved puzzles within the framework of the widely-accepted ΛCDM cosmological model. Several observations do not seem to agree with the results of numerical simulations, like the observed cool cores of clusters, the abundance and strength of the phenomena in the strong-lensing regime or the existence of particular interesting, mostly merging, systems like the `Bullet Cluster`. At the moment it is not clear if this mismatch is due to the wrong interpretations of the observations or to the incomplete treatment of physical processes within the numerical simulations. Most likely the problem arises from the combination of both aspects. Therefore, the observational and the numerical techniques have to calibrated and tested properly, before it can be assumed that there is indeed a problem within our cosmological picture.

In this context we developed and extensively tested a new numerical algorithm, which consistently combines two, usually distinct, aspects of gravitational lensing, being the strong and the weak-lensing regime. Sophisticated numerical strategies have to be applied to implement such a method including the fact that it is fully nonparametric. While not assuming any underlying model for e.g. the global dark-matter density profile or the stellar component of a galaxy cluster, our method is not biased by any flaws in those particular models as they e.g. appeared in our tests when extrapolating the results from strong-lensing analysis to large radii, compared to the Einstein radius of the analysed system.

The tremendous computational power of massively parallel systems helped us to keep the runtime of our computationally extremely demanding method on a very low level. Indeed, the runtime can easily compete with other well-established methods in the weak or in the strong-lensing regime and we have shown, that our method performs better than any of those methods, while reconstructing simulated clusters.

We have proven this, not only with the extensive tests using numerical simulations of clusters and realistic lensing scenarios, but also with the reconstructions of the two well-known strong-lensing clusters `MS2137` and `CL0024`. In the latter case, we showed that our result agrees with the density profiles, obtained from former weak and strong-lensing reconstructions and we would like to use our result for a careful comparison with reconstructions based on dynamical constraints, to find out if this particular cluster is really a relic of a head-on-collision between two massive objects along the line-of-sight.

To summarise the achievements of this work, we would like to explicitly mention the following points, concerning the performance of our method:

1. The comparisons between different methods using extremely realistic lensing scenarios[1] show, that our method is able to recover the mass profiles of simulated clusters with an accuracy of $\lesssim 10\%$ compared to the real profiles of the clusters. Thereby it plays no role if the clusters show complicated, substructured morphologies or if they are rather regular in shape. No other method in the comparison was able to produce results of this quality.

2. While making use of massive parallelisation, we reduced the runtime of our method by a factor of ~ 100. Moreover, while applying efficient, but conceptually simple numerical schemes, we achieved speed-up factors of ~ 1000. Both aspects combine to a runtime improvement expressed by a factor $\sim 10^5$, reducing the time needed for a reconstruction of this kind from the order of years to the order of minutes.

3. To be able to combine different constraints, confined to very different length scales, we successfully implemented the concept of adaptive-mesh-refinement into our algorithm. With this technique, our reconstruction covers a wide range of scales within which the mass profile can be recovered in one singe reconstruction step. In our last reconstruction, using the most recent implementation of our code, we recovered the mass profile of CL0024 in a radial range between $0.1' - 12'$, thereby covering more than two orders of magnitude in radius. In fact, regarding the upper bound, we are only limited by the field size of the observation[2].

But the case of CL0024 shows us also the limitations in our method, that are still present. Our numerical tests have shown, that even when they should not be extrapolated to large radii, parametric strong-lensing reconstruction of the density of the cluster's core deliver excellent results. And also the critical curve of CL0024, as it is obtained with such a method, shows more details than our own reconstruction, especially in areas which are poorly resolved by the observational constraints. We therefore plan to at least partly drop the fully nonparametric nature of our method, while incorporating the full critical line of a cluster, as it is obtained with a parametric method, into our joint reconstruction. This has the additional advantage that parity information on the reconstructed area can be used, meaning that the sign of the Jacobian determinant of the lens mapping can be added as an additional term in our total χ^2-function.

This is not the only additional term which could be added. We have already mentioned that we want to compare our results with reconstructions based on the dynamics of cluster member galaxies. Indeed, we could not only compare to such methods, but we should add the observations constraining them to our own reconstruction. These additional constraints, not based on the shape of background sources, but coming directly from cluster member galaxies, could help to improve the reconstruction of the innermost cluster core of the cluster. In the transient regime between this core and the outskirts of the cluster, we believe that gravitational flexion will play an important role, hence a reliable measurement is at hand. In this work we did not perform a reconstruction which made use of flexion, even if our method is fully able to incorporate such a constraint. We are planning to change this in future reconstructions and we are working hard on improved shape-measurement pipelines. Finally,

[1] In fact we also developed the tools to produce them.

[2] A preliminary result, reconstructing the mass distribution within the full COSMOS field shows already promising results.

in the long run, not only optical but also the tracers of the ICM, so X-ray and SZ observations should contribute to a joint reconstruction, which combines many observational aspects of galaxy clusters in several wave bands.

With such a combined method, most of the shortcomings when connecting observables to cluster properties should be removed. The goal then is to apply this method to a large sample of galaxy clusters, where the observational data should be of outstanding quality. That will deliver the basis for an extensive comparison between the properties of such a sample with numerical simulations. Thereby it is important not only to compare highly-resolved simulations of single objects, but also to compare with large cosmological volumes, including millions of cluster-sized halos, to have good statistics at hand to fix the question of e.g. the preferred dark-matter density profile.

We are sure, that our advanced method to recover the mass-distribution of galaxy clusters will play an important role in such an upcoming project, which will be a big step towards understanding the physical processes in galaxy clusters and thereby testing the ΛCDM paradigm. This will be one particular step in the quest of shedding some light on the dark components of the Universe.

Acknowledgements

I wrote a lot more than I initially intended, so it is time to end and find some last, thankful words about the persons who were important to me during the course of my work.

I would like to start with the scientific grandfathers of this PhD thesis. First, there is Matthias Bartelmann. I know Matthias for some years now and it is needless to say much about his scientific expertise. More important for me is to say, that I would like to thank Matthias for showing me how to be a good scientist, how to attack complicated problems, how to use my time effectively, to learn from others while being creative, how to exploit success and how to learn from mistakes. Matthias was always the person to whom I addressed the important questions and I am most grateful towards him as my mentor.

The same applies to Massimo Meneghetti. Not only that Massimo gave me the opportunity to come to Bologna, but we are also working together since years now and I think we can be proud of what we have achieved. Massimo always had this one excellent idea, but always gave me the freedom to follow my own way. I am looking forward to a lot more years of cooperation in the future.

Special thanks go to Lauro Moscardini who let me be a part of the cosmology group in Bologna and who always had an open ear for my problems and also a solution. Furthermore, a warm thank you to Hans-Walter Rix for being the second corrector of my thesis and to Luca Amendola for completing my defence committee.

I am deeply in debt with a lot of friends and colleagues allover the world, but especially in Heidelberg and Bologna. Because of you all, I actually like my job!

Last but not least, I would like to thank my family and all my friends at home for their support during the last two years of this PhD project.

References

Abel, T., Bryan, G. L., and Norman, M. L. (2000). The Formation and Fragmentation of Primordial Molecular Clouds. *ApJ*, 540:39–44.

Abell, G. O. (1958). The Distribution of Rich Clusters of Galaxies. *ApJS*, 3:211–+.

Abell, G. O., Corwin, Jr., H. G., and Olowin, R. P. (1989). A catalog of rich clusters of galaxies. *ApJS*, 70:1–138.

Agertz, O., Moore, B., Stadel, J., Potter, D., Miniati, F., Read, J., Mayer, L., Gawryszczak, A., Kravtsov, A., Nordlund, Å., Pearce, F., Quilis, V., Rudd, D., Springel, V., Stone, J., Tasker, E., Teyssier, R., Wadsley, J., and Walder, R. (2007). Fundamental differences between SPH and grid methods. *MNRAS*, 380:963–978.

Albrow, M., Birch, P., Caldwell, J. A. R., Martin, R., Menzies, J., Pel, J., Pollard, K., Sackett, P. D., Sahu, K. C., Vreeswijk, P., Williams, A., and Zwaan, M. (1995). PLANET (Probing Lensing Anomalies NETwork). 27:1449–+.

Allen, S. W., Schmidt, R. W., and Fabian, A. C. (2001). The X-ray virial relations for relaxed lensing clusters observed with Chandra. *MNRAS*, 328:L37–L41.

Andrae, R., Melchior, P., and Bartelmann, M. (2010). Soft clustering analysis of galaxy morphologies: A worked example with SDSS. *ArXiv e-prints*, 1002.0676.

Aubert, D., Amara, A., and Metcalf, R. B. (2007). Smooth Particle Lensing. *MNRAS*, 376:113–124.

Bacon, D. J., Goldberg, D. M., Rowe, B. T. P., and Taylor, A. N. (2006). Weak gravitational flexion. *MNRAS*, 365:414–428.

Barnes, J. and Hut, P. (1986). A hierarchical O(N log N) force-calculation algorithm. *Nature*, 324:446–449.

Barsdell, B. R., Barnes, D. G., and Fluke, C. J. (2010). Advanced Architectures for Astrophysical Supercomputing. *ArXiv e-prints*, 1001.2048.

Bartelmann, M. (1996). Arcs from a universal dark-matter halo profile. *A&A*, 313:697–702.

Bartelmann, M. (2010). The dark Universe. *Reviews of Modern Physics*, 82:331–382.

Bartelmann, M., Huss, A., Colberg, J. M., Jenkins, A., and Pearce, F. R. (1998). Arc statistics with realistic cluster potentials. IV. Clusters in different cosmologies. *A&A*, 330:1–9.

Bartelmann, M., Narayan, R., Seitz, S., and Schneider, P. (1996). Maximum-likelihood Cluster Reconstruction. *ApJ*, 464:L115+.

Bartelmann, M. and Schneider, P. (2001). Weak gravitational lensing. *Phys. Rep.*, 340:291–472.

