Projection effects in clusters of galaxies

Dissertation der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von Katrin Reblinsky
aus Lübeck

München, den 17. Januar 2000

1. Gutachter: Priv.-Doz. Dr. Matthias Bartelmann

2. Gutachter: Prof. Dr. Harald Lesch

Tag der mündlichen Prüfung: 9. März 2000

Meiner Mutter

Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

c	speed of light ($c = 299792458\ \mathrm{m\,s^{-1}}$)
G	Newtonian gravitational constant ($G = 6.67259 \times 10^{-11}\ \mathrm{m^3 kg^{-1} s^{-2}}$)
H_0	Hubble constant ($H_0 = 100h\ \mathrm{km\,s^{-1} Mpc^{-1}}$)
h	normalized Hubble constant ($0.5 < h < 1$)
$L_\odot$	solar luminosity ($L_\odot = 3.846 \times 10^{26}\ \mathrm{W}$)
$M_\odot$	solar mass ($M_\odot = 1.98892 \times 10^{24}\ \mathrm{kg}$)

CDM	cold dark matter
CMB	cosmic microwave background
EdS	Einstein–de–Sitter
FLRW	Friedmann–Lemaître–Robertson–Walker
HDM	hot dark matter
los	line–of–sight
LSS	large scale structure
MDRL	multiple data Richardson–Lucy
Mpc	Megaparsec ($1\ \mathrm{Mpc} = 3.0856775807 \times 10^{22}\ \mathrm{m}$)
NFW	Navarro–Frenk–White
PDF	probability distribution function
rms	root mean square
SZ	Sunyaev–Zeldovich
VT	(kinematical) virial theorem

Einleitung und Zusammenfassung

Während des letzten Jahrzehnts hat sich die Untersuchung der Bildung und Entwicklung großskaliger Strukturen in unserem Universum zu einem der interessantesten Forschungszweige in der Kosmologie entwickelt (Padmanabhan 1993; Silk 1994, Longair 1996). In engem Zusammenhang damit steht auch die Frage der Entstehung und Entwicklung von Galaxienhaufen, die die größten gravitativ gebundenen Objekte im Universum sind (Rood 1981; Oort 1983; Bahcall 1988; Sarazin 1986, 1988). Da dunkle Materie vermutlich die beherrschende Rolle auf den hier angesprochenen großen Längenskalen spielt (Trimble 1987; Kormendy & Knapp 1987; Sciama 1993), ist der Gravitationslinseneffekt nicht nur hervorragend zur Untersuchung der großräumigen Struktur des Universums selbst geeignet, sondern auch zum Studium einzelner Galaxienhaufen (Blandford & Narayan 1992; Fort & Mellier 1994). In diesem Zusammenhang bietet insbesondere der schwache Gravitationslinseneffekt die Möglichkeit, die Verteilung der dunklen Materie zu kartieren, ohne dabei Annahmen über ihren dynamischen Zustand machen zu müssen (Bartelmann & Schneider 1999).

Der schwache Gravitationslinseneffekt beschreibt die beobachtbare, geringfügige Verzerrung der Bilder entfernter Hintergrundgalaxien durch eine zwischen dem Beobachter und den Galaxien liegende Materiekonzentration (Bartelmann & Schneider 1999). Dies impliziert, daß der schwache Gravitationslinseneffekt ausschließlich von der projizierten Materie zwischen Hintergrundgalaxien und Beobachter abhängt. Deshalb ist es möglich, durch die Analyse der verzerrten Bilder Informationen über die "Linsen" selbst, also über die dazwischenliegende Materieverteilung zu erhalten (Gunn 1967a,b; Blandford et al. 1991; Miralda-Escude 1991; Kaiser 1992; Schneider et al. 1998).

Werden nun Galaxienhaufen als Linsen interpretiert, so stellt sich die Frage, ob sie sich nicht nur durch das schwache Linsensignal detektieren lassen, welche sie auf die Bilder der im Hintergrund verteilten Galaxien aufprägen, was letztendlich einer Auswahl über ihre Masse gleichkäme (Fahlmann et al. 1994; Schneider 1996). Dies wäre ein erster, wichtiger Schritt im Hinblick auf die Konstruktion einer Massenfunktion über den Gravitationslinseneffekt, aus welcher weitere Einschränkungen der kosmologischen Parameter, insbesondere für den Bereich höherer Rotverschiebungen $z \approx 0.4 - 0.6$, gewonnen werden könnten. Unter diesem Gesichtspunkt wäre eine solche Gravitationslinsen-basierte Massenfunktion eine ideale Ergänzung zu bereits bekannten Massenfunktionen, die über eine Auswahl im optischen oder Röntgenbereich gewonnen werden, und die anderen systematischen Fehlern unterliegen (Frenk et al. 1990; Bartelmann & Steinmetz 1996; Cen 1997; van Haarlem et al. 1997). Allerdings ist es wegen der über einen großen Rotverschiebungsbereich erfolgenden Beiträge zum Linsensignal notwendig, die Kontamination von

solchen massenselektierten Galaxienkatalogen aufgrund von Projektionseffekten zu untersuchen.

Diese Überlegungen haben mich veranlaßt, im Rahmen meiner Promotion die Detektion und die Auswahl von Galaxienhaufen ausschließlich über ihre Masse zu untersuchen. Dabei wurde von mir erstmalig eine neue Auswahlmethode — die Aperturmasse M_{ap} (Schneider 1996) — angewendet, die allein auf dem schwachen Gravitationslinseneffekt basiert.

Ein Vergleich mit anderen bekannten Auswahlverfahren für Galaxienhaufen, wie z.B. dem optischen Abell-Kriterium (Abell 1958), zeigt, daß die Auswahl vermittels des schwachen Gravitationslinseneffekts deutlich geringer durch Projektionseffekte gestört wird, sowohl was die Vollständigkeit der erstellten Kataloge betrifft, als auch die Anzahl der Fehldetektionen (Reblinsky & Bartelmann 1999a). Eine weitere in diesem Zusammenhang interessante Frage ist die nach der Genauigkeit der verschiedenen bekannten Schätzer für die Massen der detektierten Galaxienhaufen. Ich habe den auf dem kinematischen Virialtheorem basierenden Massenschätzer mit dem über den Gravitationslinseneffekt abgeleiteten Schätzer, der ζ-Statistik, verglichen, wobei sich herausstellte, daß der Gravitationslinsen-basierte Schätzer eine erheblich geringere Dispersion hat. Dieses Ergebnis ist von großer Bedeutung, vor allem für die verschiedenen geplanten Himmelsdurchmusterungen, MEGACAM (Mellier et al. 1999) sei hier als Beispiel genannt, in denen Objekte ausschließlich über ihr Linsensignal detektiert und kartiert werden sollen.

Die bereits eingangs erwähnte Aperturmasse M_{ap} ist nicht nur für die Detektion von Galaxienhaufen unabhängig von ihrer Leuchtkraft geeignet, sondern auch für eine Untersuchung der kosmischen Scherung, also der kohärenten Verzerrung der weit entfernten Hintergrundgalaxien durch eine dazwischenliegende großskalige Massenverteilung, welche dafür letztendlich wieder selbst als eine Linse betrachtet wird, und so eine Untersuchung der statistischen Eigenschaften des zugrundeliegenden Dichtefeldes ermöglicht. Mit Hilfe von M_{ap} habe ich die kosmische Scherung in großvolumigen, hochaufgelösten kosmologischen Simulationen studiert und konnte dabei zeigen, daß die Aperturmasse sinnvolle Aussagen über großskalige Strukturen ermöglicht. Dabei habe ich mir zunutze gemacht, daß es sich bei der Aperturmasse ganz im Gegensatz zu den bisher in Untersuchungen zur kosmischen Scherung herangezogenen Galaxienelliptizitäten um eine rein skalare Größe handelt. Diese Eigenschaft von M_{ap} vereinfacht enorm die Definition der höheren Momente wie Schiefe und Kurtosis, die, wie ich zeigen konnte, in der Tat von dem zugrundeliegenden kosmologischen Modell abhängen. Schließlich konnte ich mit Hilfe der numerischen Simulationsdaten die auf einem semianalytischen Press-Schechter-Formalismus basierende Vermutung von Kruse & Schneider (1999) bestätigen (Reblinsky et al. 1999), daß der Schwanz der Verteilungsfunktion von M_{ap} ein exponentielles Abklingverhalten zeigt, das durch die schon kollabierten Halos dominiert wird, und somit ebenfalls wichtige kosmologische Informationen enthält.

Seit geraumer Zeit wächst die Anzahl der Galaxienhaufen stetig an, für die Observationsdaten zum schwachen Gravitationslinseneffekt verfügbar sind, was ebenso für Beobachtungen im Röntgenbereich und für den Sunyaev-Zel'dovich Effekt gilt. Für den Röntgenbereich gilt dies insbesondere seit dem Start der eigens zu diesem Zweck entworfenen Satelliten Chandra und XMM im letzten Jahr. Angeregt durch einen Vorschlag meines Betreuers, M. Bartelmann, habe ich mit der Arbeit an einer neuen

Methode begonnen, in der Beobachtungsdaten aus dem schwachen Gravitationslinsen-effekt (das Linsenpotential ψ), dem Röntgenbereich (die Röntgenleuchtkraft S_x) und aus dem Sunyaev-Zel'dovich Effekt (Temperaturdekrement ΔT_{SZ}) kombiniert werden sollen, um die Struktur einzelner Galaxienhaufen entlang der Beobachtungssichtlinie zu rekonstruieren.

Der Ansatz dieser Deprojektionsmethode basiert auf dem bekannten Dekonvolu-tionsalgorithmus von Richardson und Lucy (1974), der in der optischen Astronomie häufig für die Bildrekonstruktion eingesetzt wird. Da der Richardson-Lucy Algorith-mus statistischer Natur ist, eignet er sich besonders für die Rekonstruktion verrauschter Daten; allerdings ist er bisher weder im Zusammenhang mit dem Gravitationslinsen-effekt, noch in der Kombination von Daten aus Linsen-, Röntgen-, oder Sunyaev-Zel'dovich-Beobachtungen benutzt worden. Ausgehend von geeigneten Modellan-nahmen, in denen alle diese drei Observablen ψ, S_x, und ΔT_{SZ} eine funktionale Abhängigkeit vom Gravitationspotential φ entlang der Sichtlinie haben, gelang es mir, einen sogenannten *multiple-data Richardson-Lucy* (MDRL)-Algorithmus abzuleiten, der auf dem Richardson-Lucy Verfahren von Binney et al. (1990) zur Deprojektion von elliptischen Galaxien aus photometrischen Daten beruht.

Eine erste Implementation der MDRL Deprojektionsmethode habe ich mittlerweile erfolgreich auf synthetische Eingabedaten angewandt, die aus gasdynamischen Simula-tionen generiert wurden (Reblinsky & Bartelmann 1999b, 2000). Die Rekonstruktions-methode funktioniert schon für den Spezialfall der Rekonstruktion aus einem der drei Eingabedatensätze sehr gut, und die Kombination aller drei Eingabedatensätze führt tatsächlich zu einer nochmaligen Verbesserung.

Die vorliegende Dissertation ist wie folgt gegliedert: In den ersten beiden Kapiteln wird der nötige theoretische Rahmen dargelegt: Das Standardbild der Strukturent-stehung und –entwicklung in Kapitel 1 und alle wichtigen Teile der Gravitationslin-sentheorie in Kapitel 2, insbesondere wird dort die M_{ap} Statistik eingeführt. In Kapitel 3 wird die M_{ap} Statistik benutzt, um die statistischen Eigenschaften der kosmischen Scherung mit Hilfe großer, hochaufgelöster Simulationen zu untersuchen. Außerdem wird die Anzahl der Halos, die in diesen Simulationen gefunden wurden, mit der aus semianalytischen Abschätzungen gewonnenen Anzahl verglichen. In Kapitel 4 wird der mögliche Einsatz der M_{ap} Statistik für die massenbasierte Auswahl von Galaxien-haufen im Hinblick auf Projektionseffekte untersucht und mit dem optischen Abell–Kriterium verglichen. In Kapitel 5 wird dann der neue Richardson–Lucy Rekonstruk-tionsalgorithmus abgeleitet und mit synthetischen Daten getestet. Schließlich werden in Kapitel 6 die Hauptergebnisse dieser Arbeit zusammengefaßt und in einen kosmolo-gischen Zusammenhang gestellt. Ferner wird ein Ausblick auf weitere Forschungspro-jekte gegeben.

Introduction and Summary

Over the last decade the formation and evolution of large scale structure (LSS) emerged as one of the most interesting branches of cosmology (Padmanabhan 1993; Silk 1994; Longair 1996). Closely related to this topic is the formation and evolution of galaxy clusters as the largest gravitationally bound objects in the universe (Rood 1981; Oort 1983; Bahcall 1988; Sarazin 1986, 1988). Since dark matter is believed to play the dominant role on such large scales (Trimble 1987; Kormendy & Knapp 1987; Sciama 1993), gravitational lensing is a superb tool for investigating questions related to LSS and galaxy clusters (Blandford & Narayan 1992; Fort & Mellier 1994). In this context weak gravitational lensing offers the possibility to map the dark matter component without prior assumptions about the dynamical state of the matter (Bartelmann & Schneider 1999).

Weak gravitational lensing describes the weak distortions of the images of faint background galaxies caused by any intervening mass concentrations (Bartelmann & Schneider 1999). This means that weak lensing is sensitive only to the projected matter between the background galaxies and the observer. In this way the analysis of the distorted background galaxies allows one to obtain information about the "lenses", namely the matter in between (Gunn 1967a,b; Blandford et al. 1991; Miralda–Escude 1991; Kaiser 1992; Schneider et al. 1998).

Considering clusters of galaxies as lenses, the question arises as to whether it is possible to detect these clusters solely by their weak gravitational lensing imprint on the background population, thus in practice selecting them by mass only (Fahlmann et al. 1994; Schneider 1996). This would be an important first step in constructing a cluster mass function from weak lensing, which could provide additional cosmological constraints, especially at higher redshifts of $z \approx 0.4 - 0.6$. In this sense such a weak–lensing–based mass function would ideally supplement mass functions constructed from X–ray selections or optical selections, which are hampered by different systematic errors (Frenk et al. 1990; Bartelmann & Steinmetz 1996; Cen 1997; van Haarlem et al. 1997). However, because of the very broad contribution to the lensing signal in redshift space, it is necessary to address the important question of contaminations introduced by projection effects into such cluster catalogues.

With this motivation in mind as part of my Ph.D. work I investigated the detection and selection of galaxy clusters solely by their mass using for the first time a selection method which is based on the weak lensing effect alone — the second–order aperture mass measure M_{ap} proposed by Schneider in 1996.

A comparison with other common selection methods for galaxy clusters, such as the optical Abell criterion (Abell 1958), showed that the weak-lensing-based selection is evidently much less affected by projection effects. I was able to demonstrate that the

selection of galaxy clusters according to their mass indeed leads to more reliable cluster catalogues than optical selection, both with regards to the completeness of the catalogue and the number of spurious detections (Reblinsky & Bartelmann 1999a). It also proved valuable to study different mass estimates and their reliability. Comparing the kinematical virial theorem (VT)-based mass estimate to the weak-lensing-based mass estimate, the ζ-statistics, I found that the latter has a substantially smaller dispersion. This is an important result, especially in the light of several wide-field lensing surveys, with MEGACAM (Mellier et al. 1999) being only one example, which try to map and detect objects by their weak gravitational lensing effects alone.

The above-mentioned aperture mass measure M_{ap} is not only a useful tool for the detection of clusters of galaxies irrespective of their luminous properties; it can also be used to investigate cosmic shear, i.e. the coherent distortion of faint background galaxies due to the intervening large scale mass distribution. Ultimately, in this case the projected density field of the LSS is considered as a lens. This is a means to investigate the statistical properties of the underlying density field. I studied the cosmic shear as obtained from large ray-tracing simulations of weak lensing, which in turn were computed from large N-body simulations. I was able to demonstrate the usefulness of the aperture mass measure M_{ap} in the context of large scale structure. A further advantage of the aperture mass measure is the fact that it is a scalar quantity, which is not the case for the mean galaxy ellipticity commonly used to map cosmic shear. This property of M_{ap} tremendously simplifies the definition and evaluation of higher order moments like skewness and kurtosis, which do indeed, as I was able to show, depend on the underlying cosmology. In addition, I could also demonstrate that the far tail of the probability distribution function of M_{ap} has an exponential decline dominated by the already collapsed haloes, thus by itself containing interesting information about cosmology. Finally, in the course of this study the semianalytical estimates by Kruse & Schneider (1999) for M_{ap} based on Press-Schechter-type arguments could be validated (Reblinsky et al. 1999).

The number of clusters for which a weak-lensing analysis is available is steadily increasing, and more and more data from Sunyaev–Zel'dovich and X-ray based observations for galaxy clusters become available. This applies in particular to X-ray data with the projected launches of *Chandra* and *XMM* in 1999. Motivated by a suggestion of my supervisor, M. Bartelmann, I started to work on a novel deprojection method for clusters of galaxies, in which the structure of individual clusters along the line–of–sight, namely the 3–dimensional gravitational potential φ, is reconstructed by combining data from three sources: weak lensing (lensing potential ψ), X-ray (X-ray surface brightness S_{x}, and the Sunyaev-Zel'dovich effect (temperature decrement ΔT_{SZ}).

The ansatz for our deprojection method is based on the Richardson–Lucy (1974) deconvolution algorithm, which is well known in optical astronomy, where it is used for image recovery. As the Richardson–Lucy algorithm is statistical in nature, it is well suited for the rectification of noisy data. However, it has never before been used in the context of weak lensing, let alone a combination of lensing data with X-ray or Sunyaev-Zel'dovich data. Exploiting the fact that all of these three observable distributions do have, within a suitable model, a functional dependence on the gravitational potential along the line–of–sight, a formulation for a *multiple–data Richardson–Lucy* (MDRL) algorithm was derived, which is based on the earlier deprojection method for elliptical galaxies from photometric data by Binney et al. in 1990.

A first implementation of this MDRL deprojection method was successfully applied (Reblinsky & Bartelmann 1999b, 2000) to synthetic input data created from gas–dynamical simulations. The reconstruction method not only works surprisingly well if it is used for each of the three types of input data separately, where it is able to recover all important features of the original 3-dimensional cluster, but also the combination of the different data sets was found to improve upon the deficiencies of the single-data reconstructions.

This dissertation is organized as follows. In the first two chapters the necessary theoretical framework is briefly reviewed: the standard picture of structure formation and evolution in chapter 1 and the relevant details about gravitational lensing in chapter 2, where especially the M_{ap} statistics is introduced. The M_{ap} statistics is then used in chapter 3 to investigate the statistical properties of cosmic shear using LSS ray tracing simulations from numerical N–body data. In addition, the number of haloes computed from semianalytical estimates are compared to results obtained from the ray tracing simulations. In chapter 4 the possible use of the M_{ap} statistic for a mass based selection of clusters of galaxies is analyzed with respect to projection effects and compared to the optical Abell selection criterion. Then, in chapter 5 the novel multiple–data Richardson–Lucy cluster reconstruction algorithm is derived and tested using synthetical data generated from gas-dynamical simulations. Finally, chapter 6 discusses the main results obtained in this thesis, puts them into a broader cosmological context and proposes future research.

Chapter 1

Standard Cosmology and Structure Formation

In this chapter I briefly introduce those aspects of cosmology relevant to this thesis. The standard model of cosmology (Weinberg, 1972; Misner, Thorne and Wheeler, 1973; Sexl and Urbantke, 1995) consists of two major parts: On large scales, i.e. regions comparable to the Hubble volume, whose size is characterized by $cH_0^{-1} = 9.25 \times 10^{27}$ cm $\approx 3000h^{-1}$Mpc [1], the Universe is assumed to be homogeneous and isotropic. This assumption is supported by several independent observations: the isotropy of the cosmic microwave background (CMB) (Banday A.J., Gorski K.M., Bennett C.L., et al., 1997), of the X-ray background (Treyer M., Scharf C., Lahav O., et al., 1998), and in the number counts of distant radio sources (Gregory P.C. & Condon J.J., 1991). Therefore it can be described by the homogeneous and isotropic solutions of the field equations of General Relativity, namely by the Friedmann-Lemaître cosmological models.

On much smaller scales, i.e. on scales of galaxies and clusters of galaxies, the Universe appears to be very lumpy. In order to understand this structure and its evolution, we need a model for structure formation (Padmanabhan, 1993; Peebles, 1993; Kolb and Turner, 1994; Peacock, 1999). The current picture of structure formation assumes that structures grow via gravitational instability from initial perturbations of the mass density. Most of the current theories model the initial seed field as a Gaussian random field.