Beaulieu, J., Bennett, D. P., Fouqué, P., Williams, A., Dominik, M., Jørgensen, U. G., Kubas, D., Cassan, A., Coutures, C., Greenhill, J., Hill, K., Menzies, J., Sackett, P. D., Albrow, M., Brillant, S., Caldwell, J. A. R., Calitz, J. J., Cook, K. H., Corrales, E., Desort, M., Dieters, S., Dominis, D., Donatowicz, J., Hoffman, M., Kane, S., Marquette, J., Martin, R., Meintjes, P., Pollard, K., Sahu, K., Vinter, C., Wambsganss, J., Woller, K., Horne, K., Steele, I., Bramich, D. M., Burgdorf, M., Snodgrass, C., Bode, M., Udalski, A., Szymański, M. K., Kubiak, M., Więckowski, T., Pietrzyński, G., Soszyński, I., Szewczyk, O., Wyrzykowski, Ł., Paczyński, B., Abe, F., Bond, I. A., Britton, T. R., Gilmore, A. C., Hearnshaw, J. B., Itow, Y., Kamiya, K., Kilmartin, P. M., Korpela, A. V., Masuda, K., Matsubara, Y., Motomura, M., Muraki, Y., Nakamura, S., Okada, C., Ohnishi, K., Rattenbury, N. J., Sako, T., Sato, S., Sasaki, M., Sekiguchi, T., Sullivan, D. J., Tristram, P. J., Yock, P. C. M., and Yoshioka, T. (2006). Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. *Nature*, 439:437–440.

Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., Caldwell, J. A. R., Ferguson, H. C., Hook, R., Lucas, R. A., Bergeron, L. E., Corbin, M., Jogee, S., Panagia, N., Robberto, M., Royle, P., Somerville, R. S., and Sosey, M. (2006). The Hubble Ultra Deep Field. *AJ*, 132:1729–1755.

Benson, B. A., Church, S. E., Ade, P. A. R., Bock, J. J., Ganga, K. M., Hinderks, J. R., Mauskopf, P. D., Philhour, B., Runyan, M. C., and Thompson, K. L. (2003). Peculiar Velocity Limits from Measurements of the Spectrum of the Sunyaev-Zeldovich Effect in Six Clusters of Galaxies. *ApJ*, 592:674–691.

Bernstein, G. M. and Jarvis, M. (2002). Shapes and Shears, Stars and Smears: Optimal Measurements for Weak Lensing. *AJ*, 123:583–618.

Binney, J. and Tremaine, S. (2008). *Galactic Dynamics*. Princeton University Press.

Birkinshaw, M. (1999). The Sunyaev-Zel'dovich effect. *Phys. Rep.*, 310:97–195.

Birkinshaw, M., Hughes, J. P., and Arnaud, K. A. (1991). A measurement of the value of the Hubble constant from the X-ray properties and the Sunyaev-Zel'dovich effect of Abell 665. *ApJ*, 379:466–481.

Blandford, R. D. and Kochanek, C. S. (1987). Gravitational imaging by isolated elliptical potential wells. I - Cross sections. II - Probability distributions. *ApJ*, 321:658–675.

Boehringer, H., Pratt, G. W., Arnaud, M., Borgani, S., Croston, J. H., Ponman, T. J., Ameglio, S., Temple, R. F., and Dolag, K. (2009). Substructure of the galaxy clusters in the REXCESS sample: observed statistics and comparison to numerical simulations. *ArXiv e-prints*, 0912.4667.

Bolzonella, M., Miralles, J., and Pelló, R. (2000). Photometric redshifts based on standard SED fitting procedures. *A&A*, 363:476–492.

Bond, J. R., Cole, S., Efstathiou, G., and Kaiser, N. (1991). Excursion set mass functions for hierarchical Gaussian fluctuations. *ApJ*, 379:440–460.

Borgani, S. and Kravtsov, A. (2009). Cosmological simulations of galaxy clusters. *ArXiv e-prints*, 0906.4370.

Bouchet, F. R. and Hernquist, L. (1988). Cosmological simulations using the hierarchical tree method. *ApJS*, 68:521–538.

Bradač, M., Allen, S. W., Treu, T., Ebeling, H., Massey, R., Morris, R. G., von der Linden, A., and Applegate, D. (2008). Revealing the Properties of Dark Matter in the Merging Cluster MACS J0025.4-1222. *ApJ*, 687:959–967.

Bradač, M., Clowe, D., Gonzalez, A. H., Marshall, P., Forman, W., Jones, C., Markevitch, M., Randall, S., Schrabback, T., and Zaritsky, D. (2006). Strong and Weak Lensing United. III. Measuring the Mass Distribution of the Merging Galaxy Cluster 1ES 0657-558. *ApJ*, 652:937–947.

Bradač, M., Lombardi, M., and Schneider, P. (2004). Mass-sheet degeneracy: Fundamental limit on the cluster mass reconstruction from statistical (weak) lensing. *A&A*, 424:13–22.

Bradač, M., Schneider, P., Lombardi, M., and Erben, T. (2005). Strong and weak lensing united. *A&A*, 437:39–48.

Bradač, M., Treu, T., Applegate, D., Gonzalez, A. H., Clowe, D., Forman, W., Jones, C., Marshall, P., Schneider, P., and Zaritsky, D. (2009). Focusing Cosmic Telescopes: Exploring Redshift z ˜ 5-6 Galaxies with the Bullet Cluster 1E0657 - 56. *ApJ*, 706:1201–1212.

Brainerd, T. G., Blandford, R. D., and Smail, I. (1996). Weak Gravitational Lensing by Galaxies. *ApJ*, 466:623–+.

Bridle, S., Balan, S. T., Bethge, M., Gentile, M., Harmeling, S., Heymans, C., Hirsch, M., Hosseini, R., Jarvis, M., Kirk, D., Kitching, T., Kuijken, K., Lewis, A., Paulin-Henriksson, S., Scholkopf, B., Velander, M., Voigt, L., Witherick, D., Amara, A., Bernstein, G., Courbin, F., Gill, M., Heavens, A., Mandelbaum, R., Massey, R., Moghaddam, B., Rassat, A., Refregier, A., Rhodes, J., Schrabback, T., Shawe-Taylor, J., Shmakova, M., van Waerbeke, L., and Wittman, D. (2009). Results of the GREAT08 Challenge: An image analysis competition for cosmological lensing. *ArXiv e-prints*, 0908.0945.

Bridle, S., Shawe-Taylor, J., Amara, A., Applegate, D., Balan, S. T., Berge, J., Bernstein, G., Dahle, H., Erben, T., Gill, M., Heavens, A., Heymans, C., High, F. W., Hoekstra, H., Jarvis, M., Kirk, D., Kitching, T., Kneib, J., Kuijken, K., Lagatutta, D., Mandelbaum, R., Massey, R., Mellier, Y., Moghaddam, B., Moudden, Y., Nakajima, R., Paulin-Henriksson, S., Pires, S., Rassat, A., Refregier, A., Rhodes, J., Schrabback, T., Semboloni, E., Shmakova, M., van Waerbeke, L., Witherick, D., Voigt, L., and Wittman, D. (2008). Handbook for the GREAT08 Challenge: An image analysis competition for cosmological lensing. *ArXiv e-prints*, 0802.1214.

Bridle, S. L., Hobson, M. P., Lasenby, A. N., and Saunders, R. (1998). A maximum-entropy method for reconstructing the projected mass distribution of gravitational lenses. *MNRAS*, 299:895–903.

Broadhurst, T., Benítez, N., Coe, D., Sharon, K., Zekser, K., White, R., Ford, H., Bouwens, R., Blakeslee, J., Clampin, M., Cross, N., Franx, M., Frye, B., Hartig, G., Illingworth, G., Infante, L., Menanteau, F., Meurer, G., Postman, M., Ardila, D. R., Bartko, F., Brown, R. A., Burrows, C. J., Cheng, E. S., Feldman, P. D., Golimowski, D. A., Goto, T., Gronwall, C., Herranz, D., Holden, B., Homeier, N., Krist, J. E., Lesser, M. P., Martel, A. R., Miley, G. K., Rosati, P., Sirianni, M., Sparks, W. B., Steindling, S., Tran, H. D., Tsvetanov, Z. I., and Zheng, W. (2005). Strong-Lensing Analysis of A1689 from Deep Advanced Camera Images. *ApJ*, 621:53–88.

Broadhurst, T., Huang, X., Frye, B., and Ellis, R. (2000). A Spectroscopic Redshift for the Cl 0024+16 Multiple Arc System: Implications for the Central Mass Distribution. *ApJ*, 534:L15–L18.

Broadhurst, T., Umetsu, K., Medezinski, E., Oguri, M., and Rephaeli, Y. (2008). Comparison of Cluster Lensing Profiles with ΛCDM Predictions. *ApJ*, 685:L9–L12.

Broadhurst, T. J., Taylor, A. N., and Peacock, J. A. (1995). Mapping cluster mass distributions via gravitational lensing of background galaxies. *ApJ*, 438:49–61.

Cacciato, M., Bartelmann, M., Meneghetti, M., and Moscardini, L. (2006). Combining weak and strong lensing in cluster potential reconstruction. *A&A*, 458:349–356.

Carlstrom, J. E. (2002). Cosmology with the Sunyaev-Zel'dovich Effect. *APS Meeting Abstracts*.

Carlstrom, J. E., Ade, P. A. R., Aird, K. A., Benson, B. A., Bleem, L. E., Busetti, S., Chang, C. L., Chauvin, E., Cho, H., Crawford, T. M., Crites, A. T., Dobbs, M. A., Halverson, N. W., Heimsath, S., Holzapfel, W. L., Hrubes, J. D., Joy, M., Keisler, R., Lanting, T. M., Lee, A. T., Leitch, E. M., Leong, J., Lu, W., Lueker, M., McMahon, J. J., Mehl, J., Meyer, S. S., Mohr, J. J., Montroy, T. E., Padin, S., Plagge, T., Pryke, C., Ruhl, J. E., Schaffer, K. K., Schwan, D., Shirokoff, E., Spieler, H. G., Staniszewski, Z., Stark, A. A., and Vieira, K. V. J. D. (2009). The 10 Meter South Pole Telescope. *ArXiv e-prints*, 0907.4445.