1.1 Friedmann–Lemaître–Robertson–Walker Models

1.1.1 Kinematics and Geometry

The metric of a space–time with homogeneous and isotropic spatial sections is described by the maximally symmetric Robertson–Walker (RW)–metric

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\mathrm{d}\sigma^2 \tag{1.1}$$

---

[1] h is the normalized Hubble constant, $0.5 < h < 1$

where $a(t)$ is the cosmic scale function, while $\mathrm{d}\sigma^2$ is the line element of a homogeneous and isotropic three–space. The most general form of the line element can be written as

$$\mathrm{d}\sigma^2 = \mathrm{d}w^2 + f_K^2(w)\left(\mathrm{d}\phi^2 + \sin^2(\theta)\,\mathrm{d}\theta^2\right) \equiv \mathrm{d}w^2 + f_K^2(w)\mathrm{d}\omega^2. \quad (1.2)$$

The radial function $f_K(w)$ is linked to the spatial hypersurfaces of the Universe: it must be either trigonometric, linear, or hyperbolic for a positive, zero, or negative curvature constant K, respectively. Specifically,

$$f_K(w) = \begin{cases} K^{-1/2}\sin\left(K^{1/2}w\right) & (K > 0) \\ w & (K = 0) \\ (-K)^{-1/2}\sinh\left[(-K)^{1/2}w\right] & (K < 0) \end{cases} \quad (1.3)$$

In a homogeneous and isotropic Universe there exists a set of *fundamental observers* and a *cosmic time*, such that the Universe appears homogeneous and isotropic in all its properties to all fundamental observers at a given cosmic time (Gunn, 1978).

A space described by the RW-metric (1.1) is not static, but expands or contracts with time, which is implicitly expressed by the time dependence of the cosmic scale factor $a(t)$. The expansion of space leads to the redshifting of photons travelling from source to observer. A photon of frequency ν_e emitted by a fundamental observer at time t_e, and received by a second fundamental observer at the present time t_0 at frequency ν_0 is given by

$$\frac{\nu_e}{\nu_0} = \frac{a(t_0)}{a(t_e)} \equiv 1 + z. \quad (1.4)$$

Here z is the redshift, which is an observable. Therefore I will describe the time dependence in terms of redshift.

1.1.2 Dynamics

The temporal evolution of the cosmic scale factor $a(t)$ and the dependence of the curvature constant K on the matter content is determined by Einstein's field equations. Einstein's field equations

$$G_{\alpha\beta} = \frac{8\pi G}{c^2}T_{\alpha\beta} + \Lambda g_{\alpha\beta} \quad (1.5)$$

relate the Einstein tensor $G_{\alpha\beta}$ to the stress energy tensor $T_{\alpha\beta}$ for all fields present (matter, radiation, etc.) The cosmological constant Λ was historically introduced by Einstein to allow for static cosmological solutions to the field equations. In modern theories the cosmological constant can be interpreted as vacuum energy density. As the metric given in (1.1) and (1.2) is highly symmetric, Einstein's field equations imply that $T_{\alpha\beta}$ has the form of a stress energy tensor for a homogeneous perfect fluid. Such a perfect fluid is completely characterized by its density $\rho(t)$ and its pressure $p(t)$. For the RW-metric (1.1) and (1.2) the equations (1.5) reduce to two independent equations

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3}, \quad (1.6)$$

and

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

where (1.6) is Friedmann's equation. Once the equation of state $p = p(\rho)$ is specified, the two differential equations (1.6) and (1.7) are sufficient for determining the complete temporal evolution of the universe. The equation of state for ordinary matter, which is often called dust, has $p \ll \rho c^2$, while radiation and other forms of relativistic matter have $p = \rho c^2/3$. The scale factor is chosen to be $a = 1$ at the present epoch t_0.

The relative expansion rate $\dot{a}a^{-1} = H(t)$ is called the Hubble parameter. The value of the Hubble parameter at the present epoch $t = t_0$ is the Hubble constant $H(t_0) = H_0$. Since the value of the Hubble constant is still uncertain, it is usually expressed as

$$H_0 = 100\, h \text{ km s}^{-1}\text{Mpc}^{-1} \tag{1.8}$$

with $h \in [0.5, 1.0]$. From the Hubble constant one can obtain the critical density of the universe ρ_{cr}

$$\frac{3H_0^2}{8\pi G} \equiv \rho_{\text{cr}} \approx 1.88 \times 10^{-29} h^2 \frac{\text{gr}}{\text{cm}^3}. \tag{1.9}$$

The density $\rho(t_0) \equiv \rho_0$ in units of the critical density ρ_{cr} is called the density parameter Ω_0. Furthermore the density parameter of the vacuum energy is defined as

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2}. \tag{1.10}$$

Neglecting the energy density of relativistic matter the Friedmann equation (1.6) can be recast in terms of the parameters introduced above as

$$H^2(t) = H_0^2\left(a^{-3}(t)\Omega_0 - a^{-2}(t)\frac{Kc^2}{H_0^2} + \Omega_\Lambda\right). \tag{1.11}$$

Since $H(t_0) \equiv H_0$ the condition

$$K = \left(\frac{H_0}{c}\right)^2 (\Omega_0 + \Omega_\Lambda - 1) \tag{1.12}$$

for the curvature constant K arises. Thus the geometry of the spatial hypersurface is determined by the density contributions from matter, Ω_0, and the cosmological constant, Ω_Λ:

$$\Omega = \Omega_0 + \Omega_\Lambda > 1 \implies \text{closed} \tag{1.13a}$$
$$\Omega = \Omega_0 + \Omega_\Lambda = 1 \implies \text{flat} \tag{1.13b}$$
$$\Omega = \Omega_0 + \Omega_\Lambda < 1 \implies \text{open.} \tag{1.13c}$$

Both, Ω_0 and Ω_Λ, determine the dynamics of the universe. A universe with $\rho < \rho_{\text{cr}}$ will expand forever, while in a universe with $\rho > \rho_{\text{cr}}$ the expansion will stop, followed by a contracting phase.

The time dependence of the Hubble constant in the rewritten Friedmann equation (1.11) translates by means of equation (1.4) into a redshift dependence of the density parameter. Due to the equation of state of ordinary matter, $p \ll \rho c^2$, the evolution of the density is determined by equations (1.6) and (1.7). Therefore the density in a matter–dominated universe scales as $\rho = \rho_0 a^{-3}(t)$, which leads to

$$\Omega(z) = \frac{8\pi G}{3H^2(z)}\rho_0(1+z)^3 = \frac{(1+z)^3\Omega_0}{(1+z)^3\Omega_0 + (1+z)^2(1-\Omega_0-\Omega_\Lambda)+\Omega_\Lambda},$$
(1.14)

while the density parameter of the cosmological constant can be written as

$$\Omega_\Lambda(z) = \frac{\Lambda}{3H^2(z)} = \frac{\Omega_\Lambda}{(1+z)^3\Omega_0 + (1+z)^2(1-\Omega_0-\Omega_\Lambda)+\Omega_\Lambda}.$$
(1.15)

1.2   Cosmological Distances

In curved space–time distances are not uniquely defined. Contrary to Euclidian space, different distance measures lead to different results. Therefore distance measurement prescriptions are defined in analogy to relations between measurable quantities in Euclidian space. In this section I define four different distances used later on: the proper distance, the comoving distance, the angular diameter distance, and the luminosity distance. These distances are parameterized by the redshift between two events z_1 and z_2 with $z_1 < z_2$, and the observer is assumed to be at the origin of the coordinate system.

The *proper distance* $D_\text{prop}(z_1, z_2)$ is the distance measured by the travel time of a light ray propagating from a source at $z = z_2$ to an observer at $z = z_1 < z_2$. It is defined as $\mathrm{d}D_\text{prop} = -c\,\mathrm{d}t = -c\,\mathrm{d}a\,\dot{a}^{-1} = -c\,\mathrm{d}a(aH)^{-1}$. This leads to

$$D_\text{prop}(z_1, z_2) = \frac{c}{H_0}\int_{a(z_2)}^{a(z_1)} \left(a^{-1}\Omega_0 + (1-\Omega_0-\Omega_\Lambda) + a^2\Omega_\Lambda\right)^{-1/2}\mathrm{d}a.$$
(1.16)

The *comoving distance* $D_\text{com}(z_1, z_2)$ is the distance on the spatial hypersurface $t = t_0$ between the world lines of the source and an observer comoving with the cosmic flow. Due to the choice of coordinates, it is the coordinate distance between a source at z_2 and an observer at z_1, $\mathrm{d}D_\text{com} = \mathrm{d}w$. Since light rays propagate with $\mathrm{d}s = 0$, we derive $c\mathrm{d}t = -a\mathrm{d}w$ from the metric, and therefore $\mathrm{d}D_\text{com} = -a^{-1}c\mathrm{d}t = -c\mathrm{d}a(a\dot{a})^{-1} = c\mathrm{d}a(a^2H)^{-1}$. Thus

$$D_\text{com}(z_1, z_2) = \frac{c}{H_0}\int_{a(z_2)}^{a(z_1)} \left(a\Omega_0 + a^2(1-\Omega_0-\Omega_\Lambda) + a^4\Omega_\Lambda\right)^{-1/2}\mathrm{d}a.$$
(1.17)

The *angular–diameter distance* $D_\text{ang}(z_1, z_2)$ is defined as in Euclidian space. It relates the physical cross section δA of an object at z_2 to the angle $\delta\omega$ subtended for an observer at z_1, $\delta\omega D_\text{ang}^2 = \delta A$. From

$$\frac{\delta A}{4\pi a^2(z_2)f_K^2[w(z_1,z_2)]} = \frac{\delta\omega}{4\pi},$$
(1.18)

where $a(z_2)$ is the scale factor at emission time and $f_K[w(z_1, z_2)]$ is the radial coordinate distance between the observer and the source it follows that

$$D_{\mathrm{ang}}(z_1, z_2) = \left(\frac{\delta A}{\delta \omega}\right)^{1/2} = a(z_2) f_K[D_{\mathrm{com}}(z_1, z_2)]. \tag{1.19}$$

The *luminosity distance* $D_{\mathrm{lum}}(z_1, z_2)$ is defined by the relation between the luminosity L of an object at z_2 and the flux S received by the observer at z_1. This leads to the following relation of the luminosity distance to the angular diameter distance (Etherington, 1933)

$$D_{\mathrm{lum}}(z_1, z_2) = \left(\frac{a(z_1)}{a(z_2)}\right)^2 D_{\mathrm{ang}}(z_1, z_2). \tag{1.20}$$

Equation (1.20) can be understood considering that photons are redshifted by $a(z_1)a(z_2)^{-1}$, their arrival times delayed by another factor $a(z_1)a(z_2)^{-1}$, and the area of the observer's sphere on which the photons are distributed are grows between emission and absorption in proportion to $[a(z_1)a(z_2)^{-1}]^2$. This accounts for a total factor of $[a(z_1)a(z_2)^{-1}]^4$ in the flux. This leads to a factor of $[a(z_1)a(z_2)^{-1}]^2$ in the distance relative to the angular diameter distance.