Carroll, S. (2003). *Spacetime and Geometry: Introduction to General Relativity*. Prentice Hall.

Cavaliere, A. and Fusco-Femiano, R. (1978). The Distribution of Hot Gas in Clusters of Galaxies. *A&A*, 70:677–+.

Cen, R. (1997). Toward Understanding Galaxy Clusters and Their Constituents: Projection Effects on Velocity Dispersion, X-Ray Emission, Mass Estimates, Gas Fraction, and Substructure. *ApJ*, 485:39–+.

Chandrasekhar, S. (1943). Stochastic Problems in Physics and Astronomy. *Reviews of Modern Physics*, 15:1–89.

Charbonneau, P. (1995). Genetic Algorithms in Astronomy and Astrophysics. *ApJS*, 101:309–+.

Cimatti, A., Robberto, M., Baugh, C., Beckwith, S. V. W., Content, R., Daddi, E., De Lucia, G., Garilli, B., Guzzo, L., Kauffmann, G., and 113 coauthors. (2009). SPACE: the spectroscopic all-sky cosmic explorer. *Experimental Astronomy*, 23:39–66.

Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C., and Zaritsky, D. (2006). A Direct Empirical Proof of the Existence of Dark Matter. *ApJ*, 648:L109–L113.

Clowe, D., De Lucia, G., and King, L. (2004a). Effects of asphericity and substructure on the determination of cluster mass with weak gravitational lensing. *MNRAS*, 350:1038–1048.

Clowe, D., Gonzalez, A., and Markevitch, M. (2004b). Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter. *ApJ*, 604:596–603.

Clowe, D., Luppino, G. A., Kaiser, N., and Gioia, I. M. (2000). Weak Lensing by High-Redshift Clusters of Galaxies. I. Cluster Mass Reconstruction. *ApJ*, 539:540–560.

Clowe, D., Luppino, G. A., Kaiser, N., Henry, J. P., and Gioia, I. M. (1998). Weak Lensing by Two Z approximately 0.8 Clusters of Galaxies. *ApJ*, 497:L61+.

Coe, D., Benítez, N., Sánchez, S. F., Jee, M., Bouwens, R., and Ford, H. (2006). Galaxies in the Hubble Ultra Deep Field. I. Detection, Multiband Photometry, Photometric Redshifts, and Morphology. *AJ*, 132:926–959.

Comerford, J. M., Meneghetti, M., Bartelmann, M., and Schirmer, M. (2006). Mass Distributions of Hubble Space Telescope Galaxy Clusters from Gravitational Arcs. *ApJ*, 642:39–47.

Corless, V. L. and King, L. J. (2007). A statistical study of weak lensing by triaxial dark matter haloes: consequences for parameter estimation. *MNRAS*, 380:149–161.

Couchman, H. M. P. (1991). Mesh-refined P3M - A fast adaptive N-body algorithm. *ApJ*, 368:L23–L26.

Crittenden, R. G., Natarajan, P., Pen, U., and Theuns, T. (2002). Discriminating Weak Lensing from Intrinsic Spin Correlations Using the Curl-Gradient Decomposition. *ApJ*, 568:20–27.

Crocce, M., Fosalba, P., Castander, F. J., and Gaztañaga, E. (2010). Simulating the Universe with MICE: the abundance of massive clusters. *MNRAS*, pages 198–+.

Crocce, M. and Scoccimarro, R. (2006). Renormalized cosmological perturbation theory. *Phys. Rev. D*, 73(6):063519–+.

Czoske, O., Kneib, J., Soucail, G., Bridges, T. J., Mellier, Y., and Cuillandre, J. (2001). A wide-field spectroscopic survey of the cluster of galaxies <ASTROBJ>Cl0024+1654</ASTROBJ>. I. The catalogue. *A&A*, 372:391–405.

Czoske, O., Moore, B., Kneib, J., and Soucail, G. (2002). A wide-field spectroscopic survey of the cluster of galaxies <ASTROBJ>Cl0024+1654</ASTROBJ>. II. A high-speed collision? *A&A*, 386:31–41.

Davison A.C., H. D. (1997). *Bootstrap Methods and Their Application*. Cambridge University Press.

Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, Inc., New York.

Deb, S., Goldberg, D. M., and Ramdass, V. J. (2008). Reconstruction of Cluster Masses Using Particle based Lensing. I. Application to Weak Lensing. *ApJ*, 687:39–49.

Diaferio, A. (1999). Mass estimation in the outer regions of galaxy clusters. *MNRAS*, 309:610–622.

Diaferio, A., Geller, M. J., and Rines, K. J. (2005). Caustic and Weak-Lensing Estimators of Galaxy Cluster Masses. *ApJ*, 628:L97–L100.

Diego, J. M., Tegmark, M., Protopapas, P., and Sandvik, H. B. (2007). Combined reconstruction of weak and strong lensing data with WSLAP. *MNRAS*, 375:958–970.

Dine, M. (2007). *Supersymmetry and String Theory: Beyond the Standard Model*. Cambridge University Press.

Dodelson, S. (2003). *Modern Cosmology*. Academic Press.

Dolag, K., Borgani, S., Murante, G., and Springel, V. (2009). Substructures in hydrodynamical cluster simulations. *MNRAS*, 399:497–514.

Dolag, K. and Stasyszyn, F. (2009). An MHD GADGET for cosmological simulations. *MNRAS*, 398:1678–1697.

Dolag, K., Vazza, F., Brunetti, G., and Tormen, G. (2005). Turbulent gas motions in galaxy cluster simulations: the role of smoothed particle hydrodynamics viscosity. *MNRAS*, 364:753–772.

Donnarumma, A., Ettori, S., Meneghetti, M., Gavazzi, R., Fort, B., Moscardini, L., Romano, A., Fu, L., Giordano, F., Radovich, M., Maoli, R., Scaramella, R., and Richard, J. (2010). Abell 611. II. X-ray and strong lensing analyses. *ArXiv e-prints*, 1002.1625.

Donnarumma, A., Ettori, S., Meneghetti, M., and Moscardini, L. (2009). X-ray and strong lensing mass estimate of MS2137.3-2353. *MNRAS*, 398:438–450.

Duffy, A. R., Schaye, J., Kay, S. T., Dalla Vecchia, C., Battye, R. A., and Booth, C. M. (2010). Impact of baryon physics on dark matter structures: a detailed simulation study of halo density profiles. *ArXiv e-prints*, 1001.3447.

Dunkley, J., Spergel, D. N., Komatsu, E., Hinshaw, G., Larson, D., Nolta, M. R., Odegard, N., Page, L., Bennett, C. L., Gold, B., Hill, R. S., Jarosik, N., Weiland, J. L., Halpern, M., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., Wollack, E., and Wright, E. L. (2009). Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Bayesian Estimation of Cosmic Microwave Background Polarization Maps. *ApJ*, 701:1804–1813.

Durrer, R. (2008). *The Cosmic Microwave Background*. Cambridge University Press.

Einasto, J. and Haud, U. (1989). Galactic models with massive corona. I - Method. II - Galaxy. *A&A*, 223:89–106.

Eisenstein, D. J. (2005). Dark energy and cosmic sound [review article]. *New Astronomy Review*, 49:360–365.

Ettori, S., De Grandi, S., and Molendi, S. (2002). Gravitating mass profiles of nearby galaxy clusters and relations with X-ray gas temperature, luminosity and mass. *A&A*, 391:841–855.

Fahlman, G., Kaiser, N., Squires, G., and Woods, D. (1994). Dark matter in MS 1224 from distortion of background galaxies. *ApJ*, 437:56–62.

Falco, E. E., Gorenstein, M. V., and Shapiro, I. I. (1985). On model-dependent bounds on H(0) from gravitational images Application of Q0957 + 561A,B. *ApJ*, 289:L1–L4.

Fedeli, C. and Bartelmann, M. (2007). Effects of early dark energy on strong cluster lensing. *A&A*, 461:49–57.

Fedeli, C., Bartelmann, M., Meneghetti, M., and Moscardini, L. (2008). Strong lensing statistics and the power spectrum normalisation. *A&A*, 486:35–44.

Fedeli, C., Meneghetti, M., Bartelmann, M., Dolag, K., and Moscardini, L. (2006). A fast method for computing strong-lensing cross sections: application to merging clusters. *A&A*, 447:419–430.

Fort, B., Le Fevre, O., Hammer, F., and Cailloux, M. (1992). An arc system with a radial gravitational image in the cluster MS 2137-23. *ApJ*, 399:L125–L127.

Freedman, W. L., Madore, B. F., Gibson, B. K., Ferrarese, L., Kelson, D. D., Sakai, S., Mould, J. R., Kennicutt, Jr., R. C., Ford, H. C., Graham, J. A., Huchra, J. P., Hughes, S. M. G., Illingworth, G. D., Macri, L. M., and Stetson, P. B. (2001). Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant. *ApJ*, 553:47–72.

Fryxell, B., Olson, K., Ricker, P., Timmes, F. X., Zingale, M., Lamb, D. Q., MacNeice, P., Rosner, R., Truran, J. W., and Tufo, H. (2000). FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes. Technical report.

Fu, L., Semboloni, E., Hoekstra, H., Kilbinger, M., van Waerbeke, L., Tereno, I., Mellier, Y., Heymans, C., Coupon, J., Benabed, K., Benjamin, J., Bertin, E., Doré, O., Hudson, M. J., Ilbert, O., Maoli, R., Marmo, C., McCracken, H. J., and Ménard, B. (2008). Very weak lensing in the CFHTLS wide: cosmology from cosmic shear in the linear regime. *A&A*, 479:9–25.

Gardini, A., Rasia, E., Mazzotta, P., Tormen, G., De Grandi, S., and Moscardini, L. (2004). Simulating Chandra observations of galaxy clusters. *MNRAS*, 351:505–514.

Gavazzi, R. (2005). Projection effects in cluster mass estimates: the case of MS2137-23. *A&A*, 443:793–804.

Gavazzi, R., Fort, B., Mellier, Y., Pelló, R., and Dantel-Fort, M. (2003). A radial mass profile analysis of the lensing cluster MS 2137.3-2353. *A&A*, 403:11–27.

Gavazzi, R., Mellier, Y., Fort, B., Cuillandre, J., and Dantel-Fort, M. (2004). Mass and light in the supercluster of galaxies MS0302+17. *A&A*, 422:407–422.

Gavazzi, R., Treu, T., Koopmans, L. V. E., Bolton, A. S., Moustakas, L. A., Burles, S., and Marshall, P. J. (2008). The Sloan Lens ACS Survey. VI. Discovery and Analysis of a Double Einstein Ring. *ApJ*, 677:1046–1059.

Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., Dickinson, M., Alexander, D. M., Bauer, F. E., Bergeron, J., Biagetti, C., Brandt, W. N., Casertano, S., Cesarsky, C., Chatzichristou, E., Conselice, C., Cristiani, S., Da Costa, L., Dahlen, T., de Mello, D., Eisenhardt, P., Erben, T., Fall, S. M., Fassnacht, C., Fosbury, R., Fruchter, A., Gardner, J. P., Grogin, N., Hook, R. N., Hornschemeier, A. E., Idzi, R., Jogee, S., Kretchmer, C., Laidler, V., Lee, K. S., Livio, M., Lucas, R., Madau, P., Mobasher, B., Moustakas, L. A., Nonino, M., Padovani, P., Papovich, C., Park, Y., Ravindranath, S., Renzini, A., Richardson, M., Riess, A., Rosati, P., Schirmer, M., Schreier, E., Somerville, R. S., Spinrad, H., Stern, D., Stiavelli, M., Strolger, L., Urry, C. M., Vandame, B., Williams, R., and Wolf, C. (2004). The Great Observatories Origins Deep Survey: Initial Results from Optical and Near-Infrared Imaging. *ApJ*, 600:L93–L98.

Gioia, I. M., Maccacaro, T., Schild, R. E., Wolter, A., Stocke, J. T., Morris, S. L., and Henry, J. P. (1990). The Einstein Observatory Extended Medium-Sensitivity Survey. I - X-ray data and analysis. *ApJS*, 72:567–619.

Gladders, M. D. and Yee, H. K. C. (2000). A New Method For Galaxy Cluster Detection. I. The Algorithm. *AJ*, 120:2148–2162.

Goldberg, D. M. and Bacon, D. J. (2005). Galaxy-Galaxy Flexion: Weak Lensing to Second Order. *ApJ*, 619:741–748.

Goldberg, D. M. and Leonard, A. (2007). Measuring Flexion. *ApJ*, 660:1003–1015.

Gorenstein, M. V., Shapiro, I. I., and Falco, E. E. (1988). Degeneracies in parameter estimates for models of gravitational lens systems. *ApJ*, 327:693–711.

Gottlöber, S., Yepes, G., Khalatyan, A., Sevilla, R., and Turchaninov, V. (2006). Dark and baryonic matter in the MareNostrum Universe. 878:3–9.

Grazian, A., Fontana, A., De Santis, C., Gallozzi, S., Giallongo, E., and Di Pangrazio, F. (2004). The Large Binocular Camera Image Simulator. *PASP*, 116:750–761.

Grazian, A., Fontana, A., de Santis, C., Nonino, M., Salimbeni, S., Giallongo, E., Cristiani, S., Gallozzi, S., and Vanzella, E. (2006). The GOODS-MUSIC sample: a multicolour catalog of near-IR selected galaxies in the GOODS-South field. *A&A*, 449:951–968.

Halkola, A., Hildebrandt, H., Schrabback, T., Lombardi, M., Bradač, M., Erben, T., Schneider, P., and Wuttke, D. (2008). The mass distribution of RX J1347-1145 from strong lensing. *A&A*, 481:65–77.

Halkola, A., Seitz, S., and Pannella, M. (2006). Parametric strong gravitational lensing analysis of Abell 1689. *MNRAS*, 372:1425–1462.

Hammer, F., Gioia, I. M., Shaya, E. J., Teyssandier, P., Le Fevre, O., and Luppino, G. A. (1997). Detailed Lensing Properties of the MS 2137-2353 Core and Reconstruction of Sources from Hubble Space Telescope Imagery. *ApJ*, 491:477–+.

Harrison, E. R. (1970). Fluctuations at the Threshold of Classical Cosmology. *Phys. Rev. D*, 1:2726–2730.

Hassan, A. H., Fluke, C. J., and Barnes, D. G. (2010). GPU-Based Volume Rendering of Noisy Multi-Spectral Astronomical Data. *ArXiv e-prints*, 1001.2049.

Heymans, C., Brown, M. L., Barden, M., Caldwell, J. A. R., Jahnke, K., Peng, C. Y., Rix, H., Taylor, A., Beckwith, S. V. W., Bell, E. F., Borch, A., Häußler, B., Jogee, S., McIntosh, D. H., Meisenheimer, K., Sánchez, S. F., Somerville, R., Wisotzki, L., and Wolf, C. (2005). Cosmological weak lensing with the HST GEMS survey. *MNRAS*, 361:160–176.

Heymans, C., Van Waerbeke, L., Bacon, D., Berge, J., Bernstein, G., Bertin, E., Bridle, S., Brown, M. L., Clowe, D., Dahle, H., Erben, T., Gray, M., Hetterscheidt, M., Hoekstra, H., Hudelot, P., Jarvis, M., Kuijken, K., Margoniner, V., Massey, R., Mellier, Y., Nakajima, R., Refregier, A., Rhodes, J., Schrabback, T., and Wittman, D. (2006). The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations. *MNRAS*, 368:1323–1339.

Hildebrandt, H., van Waerbeke, L., and Erben, T. (2009). CARS: The CFHTLS-Archive-Research Survey. III. First detection of cosmic magnification in samples of normal high-z galaxies. *A&A*, 507:683–691.

Hockney, R. W. and Eastwood, J. W. (1988). *Computer simulation using particles*.

Hoekstra, H. (2001). The effect of distant large scale structure on weak lensing mass estimates. *A&A*, 370:743–753.

Hoekstra, H. (2003). How well can we determine cluster mass profiles from weak lensing? *MNRAS*, 339:1155–1162.

Hoekstra, H. (2007). A comparison of weak-lensing masses and X-ray properties of galaxy clusters. *MNRAS*, 379:317–330.

Hoekstra, H., Franx, M., Kuijken, K., and Squires, G. (1998). Weak Lensing Analysis of CL 1358+62 Using Hubble Space Telescope Observations. *ApJ*, 504:636–+.

Hoekstra, H., Mellier, Y., van Waerbeke, L., Semboloni, E., Fu, L., Hudson, M. J., Parker, L. C., Tereno, I., and Benabed, K. (2006). First Cosmic Shear Results from the Canada-France-Hawaii Telescope Wide Synoptic Legacy Survey. *ApJ*, 647:116–127.

Horesh, A., Ofek, E. O., Maoz, D., Bartelmann, M., Meneghetti, M., and Rix, H. (2005). The Lensed Arc Production Efficiency of Galaxy Clusters: A Comparison of Matched Observed and Simulated Samples. *ApJ*, 633:768–780.

Hosseini, R. and Bethge, M. (2009). Spectral stacking: Unbiased shear estimation for weak gravitational lensing. Technical report, Max Planck Institute for Biological Cybernetics.

Hu, C., Chen, C., and Chen, P. (2010). A GPU-based Calculation Method for Near Field Effects of Cherenkov Radiation Induced by Ultra High Energy Cosmic Neutrinos. *ArXiv e-prints*, 1001.5341.

Hu, W. (2003). Self-consistency and calibration of cluster number count surveys for dark energy. *Phys. Rev. D*, 67(8):081304–+.

Hubble, E. (1929). A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173.

Iapichino, L. and Niemeyer, J. C. (2008). Hydrodynamical adaptive mesh refinement simulations of turbulent flows - II. Cosmological simulations of galaxy clusters. *MNRAS*, 388:1089–1100.

Inogamov, N. A. and Sunyaev, R. A. (2003). Turbulence in Clusters of Galaxies and X-ray Line Profiles. *Astronomy Letters*, 29:791–824.

Itoh, N. and Nozawa, S. (2004). Relativistic corrections to the Sunyaev-Zeldovich effect for extremely hot clusters of galaxies. *A&A*, 417:827–832.

Jee, M. J., Ford, H. C., Illingworth, G. D., White, R. L., Broadhurst, T. J., Coe, D. A., Meurer, G. R., van der Wel, A., Benítez, N., Blakeslee, J. P., Bouwens, R. J., Bradley, L. D., Demarco, R., Homeier, N. L., Martel, A. R., and Mei, S. (2007). Discovery of a Ringlike Dark Matter Structure in the Core of the Galaxy Cluster Cl 0024+17. *ApJ*, 661:728–749.

Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., and Yoshida, N. (2001). The mass function of dark matter haloes. *MNRAS*, 321:372–384.

Jullo, E., Kneib, J., Limousin, M., Elíasdóttir, Á., Marshall, P. J., and Verdugo, T. (2007). A Bayesian approach to strong lensing modelling of galaxy clusters. *New Journal of Physics*, 9:447–+.

Kaiser, N., Aussel, H., Burke, B. E., Boesgaard, H., Chambers, K., Chun, M. R., Heasley, J. N., Hodapp, K., Hunt, B., Jedicke, R., Jewitt, D., Kudritzki, R., Luppino, G. A., Maberry, M., Magnier, E., Monet, D. G., Onaka, P. M., Pickles, A. J., Rhoads, P. H. H., Simon, T., Szalay, A., Szapudi, I., Tholen, D. J., Tonry, J. L., Waterson, M., and Wick, J. (2002). Pan-STARRS: A Large Synoptic Survey Telescope Array. 4836:154–164.

Kaiser, N. and Squires, G. (1993). Mapping the dark matter with weak gravitational lensing. *ApJ*, 404:441–450.

Kaiser, N., Squires, G., and Broadhurst, T. (1995). A Method for Weak Lensing Observations. *ApJ*, 449:460–+.