For an Einstein-de-Sitter (EdS)–universe, i.e. a universe with a critical density of dust $\Omega_0 = 1$, $p = 0$ and vanishing cosmological constant Ω_Λ, the different distance measures simplify to

$$D_{\mathrm{prop}}(z_1, z_2) = \frac{2c}{3H_0}\left((1+z_1)^{-3/2} - (1+z_2)^{-3/2}\right), \tag{1.21a}$$

$$D_{\mathrm{com}}(z_1, z_2) = \frac{2c}{H_0}\left((1+z_1)^{-1/2} - (1+z_2)^{-1/2}\right), \tag{1.21b}$$

$$D_{\mathrm{ang}}(z_1, z_2) = \frac{2c}{H_0}\frac{1}{1+z_2}\left((1+z_1)^{-1/2} - (1+z_2)^{-1/2}\right), \tag{1.21c}$$

$$D_{\mathrm{lum}}(z_1, z_2) = \frac{2c}{H_0}\frac{1+z_2}{(1+z_1)^2}. \tag{1.21d}$$

1.3 Density Perturbations

As already mentioned in the preamble of this chapter, the Friedmann-Lemaître cosmological models are not able to describe the structure seen on scales of galaxies and galaxy clusters and the evolution of this structure. Therefore the Friedmann-Lemaître cosmological models have to be supplemented with a theory of structure formation.

The standard model of structure formation assumes that structure grows via gravitational instability from small initial mass density perturbations at early times. It is convenient to discuss the fluctuations in terms of their Fourier decomposition, because the Fourier modes do not couple. In the following I will assume that the phases of the fluctuations are uncorrelated, while their amplitudes follow a Gaussian distribution. This picture is consistent with the predictions from inflation, where it is assumed that the initial seed field of perturbations originates from quantum fluctuations in the very early universe, which were blown up during an inflationary phase.

Furthermore I will confine the discussion to the so–called cold dark matter (CDM)–scenario. In this scenario the CDM particles are assumed to be non–relativistic (therefore: cold) particles which only interact gravitationally or weakly. When CDM particles decouple from the radiation field they have a very small velocity dispersion, and therefore thermal pressure forces are negligible. As opposed to hot dark matter (HDM) particles, which have a large velocity dispersion, CDM particles do not damp structure on small scales through free streaming. Henceforth structure in a universe dominated by CDM particles builds in a bottom–up scenario: larger objects are formed via merging of smaller objects. The CDM–scenario has been tested extensively with N–body simulations (Davis et al., 1985). Currently, it seems to agree best with the observational data, while many alternative models like hot dark matter or topological defect models (Pen U.-E., Seljak U., Turok N., 1997) are currently disfavoured by observations.

Dark matter perturbations are characterized by the density contrast

$$\delta(\boldsymbol{x}, t) = \frac{\rho(\boldsymbol{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \tag{1.22}$$

which is defined as the deviation of the density $\rho(\boldsymbol{x}, t)$ from the average cosmic density $\bar{\rho}(t)$. The Fourier decomposition of the density contrast can be written as

$$\hat{\delta}(\boldsymbol{k}, t) = \int \mathrm{d}^3 x \delta(\boldsymbol{x}, t) e^{i\boldsymbol{k}\boldsymbol{x}}, \tag{1.23}$$

where the continuum limit is used.

The growth of the density perturbations is determined by the amplitude of the density contrast. Three different phases can be distinguished: the linear growth of density perturbations in the regime $\delta \ll 1$, quasi-linear theory with $\delta \simeq 1$, and the nonlinear evolution for $\delta > 1$. I will concentrate on the linear and non–linear regime. A description of perturbation theory appropriate for quasi–linear theory can be found in Padmanabhan (1993).

1.3.1   Linear Growth of Density Perturbations

Consider the standard Newtonian equations for the evolution of the density ρ and velocity $\boldsymbol{v}$ of a pressureless fluid under the influence of the gravitational field ϕ

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{1.24a}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\nabla \phi, \tag{1.24b}$$

$$\Delta \phi = 4\pi G \rho. \tag{1.24c}$$

In order to describe structure growth in the universe with mean density $\bar{\rho}(t)$ and cosmic scale factor $a(t)$ it is advisable to change to comoving positions $\boldsymbol{x} = \boldsymbol{r}/a$ and peculiar velocities $\boldsymbol{u} = \boldsymbol{v} - \dot{a}\boldsymbol{x}$ and consider the density contrast (1.22) instead of the density. Eliminating $\boldsymbol{u}$ and ϕ and neglecting all non–linear terms leads to

$$\frac{\mathrm{d}^2 \delta}{\mathrm{d}t^2} + \frac{2\dot{a}}{a} \frac{\mathrm{d}\delta}{\mathrm{d}t} - \frac{3H_0^2 \Omega_0}{2a^3} \delta = 0. \tag{1.25}$$

The general solution to the linearized Eq. (1.25) is a superposition of two linear, independent solutions

$$\delta(\boldsymbol{x}, t) = D_+(t)\Delta_+(\boldsymbol{x}) + D_-(t)\Delta_-(\boldsymbol{x}). \tag{1.26}$$

In the case of an EdS–universe with $\Omega_0 = 1$ and $\Omega_\Lambda = 0$, the growing mode can be described by $\delta \propto D_+(t) \propto t^{2/3} \propto a$, while $D_-(t)$ is rapidly decaying. The decaying mode is given by $\delta \propto t^{-1} \propto a^{-2/3}$. Assuming that all fluctuations were small at the epoch of recombination one can assume that only the growing mode is present at recent epochs.

For cosmological models other than the EdS–universe, solutions of (1.25) have a more complicated structure. For a detailed discussion see Peebles (1980) and Padmanabhan (1993).

The Power Spectrum of Density Perturbations

If the primordial density perturbations are Gaussian they are completely described by their power spectrum $P_\delta(k)$, which is defined by

$$\left\langle \hat{\delta}(\boldsymbol{k})\hat{\delta}^*(\boldsymbol{k}') \right\rangle = (2\pi)^3 \, \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}') \, P_\delta(k) \,, \tag{1.27}$$

where $\hat{\delta}(k)$ is the Fourier transform (1.23) of δ, and the asterisk denotes complex conjugation.

The primordial density spectrum as predicted by inflation can be described by a scale invariant power spectrum

$$P(k) \propto k^n \,. \tag{1.28}$$

For $n = 1$ this is known as the *Harrison–Zel'dovich* spectrum, which I use in the following. For such a spectrum the growths of the horizon and the perturbations with time have cancelling effects. Such a universe is self–similar in the sense of always appearing the same under the magnification of the cosmological expansion.

In the absence of other physical effects each perturbation mode of the primordial spectrum would simply scale with time in accordance with the growing mode of the density contrast. In reality this is not the case; physical effects do change the shape of the primordial power spectrum. This effect can be described by introducing a transfer function $T(k)$

$$P(k) = T^2(k)k \,. \tag{1.29}$$

For a CDM-model one possible fitting formula is given by Bond & Efstathiou (1984)

$$P(k) = \frac{Ak}{[1 + [aq + (bq)^{3/2} + (cq)^2]^\nu]^{2/\nu}} \,, \tag{1.30}$$

where $q = \Gamma^{-1} k$ with the shape parameter Γ, $a = 6.4\,h^{-1}\,\mathrm{Mpc}$, $b = 3\,h^{-1}\,\mathrm{Mpc}$, $c = 1.7\,h^{-1}\,\mathrm{Mpc}$, and $\nu = 1.13$. This approximation to the power spectrum is used for the simulations used later on. The normalisation of the power spectrum A has to be determined from observations. There exist several procedures for doing this:

1. Normalisation to microwave background anisotropies. This normalisation method translates the measured fluctuations in the temperature of the microwave background into an amplitude for $P_\delta(k)$. Due to the large angular scale of the measurement (COBE measured the rms level of $\Delta T/T \approx 1.3 \times 10^{-5}$ at an angular scale of $7°$ (Banday, 1997)) this method fixes the normalisation on these large physical scales ($\approx 1000h^{-1}$ Mpc).

2. Normalisation to the local variance of galaxy counts, pioneered by Davis & Peebles (1983). Galaxies are assumed to be biased tracers of the underlying dark matter distribution (Kaiser 1984; Bardeen et al. 1986; White et al. 1987). Therefore, by measuring the local variance of galaxy counts within certain volumes, and assuming an expression for the bias, the amplitude of dark matter fluctuations can be inferred. Unfortunately, the biasing mechanism of galaxy formation is not yet known exactly (Kauffmann, 1997), making the use of this method uncertain.

3. Normalisation to the local abundance of galaxy clusters (White et al., 1993; Eke et al., 1996; Vianna & Liddle, 1996). As galaxy clusters are thought to form via gravitational instability from dark matter density perturbations, their spatial number density reflects the amplitude of dark matter fluctuations on scales of the order of 10 Mpc h^{-1}.

Since gravitational lensing by large scale structure is most sensitive to scales comparable to $k_0^{-1} \approx 12(\Omega_0 h^2)$Mpc, I will use the cluster normalisation throughout this work.

1.3.2 Non-linear Evolution of Density Perturbations

The above discussion is valid for the linear evolution of the density contrast, i.e. $\delta \ll 1$. At late stages of the evolution or on small scales the density contrast can be of order $|\delta| \simeq 1$ or larger. Then it is no longer possible to linearize Eqs. (1.24a) – (1.24c): The density contrast becomes non–linear and the different Fourier modes couple. In these cases the density contrast can no longer be described by a Gaussian random field. Higher order moments become important and the density contrast can no longer be completely described by the power spectrum.

In order to understand the non–linear evolution of the density field there exist two different approaches: the numerical simulation of structure formation with N–body simulations on the one side and semi–analytical schemes on the other side. These two different approaches complement one another.

The numerical N–body simulations focus on the dark matter as the dominant driver of the evolution of the universe, whereas other physical processes, most prominently the gas dynamics of the baryons, are neglected. Furthermore, the intricate question of finding a sound physical description for the initial conditions still is a challenging task. Even with these difficulties N–body simulations are one of the most accurate methods for studying structure formation. Even though simulations are computationally quite costly, they have proven to be very valuable in calibrating, assessing, and even suggesting new analytic approaches. On the other hand the semi–analytical schemes are computationally inexpensive and concentrate on the basic principles.

From semi–analytical theory I will only describe the spherical top–hat model and Press–Schechter (PS) theory. The spherical top–hat model is the simplest possible

model for the formation of an object, while Press Schechter theory is a well–established heuristic model for describing the mass function of non–linear objects at any given red-shift.

Spherical Top-Hat Model

A rigourous discussion of the spherical top–hat model can be found in Padmanabhan (1993) or in White (1996).

Consider a spherical region with uniform overdensity $\bar{\rho}_s$ and the physical radius R with a homogeneous density contrast $\bar{\delta} = \bar{\rho}_s/\bar{\rho} - 1$. The spherical region contains the mass $M = 4\pi\bar{\rho}_s R^3/3$, and is assumed to be in an otherwise uniform universe. Using Birkhoff's theorem (Misner et al., 1973; Weinberg, 1972) stating that external matter exerts no force on the material within a spherically symmetric region of the spacetime, the time evolution of R is determined by

$$\frac{\mathrm{d}^2 R}{\mathrm{d}t^2} = -\frac{GM}{R^2} = -\frac{4\pi G}{3}\bar{\rho}(1+\bar{\delta})R \qquad (1.31)$$

for a universe with vanishing vacuum energy.