Kassiola, A. and Kovner, I. (1993). Elliptic Mass Distributions versus Elliptic Potentials in Gravitational Lenses. *ApJ*, 417:450–+.

Kassiola, A., Kovner, I., and Fort, B. (1992). Perturbations of cluster cusps by galaxies - The triple arc in CL 0024+1654. *ApJ*, 400:41–57.

Keeton, C. R. (2003). Analytic Cross Sections for Substructure Lensing. *ApJ*, 584:664–674.

Kirk, D. and Hwu, W. (2010). *Programming Massively Parallel Processors*. Morgan Kaufmann.

Kitching, T. D., Miller, L., Heymans, C. E., van Waerbeke, L., and Heavens, A. F. (2008). Bayesian galaxy shape measurement for weak lensing surveys - II. Application to simulations. *MNRAS*, 390:149–167.

Kneib, J., Ellis, R. S., Smail, I., Couch, W. J., and Sharples, R. M. (1996). Hubble Space Telescope Observations of the Lensing Cluster Abell 2218. *ApJ*, 471:643–+.

Kneib, J., Hudelot, P., Ellis, R. S., Treu, T., Smith, G. P., Marshall, P., Czoske, O., Smail, I., and Natarajan, P. (2003). A Wide-Field Hubble Space Telescope Study of the Cluster Cl 0024+1654 at z=0.4. II. The Cluster Mass Distribution. *ApJ*, 598:804–817.

Koekemoer, A. M., Aussel, H., Calzetti, D., Capak, P., Giavalisco, M., Kneib, J., Leauthaud, A., Le Fèvre, O., McCracken, H. J., Massey, R., Mobasher, B., Rhodes, J., Scoville, N., and Shopbell, P. L. (2007). The COSMOS Survey: Hubble Space Telescope Advanced Camera for Surveys Observations and Data Processing. *ApJS*, 172:196–202.

Komatsu, E., Dunkley, J., Nolta, M. R., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Limon, M., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Meyer, S. S., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2009). Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation. *ApJS*, 180:330–376.

Koo, D. C. (1988). Recent observations of distant matter - Direct clues to birth and evolution. pages 513–540.

Kuijken, K. (2006). Shears from shapelets. *A&A*, 456:827–838.

Lacey, C. and Cole, S. (1993). Merger rates in hierarchical models of galaxy formation. *MNRAS*, 262:627–649.

Lahav, O. and Liddle, A. R. (2010). The Cosmological Parameters 2010. *ArXiv e-prints*, 1002.3488.

Lahav, O., Lilje, P. B., Primack, J. R., and Rees, M. J. (1991). Dynamical effects of the cosmological constant. *MNRAS*, 251:128–136.

Landau, L. and Lifshitz, E. (1987). *Fluid Mechanics, Second Edition: Volume 6 (Course of Theoretical Physics)*. Butterworth-Heinemann.

Lee, J. and Komatsu, E. (2010). Bullet Cluster: A Challenge to LCDM Cosmology. *ArXiv e-prints*, 1003.0939.

Leonard, A., Goldberg, D. M., Haaga, J. L., and Massey, R. (2007). Gravitational Shear, Flexion, and Strong Lensing in Abell 1689. *ApJ*, 666:51–63.

Levine, E. S., Schulz, A. E., and White, M. (2002). Future Galaxy Cluster Surveys: The Effect of Theory Uncertainty on Constraining Cosmological Parameters. *ApJ*, 577:569–578.

Lewis, A. (2009). Galaxy shear estimation from stacked images. *MNRAS*, 398:471–476.

Li, G., Mao, S., Jing, Y. P., Bartelmann, M., Kang, X., and Meneghetti, M. (2005). Is the Number of Giant Arcs in ΛCDM Consistent with Observations? *ApJ*, 635:795–805.

Liddle, A. and Lyth, D. (2000). *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press.

Liesenborgs, J., De Rijcke, S., and Dejonghe, H. (2006). A genetic algorithm for the non-parametric inversion of strong lensing systems. *MNRAS*, 367:1209–1216.

Liesenborgs, J., de Rijcke, S., Dejonghe, H., and Bekaert, P. (2007). Non-parametric inversion of gravitational lensing systems with few images using a multi-objective genetic algorithm. *MNRAS*, 380:1729–1736.

Liesenborgs, J., de Rijcke, S., Dejonghe, H., and Bekaert, P. (2008). A generalization of the mass-sheet degeneracy producing ring-like artefacts in the lens mass distribution. *MNRAS*, 386:307–312.

Limousin, M., Richard, J., Jullo, E., Kneib, J., Fort, B., Soucail, G., Elíasdóttir, Á., Natarajan, P., Ellis, R. S., Smail, I., Czoske, O., Smith, G. P., Hudelot, P., Bardeau, S., Ebeling, H., Egami, E., and Knudsen, K. K. (2007). Combining Strong and Weak Gravitational Lensing in Abell 1689. *ApJ*, 668:643–666.

Limousin, M., Richard, J., Kneib, J., Brink, H., Pelló, R., Jullo, E., Tu, H., Sommer-Larsen, J., Egami, E., Michałowski, M. J., Cabanac, R., and Stark, D. P. (2008). Strong lensing in Abell 1703: constraints on the slope of the inner dark matter distribution. *A&A*, 489:23–35.

Lombardi, M. and Bertin, G. (1998). Improving the accuracy of mass reconstructions from weak lensing: from the shear map to the mass distribution. *A&A*, 335:1–11.

LSST Science Collaborations (2009). LSST Science Book, Version 2.0. *ArXiv e-prints*, 0912.0201.

Luppino, G. A. and Kaiser, N. (1997). Detection of Weak Lensing by a Cluster of Galaxies at Z = 0.83. *ApJ*, 475:20–+.

Mahdavi, A., Hoekstra, H., Babul, A., Balam, D. D., and Capak, P. L. (2007a). A Dark Core in Abell 520. *ApJ*, 668:806–814.

Mahdavi, A., Hoekstra, H., Babul, A., and Henry, J. P. (2008). Evidence for non-hydrostatic gas from the cluster X-ray to lensing mass ratio. *MNRAS*, 384:1567–1574.

Mahdavi, A., Hoekstra, H., Babul, A., Sievers, J., Myers, S. T., and Henry, J. P. (2007b). Joint Analysis of Cluster Observations. I. Mass Profile of Abell 478 from Combined X-Ray, Sunyaev-Zel'dovich, and Weak-Lensing Data. *ApJ*, 664:162–180.

Majumdar, S. and Mohr, J. J. (2003). Importance of Cluster Structural Evolution in Using X-Ray and Sunyaev-Zeldovich Effect Galaxy Cluster Surveys to Study Dark Energy. *ApJ*, 585:603–610.

Majumdar, S. and Mohr, J. J. (2004). Self-Calibration in Cluster Studies of Dark Energy: Combining the Cluster Redshift Distribution, the Power Spectrum, and Mass Measurements. *ApJ*, 613:41–50.

Makino, J. and Taiji, M. (1998). *Scientific Simulations with Special-Purpose Computers - the GRAPE Systems*. John Wiley and Sons.

Massey, R. and Refregier, A. (2005). Polar shapelets. *MNRAS*, 363:197–210.

Massey, R., Refregier, A., Conselice, C. J., David, J., and Bacon, J. (2004). Image simulation with shapelets. *MNRAS*, 348:214–226.

Massey, R., Rhodes, J., Ellis, R., Scoville, N., Leauthaud, A., Finoguenov, A., Capak, P., Bacon, D., Aussel, H., Kneib, J., Koekemoer, A., McCracken, H., Mobasher, B., Pires, S., Refregier, A., Sasaki, S., Starck, J., Taniguchi, Y., Taylor, A., and Taylor, J. (2007). Dark matter maps reveal cosmic scaffolding. *Nature*, 445:286–290.

Mastropietro, C. and Burkert, A. (2008). Simulating the Bullet Cluster. *MNRAS*, 389:967–988.

Matarrese, S. and Pietroni, M. (2007). Resumming cosmic perturbations. *Journal of Cosmology and Astro-Particle Physics*, 6:26–+.

Mazzotta, P., Rasia, E., Moscardini, L., and Tormen, G. (2004). Comparing the temperatures of galaxy clusters from hydrodynamical N-body simulations to Chandra and XMM-Newton observations. *MNRAS*, 354:10–24.

McNamara, B. R. and Nulsen, P. E. J. (2007). Heating Hot Atmospheres with Active Galactic Nuclei. *ARA&A*, 45:117–175.

Medezinski, E., Broadhurst, T., Umetsu, K., Coe, D., Benítez, N., Ford, H., Rephaeli, Y., Arimoto, N., and Kong, X. (2007). Using Weak-Lensing Dilution to Improve Measurements of the Luminous and Dark Matter in A1689. *ApJ*, 663:717–733.

Medezinski, E., Broadhurst, T., Umetsu, K., Oguri, M., Rephaeli, Y., and Benítez, N. (2009). Detailed Cluster Mass and Light profiles of A1703, A370 and RXJ1347-11 from Deep Subaru Imaging. *ArXiv e-prints*, 0906.4791.

Melchior, P. (2010). *Shapelets for gravitational lensing and galaxy morphology studies*. PhD thesis, University of Heidelberg.

Melchior, P., Böhnert, A., Lombardi, M., and Bartelmann, M. (2010). Limitations on shapelet-based weak-lensing measurements. *A&A*, 510:A75+.

Mellier, Y., Fort, B., and Kneib, J. (1993). The dark matter distribution in MS 2137-23 from the modeling of the multiple arc systems. *ApJ*, 407:33–45.

Meneghetti, M. (2006). Introduction to gravitational lensing. Lecture Notes (University of Heidelberg).

Meneghetti, M., Argazzi, R., Pace, F., Moscardini, L., Dolag, K., Bartelmann, M., Li, G., and Oguri, M. (2007a). Arc sensitivity to cluster ellipticity, asymmetries, and substructures. *A&A*, 461:25–38.