Comparing Eq. (1.31) with Eq. (1.6) it can be seen that the radius of the sphere R evolves like the cosmic scale factor for a universe with a different density, but the same inital time and initial expansion rate.

The first integral of the evolution equation (1.31) is given by

$$\frac{1}{2}\left(\frac{\mathrm{d}R}{\mathrm{d}t}\right)^2 - \frac{GM}{R} = E, \qquad (1.32)$$

where E is the constant of integration. If $E > 0$ the sphere will expand forever. On the other hand, if $E < 0$ then as R increases $\dot{R}$ will become zero and at later times negative, thus describing a contraction and collapse.

Considering the case $E < 0$ in more detail Eq. (1.32) has the parametric solution

$$R = A(1 - \cos(\eta)) \quad , \quad t = B(\eta - \sin(\eta)) \quad ; \quad A^3 = GMB^2 \quad , \qquad (1.33)$$

where A and B are constants related to each other as shown. The parameter η increases with increasing t, while R increases to a maximum value before decreasing to zero. According to Eq. (1.33) the maximum is reached for $\eta = \pi$. At this turn around point we have $\frac{\mathrm{d}R}{\mathrm{d}t} = 0$ and $R = R_{\max}$.

Now we turn to the evolution of the mean density within each mass shell. Since M is constant for each shell, the mean density within a shell is

$$\bar{\rho} = \frac{3M}{4\pi R^3} = \frac{3M}{4\pi A^3 (1 - \cos(\eta))^3}. \qquad (1.34)$$

In order to work out the time evolution of the density contrast $\bar{\delta}(r, t)$ we need to know the evolution of the background density. For an EdS universe the expansion factor $a(t)$ and the density of the background ρ_b is given by

$$a \propto t^{2/3} \quad , \quad \rho_{\mathrm{b}}(t) = \frac{1}{6\pi G t^2} \qquad (1.35)$$

Dividing the mean density $\bar{\rho}(r, t)$ in Eq. (1.34) by the background density, we derive for the mean background density contrast

$$\frac{\bar{\rho}(r,t)}{\rho_{\mathrm{b}}(t)} = 1 + \bar{\delta}(r,t) = \frac{3M}{4\pi A^3} \frac{6\pi G B^2 (\eta - \sin\eta)^2}{(1 - \cos(\eta))^2}. \tag{1.36}$$

Since $A^3 = GMB^2$ the mean density contrast is given by

$$\bar{\delta} = \frac{9}{2} \frac{(\eta - \sin(\eta))^2}{1 - \cos(\eta)} - 1 \tag{1.37}$$

For small t the linear evolution for the average density contrast is recovered. In this limit

$$\bar{\delta} \approx \frac{3\theta^2}{20} \quad , \quad t \approx \frac{B\theta}{6}, \tag{1.38}$$

leading to

$$\bar{\delta} = \frac{3}{20} \left(\frac{6t}{B} \right)^{2/3}. \tag{1.39}$$

For an EdS–universe at $R = R_{\mathrm{max}}$ and $t = t_{\mathrm{max}}$ Eq. (1.33) can be used to determine the constant B (Padmanabhan, 1993). Using the resulting expression we find to leading order

$$\bar{\delta} = \frac{3}{5} \bar{\delta}_i \left(\frac{t}{t_i} \right)^{2/3} \propto a_{\mathrm{EdS}}(t) \tag{1.40}$$

The collapse of the sphere to $R = 0$ occurs at $t = 2t_{\mathrm{max}}$. At this time the extrapolated *linear* overdensity is

$$\delta_{\mathrm{crit}}^0 = \bar{\delta}(2t_{\mathrm{max}}) = \frac{3}{20} (12\pi)^{2/3} = 1.686. \tag{1.41}$$

For open models and models with a cosmological constant one determines different fit formulae for the determination of δ_{crit}^0 (NFW).

The example presented above of a spherical collapse is quite unrealistic, because it is unlikely that the overdense regions are homogeneous and spherically symmetric. In a more realistic model the non–radial gravitational forces will prevent the collapse to $R = 0$ at $t = 2t_{\mathrm{max}}$. Instead the particles are scattered from their radial orbits. The collapse will stop once the system is virialized, meaning that the amount of kinetic energy E_{kin} of the system equals the amount of potential energy E_{pot}.

Press–Schechter Theory

In the following I am interested in determining the mass function of gravitationally bound objects like galaxies and clusters of galaxies in the universe. Even though these objects are non–linear, Press & Schechter (1974) were able to give a simple recipe to compute the number density of bound objects $f(M)\mathrm{d}M$ in the mass range $(M, M + \mathrm{d}M)$ starting from a Gaussian linear density field. Here I sketch the original

arguments given by Press & Schechter (1974). A detailed discussion of PS–theory can be found in White (1996), Padmanabhan (1993) and several other textbooks (Peacock, 1999; Coles & Lucchin, 1995).

The basic assumption entering PS–theory is that regions of high density in the linear field will eventually form gravitationally bound objects. Starting from a density field $\delta_R(\mathbf{x})$ filtered on spatial scale R, and assuming Gaussian statistics for the density field, the distribution of the density fluctuations is given by

$$P(\delta_R, t)\mathrm{d}\delta_R = \frac{1}{(2\pi\sigma^2(R,t))^{1/2}} \exp\left(-\frac{\delta_R^2}{2\sigma^2(R,t)}\right)\mathrm{d}\delta_R, \qquad (1.42)$$

where

$$\sigma^2(R,t) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}|\hat{\delta}_k(t)|^2 W_k^2(R) \qquad (1.43)$$

is the *rms density fluctuation* on scale R at time t. $W_k R$ denotes the Fourier transform of a top–hat window function. There is a one to one correspondence between the filter scale $R = R(M)$ and the mass M. The fraction of objects with mass greater than M is then obtained from the distribution function (1.42) as

$$F(M) = \int_{\delta_{\mathrm{crit}}^0}^{\infty} P(\delta_R, z)\mathrm{d}\delta_R = \frac{1}{2}\mathrm{erfc}\left(\frac{\delta_{\mathrm{crit}}^0}{\sqrt{2}\sigma(R,z)}\right), \qquad (1.44)$$

where $\mathrm{erfc}(x)$ is the complementary error function, and a parameterization with redshift z instead of time t is used. Press & Schechter (1974) made the *assumption* that this fraction be identified with the fraction of particles which are part of a non linear lump with a mass excceding $M = 4\pi\bar{\rho}a^3R^3/3$. A reasonable choice for δ_{crit} is the linear overdensity at collapse of a spherical perturbation of 1.688 given in Eq. (1.41).

$\sigma_{\mathrm{crit}}(z)$ depends on the growing mode $D_+(z, \Omega_0, \Omega_\Lambda)$

$$\sigma(R, z) = \sigma(R, 0)D_+(z). \qquad (1.45)$$

Fit formulae for the computation of the growing mode $D_+(z, \Omega_0, \Omega_\Lambda)$ for different cosmologies can be found in NFW. The mass function $f(M)$ is determined by simply differentiating equation (1.44) with respect to M

$$f(M) = \frac{\partial F}{\partial M}. \qquad (1.46)$$

Under the preposition that high density regions will form bound objects we derive the comoving number density $N_{\mathrm{halo}}(M, z)$

$$N_{\mathrm{halo}}(M, z)\mathrm{d}M = -\left(\frac{\bar{\rho}}{M}\right)\left(\frac{1}{2\pi}\right)^{1/2}\left(\frac{\delta_{\mathrm{crit}}}{\sigma}\right)\left(\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)\exp\left(-\frac{\delta_{\mathrm{crit}}^2(z)}{2\sigma^2}\right)\mathrm{d}M \qquad (1.47)$$

by multiplying the mass function (1.46) with $2\bar{\rho}/M$. In expression (1.47) an additional factor of 2 is included correcting for the fact that the integral over all masses is

$$\int_0^{\infty} f(M)\mathrm{d}M = \int_0^{\infty} \mathrm{d}F = \frac{1}{2}, \qquad (1.48)$$

even though it should be unity. This discrepancy arises because points with $\delta < \delta_{\mathrm{crit}}$ can also contribute to the number density N_{halo} if a larger filter size $R_1 > R$ is used, i.e. the probability for this to happen is non zero, a point neglected in the above arguments. An alternative derivation correctly accounting for the underdense regions can be found in Bond et al. (1991).

Once the cosmology, power spectrum $|\delta_k|$ and the filter function W_R are specified the comoving number density N_{halo} can be computed.

Chapter 2

Gravitational Lensing

Here I briefly review the ideas important for weak gravitational lensing and for weak cosmological lensing. More thorough reviews of gravitational lensing are found in Schneider, Ehlers, and Falco (1992), Blandford & Narayan (1992), and Narayan & Bartelmann (1996); a review concentrating on weak lensing is written by Bartelmann & Schneider (1999).

After describing the basics of weak gravitational lensing I present the spatially filtered mass measure M_{ap} and its properties. The aperture mass measure M_{ap} was first introduced by Schneider (1996, hereafter S96). First semi–analytical investigations were performed by Schneider et al. (1998, hereafter SvWJK). M_{ap} is well suited to determine the statistical properties of large–scale structure (LSS), but it can also be employed to detect clusters according to their mass irrespective of their luminous properties.

These applications of M_{ap} are tested in later chapters with N–body simulations. Here the reader is acquainted with the N–body simulations used. Furthermore I discuss the generation of shear maps from ray tracing simulations of weak lensing needed for determining the statistical properties of LSS, the cosmic shear.

2.1 Weak Gravitational Lensing

In Einstein's Theory of General Relativity, light propagates on null geodesics of the space–time. Fortunately, for almost all astrophysical relevant cases, the overall geometry of the universe can be described by the FLRW-metric introduced in Sec. 1.1, where the matter inhomogeneities responsible for the lensing can be considered as local perturbations. Within this approximation a viable picture to describe gravitational lensing is to assume that light propagates through an unperturbed space–time up to a point close to the matter inhomogeneity acting as lens. The light deflection close to the lens is characterized by a locally flat Minkowskian space–time weakly perturbed by the Newtonian gravitational potential of the mass distribution of the lens. The prerequisites for this approach are that the Newtonian potential ϕ is small: $|\phi| \ll c^2$, that the peculiar velocity v of the lens is small, $v \ll c$, and that the lens is small compared to the Hubble radius $\ll c/H_0$. These conditions are satisfied in virtually all situations of astrophysical interest, e.g. for a cluster of galaxies typical velocities are of order 10^3 km s^{-1}, and typical Newtonian potentials are of order $\phi < 10^{-5}c^2$.