Meneghetti, M., Bartelmann, M., Jenkins, A., and Frenk, C. (2007b). The effects of ellipticity and substructure on estimates of cluster density profiles based on lensing and kinematics. *MNRAS*, 381:171–186.

Meneghetti, M., Bartelmann, M., and Moscardini, L. (2003a). cD galaxy contribution to the strong lensing cross-sections of galaxy clusters. *MNRAS*, 346:67–77.

Meneghetti, M., Bartelmann, M., and Moscardini, L. (2003b). Cluster cross-sections for strong lensing: analytic and numerical lens models. *MNRAS*, 340:105–114.

Meneghetti, M., Fedeli, C., Pace, F., Gottloeber, S., and Yepes, G. (2010). Strong lensing in the MareNostrum Universe: biases in the cluster lens population. *ArXiv e-prints*, 1003.4544.

Meneghetti, M., Melchior, P., Grazian, A., De Lucia, G., Dolag, K., Bartelmann, M., Heymans, C., Moscardini, L., and Radovich, M. (2008). Realistic simulations of gravitational lensing by galaxy clusters: extracting arc parameters from mock DUNE images. *A&A*, 482:403–418.

Meneghetti, M., Rasia, E., Merten, J., Bellagamba, F., Ettori, S., Mazzotta, P., and Dolag, K. (2009). Weighing simulated galaxy clusters using lensing and X-ray. *ArXiv e-prints*, 0912.1343.

Merritt, D., Graham, A. W., Moore, B., Diemand, J., and Terzić, B. (2006). Empirical Models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models. *AJ*, 132:2685–2700.

Merten, J. (2009). Sawlens 1.4 - a practical manual. Technical report.

Merten, J., Cacciato, M., Meneghetti, M., Mignone, C., and Bartelmann, M. (2009). Combining weak and strong cluster lensing: applications to simulations and MS 2137. *A&A*, 500:681–691.

Metzler, C. A., White, M., Norman, M., and Loken, C. (1999). Weak Gravitational Lensing and Cluster Mass Estimates. *ApJ*, 520:L9–L12.

Miley, G. K., Overzier, R. A., Zirm, A. W., Ford, H. C., Kurk, J., Pentericci, L., Blakeslee, J. P., Franx, M., Illingworth, G. D., Postman, M., Rosati, P., Röttgering, H. J. A., Venemans, B. P., and Helder, E. (2006). The Spiderweb Galaxy: A Forming Massive Cluster Galaxy at z ˜ 2. *ApJ*, 650:L29–L32.

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., and van Waerbeke, L. (2007). Bayesian galaxy shape measurement for weak lensing surveys - I. Methodology and a fast-fitting algorithm. *MNRAS*, 382:315–324.

Milosavljević, M., Koda, J., Nagai, D., Nakar, E., and Shapiro, P. R. (2007). The Cluster-Merger Shock in 1E 0657-56: Faster than a Speeding Bullet? *ApJ*, 661:L131–L134.

Miralda-Escude, J. (1995). Gravitational lensing by a cluster of galaxies and the central cD galaxy: Measuring the mass profile. *ApJ*, 438:514–526.

Monaghan, J. J. (2005). Smoothed particle hydrodynamics. *Reports on Progress in Physics*, 68(8):1703.

Morandi, A., Ettori, S., and Moscardini, L. (2007). X-ray and Sunyaev-Zel'dovich scaling relations in galaxy clusters. *MNRAS*, 379:518–534.

Morandi, A., Pedersen, K., and Limousin, M. (2010). Unveiling the Three-dimensional Structure of Galaxy Clusters: Resolving the Discrepancy Between X-ray and Lensing Masses. *ApJ*, 713:491–502.

Narayan, R. and Bartelmann, M. (1996). Lectures on Gravitational Lensing. *ArXiv Astrophysics e-prints*, 9606001.

Narayan, R. and Nityananda, R. (1986). Maximum entropy image restoration in astronomy. *ARA&A*, 24:127–170.

Navarro, J. F., Frenk, C. S., and White, S. D. M. (1996). The Structure of Cold Dark Matter Halos. *ApJ*, 462:563–+.

Norman, M. L. and Bryan, G. L. (1999). Cosmological Adaptive Mesh Refinement. In S. M. Miyama, K. Tomisaka, & T. Hanawa, editor, *Numerical Astrophysics*, volume 240 of *Astrophysics and Space Science Library*, pages 19–+.

NVIDIA (2009a). Cuda programming best practices guide. Technical report.

NVIDIA (2009b). Cuda programming guide 2.3.1. Technical report.

NVIDIA (2009c). Cuda reference manual. Technical report.

Oguri, M., Hennawi, J. F., Gladders, M. D., Dahle, H., Natarajan, P., Dalal, N., Koester, B. P., Sharon, K., and Bayliss, M. (2009). Subaru Weak Lensing Measurements of Four Strong Lensing Clusters: Are Lensing Clusters Overconcentrated? *ApJ*, 699:1038–1052.

Oguri, M., Lee, J., and Suto, Y. (2003). Arc Statistics in Triaxial Dark Matter Halos: Testing the Collisionless Cold Dark Matter Paradigm. *ApJ*, 599:7–23.

Oguri, M., Takada, M., Umetsu, K., and Broadhurst, T. (2005). Can the Steep Mass Profile of A1689 Be Explained by a Triaxial Dark Halo? *ApJ*, 632:841–846.

Okabe, N., Takada, M., Umetsu, K., Futamase, T., and Smith, G. P. (2009). LoCuSS: Subaru Weak Lensing Study of 30 Galaxy Clusters. *ArXiv e-prints*, 0903.1103.

Okura, Y., Umetsu, K., and Futamase, T. (2007). A New Measure for Weak-Lensing Flexion. *ApJ*, 660:995–1002.

Okura, Y., Umetsu, K., and Futamase, T. (2008). A Method for Weak-Lensing Flexion Analysis by the HOLICs Moment Approach. *ApJ*, 680:1–16.

Ota, N., Pointecouteau, E., Hattori, M., and Mitsuda, K. (2004). Chandra Analysis and Mass Estimation of the Lensing Cluster of Galaxies Cl 0024+17. *ApJ*, 601:120–132.

Overzier, R. A., Bouwens, R. J., Cross, N. J. G., Venemans, B. P., Miley, G. K., Zirm, A. W., Benítez, N., Blakeslee, J. P., Coe, D., Demarco, R., Ford, H. C., Homeier, N. L., Illingworth, G. D., Kurk, J. D., Martel, A. R., Mei, S., Oliveira, I., Röttgering, H. J. A., Tsvetanov, Z. I., and Zheng, W. (2008). Lyman Break Galaxies, Lyα Emitters, and a Radio Galaxy in a Protocluster at z = 4.1. *ApJ*, 673:143–162.

Padmanabhan, T. (1993). *Structure Formation in the Universe*. Cambridge University Press.

Peebles, P. J. E. and Yu, J. T. (1970). Primeval Adiabatic Perturbation in an Expanding Universe. *ApJ*, 162:815–+.

Percival, W. J., Cole, S., Eisenstein, D. J., Nichol, R. C., Peacock, J. A., Pope, A. C., and Szalay, A. S. (2007). Measuring the Baryon Acoustic Oscillation scale using the Sloan Digital Sky Survey and 2dF Galaxy Redshift Survey. *MNRAS*, 381:1053–1066.

Perkins, D. (1987). *Introduction to High Energy Physics*. Addison Wesley.

Peterson, J. R. and Fabian, A. C. (2006). X-ray spectroscopy of cooling clusters. *Phys. Rep.*, 427:1–39.

Pietroni, M. (2008). Flowing with time: a new approach to non-linear cosmological perturbations. *Journal of Cosmology and Astro-Particle Physics*, 10:36–+.

Plummer, H. C. (1911). On the problem of distribution in globular star clusters. *MNRAS*, 71:460–470.

Popesso, P., Dickinson, M., Nonino, M., Vanzella, E., Daddi, E., Fosbury, R. A. E., Kuntschner, H., Mainieri, V., Cristiani, S., Cesarsky, C., Giavalisco, M., Renzini, A., and GOODS Team (2009). The great observatories origins deep survey. VLT/VIMOS spectroscopy in the GOODS-south field. *A&A*, 494:443–460.

Postman, M., Lubin, L. M., Gunn, J. E., Oke, J. B., Hoessel, J. G., Schneider, D. P., and Christensen, J. A. (1996). The Palomar Distant Clusters Survey. I. The Cluster Catalog. *AJ*, 111:615–+.

Powell, L. C., Kay, S. T., and Babul, A. (2009). The relationship between substructure in 2D X-ray surface brightness images and weak-lensing mass maps of galaxy clusters: a simulation study. *MNRAS*, pages 1472–+.

Pratt, G. W. and Arnaud, M. (2002). The mass profile of <ASTROBJ>A1413</ASTROBJ> observed with XMM-Newton: Implications for the M-T relation. *A&A*, 394:375–393.

Pratt, G. W. and Arnaud, M. (2003). Entropy scaling in galaxy clusters: Insights from an XMM-Newton observation of the poor cluster <ASTROBJ>A1983</ASTROBJ>. *A&A*, 408:1–16.

Pratt, G. W., Croston, J. H., Arnaud, M., and Böhringer, H. (2009). Galaxy cluster X-ray luminosity scaling relations from a representative local sample (REXCESS). *A&A*, 498:361–378.

Press, W. H. and Schechter, P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). Numerical recipes 3rd edition: The art of scientific computing.

Prunet, S., Pichon, C., Aubert, D., Pogosyan, D., Teyssier, R., and Gottloeber, S. (2008). Initial Conditions For Large Cosmological Simulations. *ApJS*, 178:179–188.

Puchwein, E. and Bartelmann, M. (2006). Three-dimensional reconstruction of the intracluster medium. *A&A*, 455:791–801.

Puchwein, E. and Bartelmann, M. (2007). Probing the dynamical state of galaxy clusters. *A&A*, 474:745–754.

Puchwein, E., Bartelmann, M., Dolag, K., and Meneghetti, M. (2005). The impact of gas physics on strong cluster lensing. *A&A*, 442:405–412.

Puchwein, E. and Hilbert, S. (2009). Cluster strong lensing in the Millennium simulation: the effect of galaxies and structures along the line-of-sight. *MNRAS*, 398:1298–1308.

Rannacher, R. (2003a). Analysis i. Lecture Notes (University of Heidelberg).

Rannacher, R. (2003b). Analysis ii. Lecture Notes (University of Heidelberg).

Rannacher, R. (2006). Einfuehrung in die numerische mathematik. Lecture Notes (University of Heidelberg).

Rasia, E., Ettori, S., Moscardini, L., Mazzotta, P., Borgani, S., Dolag, K., Tormen, G., Cheng, L. M., and Diaferio, A. (2006). Systematics in the X-ray cluster mass estimators. *MNRAS*, 369:2013–2024.

Rasia, E., Mazzotta, P., Bourdin, H., Borgani, S., Tornatore, L., Ettori, S., Dolag, K., and Moscardini, L. (2008). X-MAS2: Study Systematics on the ICM Metallicity Measurements. *ApJ*, 674:728–741.

Rasia, E., Tormen, G., and Moscardini, L. (2004). A dynamical model for the distribution of dark matter and gas in galaxy clusters. *MNRAS*, 351:237–252.

Reblinsky, K. and Bartelmann, M. (1999). Projection effects in mass-selected galaxy-cluster samples. *A&A*, 345:1–16.

Refregier, A. (2003). Shapelets - I. A method for image analysis. *MNRAS*, 338:35–47.

Refregier, A., Amara, A., Kitching, T. D., Rassat, A., Scaramella, R., Weller, J., and Euclid Imaging Consortium, f. t. (2010). Euclid Imaging Consortium Science Book. *ArXiv e-prints*, 1001.0061.

Reiprich, T. H. and Böhringer, H. (2002). The Mass Function of an X-Ray Flux-limited Sample of Galaxy Clusters. *ApJ*, 567:716–740.

Ricker, P. M. and Sarazin, C. L. (2001). Off-Axis Cluster Mergers: Effects of a Strongly Peaked Dark Matter Profile. *ApJ*, 561:621–644.

Riemer-Sørensen, S., Paraficz, D., Ferreira, D. D. M., Pedersen, K., Limousin, M., and Dahle, H. (2009). Resolving the Discrepancy Between Lensing and X-Ray Mass Estimates of the Complex Galaxy Cluster Abell 1689. *ApJ*, 693:1570–1578.

Riess, A. G., Macri, L., Casertano, S., Sosey, M., Lampeitl, H., Ferguson, H. C., Filippenko, A. V., Jha, S. W., Li, W., Chornock, R., and Sarkar, D. (2009). A Redetermination of the Hubble Constant with the Hubble Space Telescope from a Differential Distance Ladder. *ApJ*, 699:539–563.

Rines, K., Diaferio, A., and Natarajan, P. (2007). The Virial Mass Function of Nearby SDSS Galaxy Clusters. *ApJ*, 657:183–196.

Rines, K., Geller, M. J., and Diaferio, A. (2010). Virial Masses from the Hectospec Cluster Survey (HeCS) and the Sunyaev-Zeldovich Effect. *ArXiv e-prints*, 1001.0006.

Rix, H., Barden, M., Beckwith, S. V. W., Bell, E. F., Borch, A., Caldwell, J. A. R., Häussler, B., Jahnke, K., Jogee, S., McIntosh, D. H., Meisenheimer, K., Peng, C. Y., Sanchez, S. F., Somerville, R. S., Wisotzki, L., and Wolf, C. (2004). GEMS: Galaxy Evolution from Morphologies and SEDs. *ApJS*, 152:163–173.

Rosati, P., Borgani, S., and Norman, C. (2002). The Evolution of X-ray Clusters of Galaxies. *ARA&A*, 40:539–577.

Rowe, B. (2010). Improving PSF modelling for weak gravitational lensing using new methods in model selection. *MNRAS*, pages 242–+.

Sainio, J. (2009). CUDAEASY - a GPU Accelerated Cosmological Lattice Program. *ArXiv e-prints*, 0911.5692.

Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution. *ApJ*, 121:161–+.

Sand, D. J., Treu, T., and Ellis, R. S. (2002). The Dark Matter Density Profile of the Lensing Cluster MS 2137-23: A Test of the Cold Dark Matter Paradigm. *ApJ*, 574:L129–L133.

Sand, D. J., Treu, T., Ellis, R. S., Smith, G. P., and Kneib, J.-P. (2008). Separating Baryons and Dark Matter in Cluster Cores: A Full Two-dimensional Lensing and Dynamic Analysis of Abell 383 and MS 2137-23. *ApJ*, 674:711–727.

Saro, A., Borgani, S., Tornatore, L., De Lucia, G., Dolag, K., and Murante, G. (2009). Simulating the formation of a protocluster at z ˜ 2. *MNRAS*, 392:795–800.

Saro, A., Borgani, S., Tornatore, L., Dolag, K., Murante, G., Biviano, A., Calura, F., and Charlot, S. (2006). Properties of the galaxy population in hydrodynamical simulations of clusters. *MNRAS*, 373:397–410.

Schechter, P. (1976). An analytic expression for the luminosity function for galaxies. *ApJ*, 203:297–306.

Schmidt, R. W. and Allen, S. W. (2007). The dark matter haloes of massive, relaxed galaxy clusters observed with Chandra. *MNRAS*, 379:209–221.

Schneider, P. and Er, X. (2008). Weak lensing goes bananas: what flexion really measures. *A&A*, 485:363–376.

Schneider, P., Kochanek, C., and Wambsganss, J. (2006). *Gravitational Lensing: Strong, Weak and Micro*. Springer, Berlin.

Schneider, P. and Seitz, C. (1995). Steps towards nonlinear cluster inversion through gravitational distortions. 1: Basic considerations and circular clusters. *A&A*, 294:411–431.

Schrabback, T., Hartlap, J., Joachimi, B., Kilbinger, M., Simon, P., Benabed, K., Bradač, M., Eifler, T., Erben, T., Fassnacht, C. D., High, F. W., Hilbert, S., Hildebrandt, H., Hoekstra, H., Kuijken, K., Marshall, P., Mellier, Y., Morganson, E., Schneider, P., Semboloni, E., Van Waerbeke, L., and Velander, M. (2009). Evidence for the accelerated expansion of the Universe from weak lensing tomography with COSMOS. *ArXiv e-prints*, 0911.0053.

Seitz, C. and Schneider, P. (1995). Steps towards nonlinear cluster inversion through gravitational distortions II. Generalization of the Kaiser and Squires method. *A&A*, 297:287–+.

Seitz, S., Schneider, P., and Bartelmann, M. (1998). Entropy-regularized maximum-likelihood cluster mass reconstruction. *A&A*, 337:325–337.

Semboloni, E., Mellier, Y., van Waerbeke, L., Hoekstra, H., Tereno, I., Benabed, K., Gwyn, S. D. J., Fu, L., Hudson, M. J., Maoli, R., and Parker, L. C. (2006). Cosmic shear analysis with CFHTLS deep data. *A&A*, 452:51–61.

Sérsic, J. L. (1963). Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy. *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, 6:41–+.

Sheth, R. K., Mo, H. J., and Tormen, G. (2001). Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *MNRAS*, 323:1–12.

Sheth, R. K. and Tormen, G. (2002). An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. *MNRAS*, 329:61–75.

Shimon, M. and Rephaeli, Y. (2004). Quantitative description of the Sunyaev-Zeldovich effect: analytic approximations. *New Astronomy*, 9:69–82.

Shu, C., Zhou, B., Bartelmann, M., Comerford, J. M., Huang, J., and Mellier, Y. (2008). Comparisons between Isothermal and NFW Mass Profiles for Strong-Lensing Galaxy Clusters. *ApJ*, 685:70–82.

Sirko, E. (2005). Initial Conditions to Cosmological N-Body Simulations, or, How to Run an Ensemble of Simulations. *ApJ*, 634:728–743.

Smail, I., Dressler, A., Kneib, J., Ellis, R. S., Couch, W. J., Sharples, R. M., and Oemler, Jr., A. (1996). Hubble Space Telescope Observations of Giant Arcs: High-Resolution Imaging of Distant Field Galaxies. *ApJ*, 469:508–+.

Sofue, Y. and Rubin, V. (2001). Rotation Curves of Spiral Galaxies. *ARA&A*, 39:137–174.

Somerville, R. S. and Kolatt, T. S. (1999). How to plant a merger tree. *MNRAS*, 305:1–14.

Spergel, D. N., Bean, R., Doré, O., Nolta, M. R., Bennett, C. L., Dunkley, J., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H. V., Verde, L., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2007). Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. *ApJS*, 170:377–408.

Springel, V. (2005). The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134.

Springel, V. (2010). E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *MNRAS*, 401:791–851.

Springel, V. and Farrar, G. R. (2007). The speed of the 'bullet' in the merging galaxy cluster 1E0657-56. *MNRAS*, 380:911–925.

Springel, V. and Hernquist, L. (2003). Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. *MNRAS*, 339:289–311.

Springel, V., Wang, J., Vogelsberger, M., Ludlow, A., Jenkins, A., Helmi, A., Navarro, J. F., Frenk, C. S., and White, S. D. M. (2008). The Aquarius Project: the subhaloes of galactic haloes. *MNRAS*, 391:1685–1711.

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636.

Springel, V., White, S. D. M., Tormen, G., and Kauffmann, G. (2001). Populating a cluster of galaxies - I. Results at z=0. *MNRAS*, 328:726–750.

Spurzem, R., Berentzen, I., Berczik, P., Merritt, D., Amaro-Seoane, P., Harfst, S., and Gualandris, A. (2008). Parallelization, Special Hardware and Post-Newtonian Dynamics in Direct N - Body Simulations. *Lecture Notes in Physics, Berlin Springer Verlag*, 760:377–+.

Srednicki, M. (2007). *Quantum Field Theory*. Cambridge University Press.