The effect of space–time curvature can be described by an effective index of refrac-

tion n,

$$n = 1 - \frac{2}{c^2}\phi = 1 + \frac{2}{c^2}|\phi|. \tag{2.1}$$

Since the Newtonian gravitational potential is defined such that it approaches zero at infinity, thus leading to a negative potential, the refractive index is larger than unity: $n > 1$. In this case light travels more slowly in a gravitational field ϕ as compared to free vacuum causing the deflection of the light rays when they pass through the gravitational field. The deflection angle is the integral along the light path of the gradient of $\ln n$ perpendicular to the light path, i.e.,

$$\hat{\boldsymbol{\alpha}} = -\int \boldsymbol{\nabla}_{\perp} \ln n \, \mathrm{d}l = \frac{2}{c^2}\int \boldsymbol{\nabla}_{\perp}\phi \, \mathrm{d}l. \tag{2.2}$$

For all astrophysically relevant situations the deflection angle is very small. Therefore $\boldsymbol{\nabla}_{\perp}\phi$ can be integrated along the unperturbed ray instead of the deflected light ray, thus simplifying computations. Using Eq. (2.2) the deflection angle of a point mass can easily be computed. The Newtonian potential of the lens is

$$\phi(b, z) = -\frac{GM}{(b^2 + z^2)^{1/2}}, \tag{2.3}$$

where b is the impact parameter of the unperturbed light ray, and z indicates the distance along the unperturbed light ray from the point of closest approach. Now the deflection angle of a point mass is given as

$$\hat{\alpha} = \frac{4GM}{c^2} \tag{2.4}$$

If the distances between source and lens and lens and observer are much larger than the spatial extensions of the lens itself, it is justified to use the *thin screen approximation*. In this case the mass distribution $\rho(\boldsymbol{r})$ of the lens can be projected along the line–of–sight onto the lens plane perpendicular to the line–of–sight. The lens is then characterized by the surface mass density

$$\Sigma(\boldsymbol{\zeta}) = \int \rho(\boldsymbol{\zeta}, z) \, \mathrm{d}z, \tag{2.5}$$

where $\boldsymbol{\zeta}$ is the two dimensional vector in the lens plane. Using the surface mass density the deflection angle $\hat{\boldsymbol{\alpha}}$ is the sum of the deflections due to all mass elements in the plane

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\zeta}) = \frac{4G}{c^2}\int \Sigma(\boldsymbol{\zeta}')\frac{\boldsymbol{\zeta} - \boldsymbol{\zeta}'}{|\boldsymbol{\zeta} - \boldsymbol{\zeta}'|^2}\mathrm{d}^2\zeta'. \tag{2.6}$$

The *lens equation* relates the true position of the source to the positions of the images on the lens plane. In Figure 2.1 the lensing geometry of a typical lens system in the thin screen approximation is displayed. D_d, D_ds, and D_s denote the angular diameter distances from the observer to the lens, the lens to the source, and the observer to the source, respectively. The source plane and lens plane are perpendicular to the line–of–sight. Then the lens equation can be read off from Fig. 2.1 as

$$\boldsymbol{\eta} = \frac{D_\mathrm{s}}{D_\mathrm{d}}\boldsymbol{\zeta} - D_\mathrm{ds}\hat{\boldsymbol{\alpha}}(\boldsymbol{\zeta}). \tag{2.7}$$
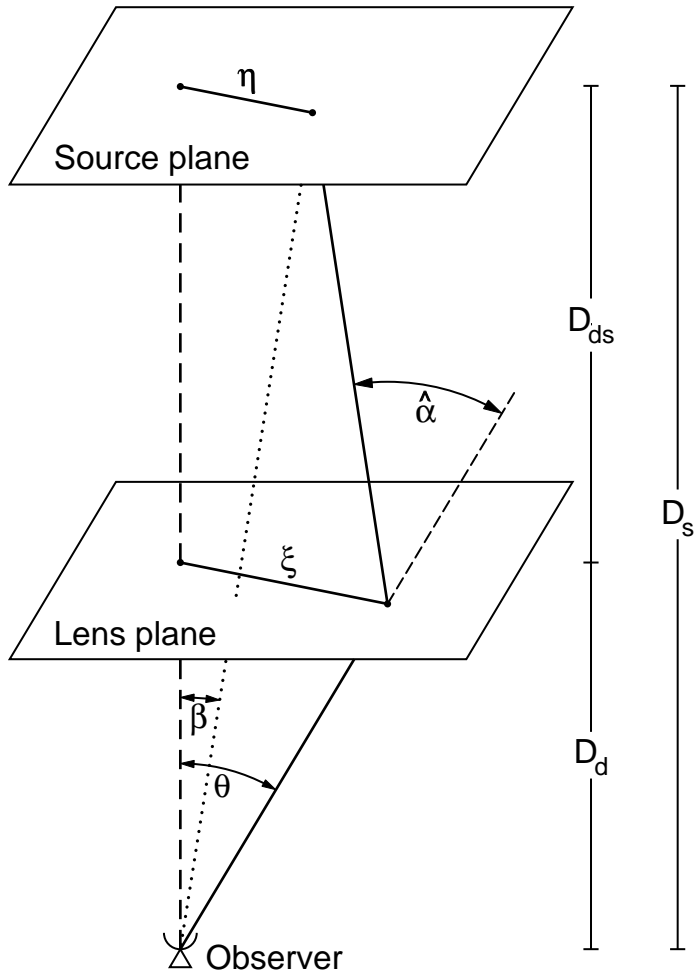
Figure 2.1: Geometrical setup of a typical gravitational lens system in the thin screen approximation. D_d, D_{ds}, and D_s denote angular diameter distances as defined in Sec. 1.2.

With angular coordinates $\eta = D_s\beta$ and $\zeta = D_d\theta$ the lens equation transforms to

$$\beta = \theta - \frac{D_{ds}}{D_s}\hat{\alpha}(D_d\theta) = \theta - \alpha(\theta), \tag{2.8}$$

with $\alpha(\theta)$ being the reduced deflection angle – i.e. the deflection angle at the observer. Equation (2.8) relates the true position β of the source image to an observed angular position θ.

Introducing the dimension-less surface mass density

$$\kappa = \frac{\Sigma(D_d\theta)}{\Sigma_{cr}} \qquad \text{with} \qquad \Sigma_{cr} = \frac{c^2}{4\pi G}\frac{D_s}{D_d D_{ds}}, \tag{2.9}$$

and using Eq. (2.6), the reduced deflection angle reads

$$\alpha(\theta) = \frac{1}{\pi}\int_{\mathbb{R}^2}\kappa(\theta')\frac{\theta - \theta'}{|\theta - \theta'|^2}\mathrm{d}^2\theta'. \tag{2.10}$$

A lens with surface mass density $\Sigma \geq \Sigma_{cr}$ for a certain source position $\boldsymbol{\beta}$ produces multiple images $\boldsymbol{\theta}$. Hence, Σ_{cr} is a characteristic value for the surface mass density and is called critical surface mass density. The dimension-less surface mass density κ — which is also called *convergence* for reasons which will become apparent later in this section — thus can be used to distinguish between cases of weak and strong lensing. In the strong lensing regime $\kappa \geq 1$ holds for some values of $\boldsymbol{\theta}$ which is a sufficient, but not necessary condition for multiple images to occur. In the weak lensing regime $\kappa \ll 1$, to which I will confine the discussion in the subsequent sections.

The reduced deflection angle can be written as the gradient

$$\alpha(\boldsymbol{\theta}) = \nabla\psi(\boldsymbol{\theta}) \tag{2.11}$$

of an effective lensing potential

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

This scalar potential $\psi(\boldsymbol{\theta})$ can also be interpreted as the appropriately scaled, projected Newtonian potential of the lens,

$$\psi(\boldsymbol{\theta}) = \frac{D_{ds}}{D_d D_s} \frac{2}{c^2} \int \phi(D_d \boldsymbol{\theta}, z) dz, \tag{2.13}$$

and the Laplacian of ψ is related to the convergence κ via the two–dimensional Poisson equation

$$\Delta\psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}). \tag{2.14}$$

Therefore the convergence κ and the deflection angle $\boldsymbol{\alpha}$ are related by means of the deflection potential as

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2}\nabla_\theta \cdot \boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{2}\frac{\partial\alpha_i(\theta)}{\partial\theta_i}. \tag{2.15}$$

Looking at angular scales on which the deflection angle is approximately constant, the lens mapping (2.8) can locally be linearized, leading to the Jacobian matrix

$$\mathcal{A}(\boldsymbol{\theta}) = \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial^2\psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \tag{2.16a}$$

$$= (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - |\gamma|\begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix}. \tag{2.16b}$$

In equation (2.16a) the *complex shear* $\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\phi}$ is introduced as

$$\gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}) \quad, \quad \gamma_2 = \psi_{,12}, \tag{2.17}$$

where the indices i following the commas denote partial derivatives with respect to θ_i.

Now it becomes obvious why the dimensionless surface mass density κ is also called convergence: The Jacobian matrix $\mathcal{A}$ describes the local distortion of images.

The first term of equation (2.16a) describes the isotropic focussing of the light rays due to the convergence κ at the position of an image in the lens plane. The second term involving the shear γ accounts for the tidal effects exerted by all the matter outside of the light beam. Thus the convergence leads to an isotropic magnification or demagnification, whereas the shear causes the distortion of the images, e.g. an intrinsically circular source is mapped to an ellipse.

Gravitational lensing conserves surface brightness I, but it changes the angle ω an image subtends on the sky as compared to the undeflected source. Therefore, the magnification μ, i.e. the ratio between area of the image and of the source, is simply the reciprocal of the determinant of $\mathcal{A}$

$$\mu = \frac{1}{\det \mathcal{A}}.$$ (2.18)

So the area distortions caused by the deflection are given by the determinant of the Jacobian matrix of the lens mapping $\boldsymbol{\theta} \longrightarrow \boldsymbol{\beta}$

2.1.1 Weak Cosmological Lensing

Now I briefly turn to weak cosmological lensing, which is needed in order to investigate cosmic shear, the distortion of light bundles from distant sources in the universe used to investigate the intervening mass distribution. Even though the thin screen approximation has to be dropped for investigating lensing by large–scale structure (LSS), it is possible to define a deflection angle $\boldsymbol{\alpha}$ in close analogy to the deflection angle (2.2) defined in the thin screen approximation.

However, before we are in a position to define the deflection angle $\boldsymbol{\alpha}$ for weak cosmological lensing, one complication has to be considered. In weakly perturbed Minkowski space, i.e. in cases where the mass inhomogeneities are well localised one choses an unperturbed ray as reference ray. For weak cosmological lensing this concept has to be generalized to large scale mass inhomogeneities as it is not obvious what an unperturbed ray is in this case. In a universe with expanding background model the meaning of '"straight" fiducial ray is not clear, because any physical fiducial ray is deflected by the potential gradients along its way. Therefore $\boldsymbol{x}(\boldsymbol{\theta}, w)$ is interpreted such that it is the comoving separation vector between an arbitrarily chosen fiducial ray and a closely neighbouring light ray. So, the light ray simply is described relative to the neighbouring, fiducial ray.

The net deflection angle at distance w from the observer between the two rays $\boldsymbol{x}'$ and $\boldsymbol{x}$, divided by the angular diameter distance to w, hence is

$$\boldsymbol{\alpha}(\boldsymbol{\theta}, w) = \frac{f_K(w)\boldsymbol{\theta} - \boldsymbol{x}(\boldsymbol{\theta}, w)}{f_K(w)} \frac{2}{c^2} \int_0^w \mathrm{d}w' \frac{f_K(w - w')}{f_K(w)} \nabla_\perp \phi[f_K(w')\boldsymbol{\theta}, w'].$$ (2.19)

This deflection angle of a light ray starts out at the observer into the direction θ relative to the nearby fiducial ray. Absolute deflection angles cannot be measured; the relative deflection between the two light rays is measured instead. All measurable effects of light deflection depend on the derivatives of the deflection angle, so the choice of the fiducial ray is not important. In analogy to the thin screen approximation (2.15) an ef-

fective convergence κ_{eff} is defined by

$$\kappa_{\text{eff}}(\boldsymbol{\theta}, w) = \frac{1}{2} \nabla_\theta \boldsymbol{\alpha}(\boldsymbol{\theta}, w) = \frac{1}{2} \int_0^w \mathrm{d}w^2 \frac{f_K(w - w') f_K(w')}{f_K(w)} \frac{\partial^2}{\partial x_i \partial x_j} \phi[f_K(w')\boldsymbol{\theta}, w].$$
(2.20)

This effective convergence κ_{eff} can be used to compute lensing effects of LSS in a analogous way to the convergence κ in the thin screen approximation.

2.2 The Aperture Mass Measure M_{ap}

In this section, I briefly summarise the properties of the aperture mass, i.e., its definition, its relation to the shear, and its signal-to-noise ratio. For more details, the reader is referred to S96 and SvWJK.

2.2.1 M_{ap} Statistics

The spatially filtered mass inside a circular aperture of angular radius θ around the point $\boldsymbol{\zeta}$ in the lens plane is defined by

$$M_{\text{ap}}(\boldsymbol{\zeta}) := \int \mathrm{d}^2\vartheta \; \kappa(\boldsymbol{\vartheta}) \; U(|\boldsymbol{\vartheta} - \boldsymbol{\zeta}|),$$
(2.21)

where the continuous weight function $U(\vartheta)$ vanishes for $\vartheta > \theta$. If $U(\vartheta)$ is a compensated filter function,

$$\int_0^\theta \mathrm{d}\vartheta \; \vartheta \; U(\vartheta) = 0,$$
(2.22)

one can express M_{ap} in terms of the tangential shear $\gamma_{\text{t}}(\boldsymbol{\xi}; \boldsymbol{\zeta})$ at position $\boldsymbol{\xi} + \boldsymbol{\zeta}$ relative to $\boldsymbol{\zeta}$ as

$$M_{\text{ap}}(\boldsymbol{\zeta}) = \int \mathrm{d}^2\xi \; \gamma_{\text{t}}(\boldsymbol{\xi}; \boldsymbol{\zeta}) \; Q(|\boldsymbol{\xi}|),$$
(2.23)

(Fahlmann et al. 1994; S96), where

$$\gamma_{\text{t}}(\boldsymbol{\xi}; \boldsymbol{\zeta}) = -\mathrm{Re}\left(\gamma(\boldsymbol{\xi} + \boldsymbol{\zeta}) \, \mathrm{e}^{-2\mathrm{i}\phi}\right),$$
(2.24)

and ϕ is the polar angle of $\boldsymbol{\xi}$. The function Q is related to U by

$$Q(\vartheta) = \frac{2}{\vartheta^2} \int_0^\vartheta \mathrm{d}\vartheta' \; \vartheta' \; U(\vartheta') \; - U(\vartheta).$$
(2.25)

2.2.2 Signal-to-Noise Ratio

An estimate of the shear field γ, and thus of the aperture mass $M_{\text{ap}}(\boldsymbol{\vartheta})$ through Eq. (2.23), is provided by the distortions of images of faint background galaxies. The complex ellipticity of galaxy images is defined in terms of second moments of the surface-brightness tensor (e.g. Tyson et al. 1990; Kaiser & Squires 1993). Specifically,

I use here the ellipticity parameter ϵ (Schneider 1995; Seitz & Schneider 1997), which is defined such that for sources with elliptical isophotes of axis ratio $r \leq 1$, the modulus of the source ellipticities is given as $|\epsilon^{(\mathrm{s})}| = (1 - r)/(1 + r)$, and the phase of the $\epsilon^{(\mathrm{s})}$ is twice the position angle of the major axis.

The complex image ellipticity ϵ can then be calculated in terms of the source ellipticity $\epsilon^{(\mathrm{s})}$ and the reduced shear $g \equiv \gamma (1 - \kappa)^{-1}$ by the transformation (Seitz & Schneider 1997)

$$\epsilon = \frac{\epsilon^{(\mathrm{s})} + g}{1 + g^* \epsilon^{(\mathrm{s})}} \; . \tag{2.26}$$

This relation is valid only for noncritical clusters with $\sigma \ll \sigma_{\mathrm{cr}}$. For critical clusters, it has to be replaced by a different transformation. However, as I am mainly interested in the weak lensing regime, the above relation is sufficient here.

It has been demonstrated (Schramm & Kayser 1995; Seitz & Schneider 1997) that the ellipticity ϵ of a galaxy image is an unbiased estimate of the local reduced shear, provided that the intrinsic orientations of the sources are random

$$\langle \epsilon^{(\mathrm{s})} \rangle = 0, \tag{2.27}$$

with the average taken over an ensemble of sources. Then all average net ellipticities reflect the gravitational tidal effects of the intervening mass distribution. In the case of weak lensing, $\kappa \ll 1$, one then has

$$\langle \epsilon \rangle = g \approx \gamma \tag{2.28}$$

by averaging (2.26) with the probability distribution of the source ellipticities. In my application of the M_{ap} statistics in Chapters 3,4 I assume a Gaussian probability distribution,

$$p_{\mathrm{s}}(|\epsilon^{(\mathrm{s})}|) = \frac{1}{\pi \sigma_\epsilon^2 \left[1 - \exp\left(-\sigma_\epsilon^{-2} \right) \right]} \, \exp\left(-\frac{|\epsilon^{(\mathrm{s})}|^2}{\sigma_\epsilon^2} \right) \; . \tag{2.29}$$

As for the tangential shear component γ_{t} occurring in (2.24), a similar quantity for the image ellipticities can be defined. Consider a galaxy image i at a position $\boldsymbol{\vartheta}_i + \boldsymbol{\zeta}$ relative to the point $\boldsymbol{\zeta}$ with a complex image ellipticity ϵ_i. In analogy to (2.24) the tangential ellipticity $\epsilon_{\mathrm{t}i}(\boldsymbol{\vartheta}_i; \boldsymbol{\zeta})$ of this galaxy is then given by

$$\epsilon_{\mathrm{t}i}(\boldsymbol{\vartheta}_i; \boldsymbol{\zeta}) = -\mathrm{Re}\left(\epsilon_i(\boldsymbol{\vartheta}_i + \boldsymbol{\zeta}) \, \mathrm{e}^{-2\mathrm{i}\phi_i} \right), \tag{2.30}$$

where ϕ_i is the polar angle of $\boldsymbol{\vartheta}_i$.

Now the integral (2.23) can be estimated by a discrete sum over galaxy images,

$$M_{\mathrm{ap}}(\boldsymbol{\zeta}) = \frac{1}{n} \sum_i \epsilon_{\mathrm{t}i}(\boldsymbol{\vartheta}_i; \boldsymbol{\zeta}) \, Q(|\boldsymbol{\vartheta}_i|), \tag{2.31}$$

where n is the number density of galaxy images. The discrete dispersion σ_{d} of the aperture mass $M_{\mathrm{ap}}(\boldsymbol{\zeta})$ is found by squaring (2.31) and taking the expectation value in the absence of lensing, which leads to

$$\sigma_{\mathrm{d}}^2 = \frac{\sigma_\epsilon^2}{2\,n^2} \sum_i Q^2(|\boldsymbol{\vartheta}_i|), \tag{2.32}$$

where $\sigma_\epsilon^2 = \langle |\epsilon^{(s)}|^2 \rangle$. Performing an ensemble average of Eq. (2.32) leads to the continuous dispersion σ_c

$$\sigma_c^2(\theta) = \frac{\pi \sigma_\epsilon^2}{n} \int_0^\theta \mathrm{d}\vartheta \, \vartheta \, Q^2(\vartheta). \tag{2.33}$$

Finally, the *signal-to-noise* ratio S at position ζ is

$$S(\zeta) \equiv \frac{M_{\mathrm{ap}}(\zeta)}{\sigma_d} = \frac{\sqrt{2}}{\sigma_\epsilon} \frac{\sum_i \epsilon_{\mathrm{t}i}(\vartheta_i; \zeta) \, Q(|\vartheta_i|)}{\left[\sum_i Q^2(|\vartheta_i|)\right]^{1/2}} \ . \tag{2.34}$$

The M_{ap} statistics in the continous (2.23) and discretized version (2.31) together with the signal-to-noise ratio (2.34) are the three main equations that will be put to use in the following two Chapters 3 and 4.

2.2.3 Weight Functions

So far, the formalism for aperture mass measures and their signal-to-noise ratios is independent of the choice for the weight function U. Specialising U now, we are led to aperture measures with different merits. Several principal choices for the filter function have been suggested in the literature. In this thesis I will use three different choices for the weight function.

For investigating the statistical properties of LSS in chapter 3 I used the generic filter function for $l = 1$ from the family given in SvWJK: writing $U(\vartheta) = u(\vartheta/\theta)/\vartheta^2$, and $Q(\vartheta) = q(\vartheta/\theta)/\vartheta^2$, I take

$$u(x) = \frac{9}{\pi} \left(1 - x^2\right) \left(\frac{1}{3} - x^2\right), \tag{2.35}$$

and

$$q(x) = \frac{6}{\pi} x^2 (1 - x^2), \tag{2.36}$$

with $u(x) = 0 = q(x)$ for $x > 1$.