Steigman, G. (2007). Primordial Nucleosynthesis in the Precision Cosmology Era. *Annual Review of Nuclear and Particle Science*, 57:463–491.

Stocke, J. T., Morris, S. L., Gioia, I. M., Maccacaro, T., Schild, R., Wolter, A., Fleming, T. A., and Henry, J. P. (1991). The Einstein Observatory Extended Medium-Sensitivity Survey. II - The optical identifications. *ApJS*, 76:813–874.

Straumann, N. (1997). Complex Formulation of Lensing Theory and Applications. *ArXiv Astrophysics e-prints*, arXiv:astro-ph/9703103.

Sun, M., Voit, G. M., Donahue, M., Jones, C., Forman, W., and Vikhlinin, A. (2009). Chandra Studies of the X-Ray Gas Properties of Galaxy Groups. *ApJ*, 693:1142–1172.

Sunyaev, R. A. and Zeldovich, Y. B. (1970). Small-Scale Fluctuations of Relic Radiation. *Ap&SS*, 7:3–19.

Sunyaev, R. A. and Zeldovich, Y. B. (1972). The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies. *Comments on Astrophysics and Space Physics*, 4:173–+.

Tinker, J., Kravtsov, A. V., Klypin, A., Abazajian, K., Warren, M., Yepes, G., Gottlöber, S., and Holz, D. E. (2008). Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. *ApJ*, 688:709–728.

Tormen, G., Bouchet, F. R., and White, S. D. M. (1997). The structure and dynamical evolution of dark matter haloes. *MNRAS*, 286:865–884.

Torri, E., Meneghetti, M., Bartelmann, M., Moscardini, L., Rasia, E., and Tormen, G. (2004). The impact of cluster mergers on arc statistics. *MNRAS*, 349:476–490.

Tung, W.-K. (1985). *Group Theory In Physics: An Introduction to Symmetry Principles, Group Representations, and Special Functions in Classical and Quantum Physics*. World Scientific Publishing Co Pte Ltd.

Umetsu, K., Birkinshaw, M., Liu, G., Wu, J., Medezinski, E., Broadhurst, T., Lemze, D., Zitrin, A., Ho, P. T. P., Huang, C., Koch, P. M., Liao, Y., Lin, K., Molnar, S. M., Nishioka, H., Wang, F., Altamirano, P., Chang, C., Chang, S., Chang, S., Chen, M., Han, C., Huang, Y., Hwang, Y., Jiang, H., Kesteven, M., Kubo, D. Y., Li, C., Martin-Cocher, P., Oshiro, P., Raffin, P., Wei, T., and Wilson, W. (2009a). Mass and Hot Baryons in Massive Galaxy Clusters from Subaru Weak-Lensing and AMiBA Sunyaev-Zel'Dovich Effect Observations. *ApJ*, 694:1643–1663.

Umetsu, K. and Broadhurst, T. (2008). Combining Lens Distortion and Depletion to Map the Mass Distribution of A1689. *ApJ*, 684:177–203.

Umetsu, K., Medezinski, E., Broadhurst, T., Zitrin, A., Okabe, N., Hsieh, B., and Molnar, S. M. (2009b). The Mass Structure of the Galaxy Cluster Cl0024+1654 from a Full Lensing Analysis of Joint Subaru and ACS/NIC3 Observations. *ArXiv e-prints*, 0908.0069.

van Waerbeke, L. (2000). Noise properties of gravitational lens mass reconstruction. *MNRAS*, 313:524–532.

van Waerbeke, L. (2010). Shear and magnification: cosmic complementarity. *MNRAS*, 401:2093–2100.

Vanderlinde, K., Crawford, T. M., de Haan, T., Dudley, J. P., Shaw, L., Ade, P. A. R., Aird, K. A., Benson, B. A., Bleem, L. E., Brodwin, M., Carlstrom, J. E., Chang, C. L., Crites, A. T., Desai, S., Dobbs, M. A., Foley, R. J., George, E. M., Gladders, M. D., Hall, N. R., Halverson, N. W., High, F. W., Holder, G. P., Holzapfel, W. L., Hrubes, J. D., Joy, M., Keisler, R., Knox, L., Lee, A. T., Leitch, E. M., Loehr, A., Lueker, M., Marrone, D. P., McMahon, J. J., Mehl, J., Meyer, S. S., Mohr, J. J., Montroy, T. E., Ngeow, C., Padin, S., Plagge, T., Pryke, C., Reichardt, C. L., Rest, A., Ruel, J., Ruhl, J. E., Schaffer, K. K., Shirokoff, E., Song, J., Spieler, H. G., Stalder, B., Staniszewski, Z., Stark, A. A., Stubbs, C. W., van Engelen, A., Vieira, J. D., Williamson, R., Yang, Y., Zahn, O., and Zenteno, A. (2010). Galaxy Clusters Selected with the Sunyaev-Zel'dovich Effect from 2008 South Pole Telescope Observations. *ArXiv e-prints*, 1003.0003.

Vanzella, E., Cristiani, S., Dickinson, M., Giavalisco, M., Kuntschner, H., Haase, J., Nonino, M., Rosati, P., Cesarsky, C., Ferguson, H. C., Fosbury, R. A. E., Grazian, A., Moustakas, L. A., Rettura, A., Popesso, P., Renzini, A., Stern, D., and GOODS Team (2008). The great observatories origins deep survey. VLT/FORS2 spectroscopy in the GOODS-South field: Part III. *A&A*, 478:83–92.

Varshalovich, D. A. and Moskalev, A. N. (1988). *Quantum theory of angular momentum*. World Scientific Publishing Company.

Verde, L. (2009). Statistical methods in cosmology. *ArXiv e-prints*, 0911.3105.

Vikhlinin, A., Kravtsov, A., Forman, W., Jones, C., Markevitch, M., Murray, S. S., and Van Speybroeck, L. (2006). Chandra Sample of Nearby Relaxed Galaxy Clusters: Mass, Gas Fraction, and Mass-Temperature Relation. *ApJ*, 640:691–709.

Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., Ebeling, H., Forman, W. R., Hornstrup, A., Jones, C., Murray, S. S., Nagai, D., Quintana, H., and Voevodkin, A. (2009). Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints. *ApJ*, 692:1060–1074.

Voit, G. M. (2005). Tracing cosmic evolution with clusters of galaxies. *Reviews of Modern Physics*, 77:207–258.

Walsh, D., Carswell, R. F., and Weymann, R. J. (1979). 0957 + 561 A, B - Twin quasistellar objects or gravitational lens. *Nature*, 279:381–384.

Wambsganss, J., Bode, P., and Ostriker, J. P. (2004). Giant Arc Statistics in Concord with a Concordance Lambda Cold Dark Matter Universe. *ApJ*, 606:L93–L96.

Wang, P., Abel, T., and Kaehler, R. (2009). Adaptive Mesh Fluid Simulations on GPU. *ArXiv e-prints*, 0910.5547.

Warren, M. S., Abazajian, K., Holz, D. E., and Teodoro, L. (2006). Precision Determination of the Mass Function of Dark Matter Halos. *ApJ*, 646:881–885.

Weinberg, S. (2008). *Cosmology*. Oxford University Press.

Werner, N., Durret, F., Ohashi, T., Schindler, S., and Wiersma, R. P. C. (2008). Observations of Metals in the Intra-Cluster Medium. *Space Science Reviews*, 134:337–362.

White, M. and Vale, C. (2004). Simulations of weak gravitational lensing. *Astroparticle Physics*, 22:19–27.

Yee, H. K. C. and López-Cruz, O. (1999). A Quantitative Measure of the Richness of Galaxy Clusters. *AJ*, 117:1985–1994.

Yoshida, N., Sheth, R. K., and Diaferio, A. (2001). Non-Gaussian cosmic microwave background temperature fluctuations from peculiar velocities of clusters. *MNRAS*, 328:669–677.

Zel'dovich, Y. B. (1970). Gravitational instability: An approximate theory for large density perturbations. *A&A*, 5:84–89.

Zel'dovich, Y. B. (1972). A hypothesis, unifying the structure and the entropy of the Universe. *MNRAS*, 160:1P–+.

Zhang, Y., Böhringer, H., Mellier, Y., Soucail, G., and Forman, W. (2005). XMM-Newton study of the lensing cluster of galaxies CL 0024+17. *A&A*, 429:85–99.

Zhang, Y., Okabe, N., Finoguenov, A., Smith, G. P., Piffaretti, R., Valdarnini, R., Babul, A., Evrard, A. E., Mazzotta, P., Sanderson, A. J. R., and Marrone, D. P. (2010). LoCuSS: A Comparison of Cluster Mass Measurements from XMM-Newton and Subaru - Testing Deviation from Hydrostatic Equilibrium and Non-thermal Pressure Support. *ApJ*, 711:1033–1043.

Zitrin, A., Broadhurst, T., Barkana, R., Rephaeli, Y., and Benitez, N. (2010). Strong-Lensing Analysis of a Complete Sample of 12 MACS Clusters at z > 0.5: Mass Models and Einstein Radii. *ArXiv e-prints*, 1002.0521.

Zitrin, A., Broadhurst, T., Umetsu, K., Coe, D., Benítez, N., Ascaso, B., Bradley, L., Ford, H., Jee, J., Medezinski, E., Rephaeli, Y., and Zheng, W. (2009). New multiply-lensed galaxies identified in ACS/NIC3 observations of Cl0024+1654 using an improved mass model. *MNRAS*, 396:1985–2002.

Zu Hone, J. A., Lamb, D. Q., and Ricker, P. M. (2009a). Rings of Dark Matter in Collisions Between Clusters of Galaxies. *ApJ*, 696:694–700.

Zu Hone, J. A., Ricker, P. M., Lamb, D. Q., and Karen Yang, H. (2009b). A Line-of-Sight Galaxy Cluster Collision: Simulated X-Ray Observations. *ApJ*, 699:1004–1014.

Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127.

Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae. *ApJ*, 86:217–+.

Zwicky, F. (1959). Clusters of Galaxies. *Handbuch der Physik*, 53:390–+.

Index