Since the filter function (2.35) and (2.36) is not designed for detecting mass concentrations, its filter function is not optimised for achieving high signal-to-noise ratios, leading to high noise levels in a signal-to-noise map. Schneider (1996) solved this problem by introducing the smooth, continuous weight function

$$U_{\mathrm{S}}(x) = \begin{cases} 1 & \text{for} \quad 0 \leq x < \nu_1 R \\ \frac{1}{1-c} \left(\frac{\nu_1 R}{\sqrt{(x - \nu_1 R)^2 + (\nu_1 R)^2}} - c \right) & \text{for} \quad \nu_1 R \leq x < \nu_2 R \\ \frac{b}{R^3} (R - x)^2 (x - \alpha R) & \text{for} \quad \nu_2 R \leq x \leq R \end{cases} \ . \tag{2.37}$$

In the following, the term S-*statistics* refers to the signal-to-noise ratio obtained from Eq. (2.34) using the filter function $U_{\mathrm{S}}(x)$, which guarantees low noise in the signal-to-noise ratio map. The parameters α, b, and c are determined once ν_1 and ν_2 are specified;

see Schneider (1996). I choose $\nu_1 = 0.05$ and $\nu_2 = 0.8$ (leading to $\alpha = 0.7133$, $b = -18.8875$, and $c = 0.1239$) in order to achieve high signal-to-noise ratios and evaluate the S-statistics for an aperture size of 2 arc minutes.

A third possible choice for the weight function leads to the ζ–statistic proposed by Kaiser (1995) and first applied by Fahlman et al. (1994). It gives a lower bound to the average surface mass density κ within a circle inside an annulus by measuring the distortions of background galaxy images inside the annulus. I discuss the ζ–statistic in more detail in chapter 4 in the context of cluster mass estimation.

2.2.4 Analytical Work Done with M_{ap}

The aperture mass has been considered in the framework of blank field surveys in a variety of earlier publications. Introduced as a convenient statistics for cosmic shear, SvWJK have calculated the *rms* of M_{ap} as a function of angular scale, using the Peacock & Dodds (1996) approximation for the non-linear evolution of the power spectrum of density fluctuations. Like other two-point statistics, the dispersion of M_{ap} is an integral over the power spectrum of the projected mass distribution, weighted by a filter function. The filter function corresponding to $\langle M_{\mathrm{ap}}^2 \rangle$ is very narrow and can be well approximated by a delta function (Bartelmann & Schneider 1999). Hence, $\langle M_{\mathrm{ap}}^2(\theta) \rangle$ reproduces the shape of the projected power spectrum and, depending on the cosmological model and the redshift distribution of the sources, it reveals a broad peak at $\theta \sim 1'$. One convenient property of the aperture mass is that the correlation function of M_{ap} of two apertures spatially separated by $\Delta\theta$ quickly decreases and already achieves values of 10^{-2} for $\Delta\theta \sim 2\theta$. This means that measurements of M_{ap} from a large coherent area can be considered independent if the apertures are densely laid out on this data field; this is in contrast to the *rms* shear in apertures which is strongly correlated, and thus must be obtained from widely separated regions on the sky.

Being a scalar quantity, M_{ap} can also be used for higher-order statistical measures of the cosmic shear. SvWJK calculated the skewness of M_{ap}, using Eulerian perturbation theory for the evolution of the three-dimensional density contrast δ. In agreement with Bernardeau et al. (1997) they found that the skewness is a sensitive function of the cosmic density factor Ω_0, and is in this approximation independent of the normalisation of the power spectrum.

A measurement of the dispersion of M_{ap} is affected by two main sources of statistical error: the intrinsic ellipticity distribution of the source galaxies, and cosmic variance. To estimate the latter, one needs to know the kurtosis of M_{ap} which cannot easily be determined analytically.

Values of M_{ap} much larger than its *rms* probe the highly non-Gaussian regime of the projected density field. From its definition, one sees that large values of M_{ap} are expected if the aperture is centred on a density peak with size comparable to the filter scale θ. Therefore, the aperture map can be used to search for such density peaks, presumably collapsed dark matter haloes, in blank field imaging surveys. In this way it is possible to obtain a mass-selected sample of such haloes (S96). Simple analytical arguments in S96 suggest that dark matter haloes with an approximately isothermal profile are detectable with a signal-to-noise ratio larger than 5 if their velocity dispersion exceeds ~ 600 km/s, assuming a number density of background sources of $n \sim 30$ arcmin^{-2}. Indeed, this theoretical expectation was verified in the lensing investigation

of the cluster MS1512+36 (Seitz et al. 1998b). This cluster has a velocity dispersion of about ~ 600 km/s, as obtained from strong lensing modelling and from spectroscopy of cluster members, and is detected in the weak lensing analysis with very high statistical significance.

Assuming that the high signal-to-noise peaks of M_{ap} are due to collapsed dark matter haloes, one can attempt to estimate the abundance of such peaks using analytic theory. KS1 have calculated the number density of haloes with aperture mass larger than M_{ap}, $N(> M_{\mathrm{ap}}, \theta)$, assuming (1) that dark matter haloes are distributed in mass according to Press & Schechter (1974) theory which yields the number density of collapsed haloes as a function of halo mass and redshift (1.47), and (2) that the azimuthally-averaged projected density profiles of these haloes can be described by the projection of the universal halo density profile found in numerical simulations by NFW. They found that the mass profile of haloes in the mass range $3 \times 10^{11} \mathrm{M}_{\odot} \leq M \leq 10^{15} \mathrm{M}_{\odot}$ is independent of the power spectrum and the cosmological parameters, and can be well described by a universal fit formula

$$\rho(r) = \frac{3 H_0^2}{8 \pi G} (1 + z)^3 \frac{\Omega_0}{\Omega(z)} \frac{\delta_{\mathrm{c}}}{r / r_{\mathrm{s}} (1 + r / r_{\mathrm{s}})^2} \qquad (2.38)$$

depending on the characteristic density δ_{c} and the scale radius r_s given as $r_{\mathrm{s}} = r_{200} / c$, where r_{200} is the virial radius of a sphere with an overdensity of $200 \times \rho_{\mathrm{cr}}$ and c is the concentration parameter. $\Omega(z)$ is determined by (1.14).

Depending on the cosmological model and on the redshift distribution of the faint galaxies, the number density of peaks of M_{ap} with a signal-to-noise ratio larger than 5 was estimated to be $\gtrsim 10$ per square degree, and the redshift distribution of these haloes is strongly dependent on the behaviour of the linear growth factor for density perturbations, and thus on Ω_0. This abundance is encouraging, since it allows one to obtain samples of haloes selected by their mass properties alone (for a first observational example, see Erben et al. 1999).

Using the same model, KS2 have calculated the probability distribution of M_{ap} for values of M_{ap} much larger than its _rms_, assuming that this non-Gaussian tail of the probability distribution is dominated by dark matter haloes. They found that the distribution is very well described by an exponential; i.e., the tail is much broader than for a Gaussian.

All these analytic predictions are based on a number of approximations and simplifying assumptions. In Chapter 3 below I compare these analytic results with those found in ray-tracing simulations through a cosmological mass distribution obtained from very large N-body calculations, as described in the next section.

2.3 N–body Simulations

The N-body simulations used are a set of adaptive particle-particle/particle-mesh (AP^{3}M) simulations. The long-range component of the gravitational force is computed by solving Poisson's equation on a grid. The grid calculation is supplemented with a short range correction computed either by a direct sum over neighbouring particles, or, in highly clustered regions, by combining a calculation on a localised refinement mesh with a direct sum over a smaller number of much closer neighbours. The parameters

Simulation	SCDM	τCDM	ΛCDM	OCDM
N_{par}	256^3	256^3	256^3	256^3
$l_{\mathrm{soft}}[h^{-1}\,\mathrm{kpc}]$	36	36	30	30
Γ	0.5	0.21	0.21	0.21
$L_{\mathrm{box}}[h^{-1}\,\mathrm{Mpc}]$	85	85	141	141
Ω_0	1.0	1.0	0.3	0.3
Λ_0	0.0	0.0	0.7	0.0
H_0 [km/s/Mpc]	50	50	70	70
σ_8	0.6	0.6	0.9	0.85
$m_{\mathrm{p}}10^{10}h^{-1}\mathrm{M}_\odot$	1.0	1.0	1.4	1.4
field size [°]	2.7	2.7	3.4	3.9

Table 2.1: Parameters of the N–body simulations.

used by the N-body simulations are given in Table 2.1, all simulations use periodic boundary conditions.

The simulation adopts the approximation to the linear CDM power spectrum (Bond & Efstathiou 1984) already given in Eq. (1.30). The normalisation constant, A, is chosen by fixing σ_8, the *rms* density contrast in spheres of $8\,h^{-1}$ Mpc radius. It is determined following the procedure outlined by White, Efstathiou & Frenk (1993) to meet the present-day local cluster abundance of $\approx 8 \times 10^{-6}\,h^3\,\mathrm{Mpc}^{-3}$ for rich galaxy clusters.

The simulations were run with the parallel adaptive AP^{3}M code HYDRA (Couchman et al. 1995; Pearce & Couchman 1997) kindly made available by the Virgo Supercomputing Consortium (e.g. Jenkins et al. 1998). They followed 256^3 particles using a force law with softening length $l_{\mathrm{soft}} \simeq 30\,h^{-1}$kpc at $z = 0$ (the force is $\sim 1/2$ its $1/r^2$ value at one softening length and is almost exactly Newtonian beyond two softening lengths). l_{soft} was kept constant in physical coordinates over the redshift range of interest to us here. The simulations were carried out using 128 or 256 processors on CRAY T3D machines at the Edinburgh Parallel Computer Centre and at the Garching Computer Centre of the Max-Planck Society. These simulations have previously been used for studies of strong lensing by Bartelmann et al. (1998), for studies of dark matter clustering by Jenkins et al. (1998), and for studies of the relation between galaxy formation and galaxy clustering by Kauffmann et al. (1999a,b), and Diaferio et al. (1999).

2.4 Generation of Shear Maps with Ray-Tracing Simulations

Simulated maps of the shear due to weak lensing by large-scale structure are made by performing ray tracing simulations through the dark matter distribution produced by

the above N-body simulations. The ray tracing simulations of weak lensing were computed by JSW. They used a multiple lens-plane calculation that implements the discrete recursion relations for the position of a given photon and for the Jacobian matrix of the lens mapping at this position (Schneider & Weiss 1988; Schneider et al. 1992; see Seitz et al. 1994 of a thorough justification for this approach). Aside from the distance factors, the main input into the recursion relations are the shear matrices at each lens plane. The ray tracing algorithm consists of three parts: constructing the dark matter lens planes, computing the shear matrix on each plane, and using these to evolve the photon trajectory from the observer to the source. The details involved at each step are as follows:

1. The dark matter distribution between source and observer is projected onto $20 - 30$ lens planes equally spaced (in comoving distance). The particle positions on each plane are interpolated onto a grid of size 2048^2. Since the three-dimensional mass distribution is taken from a single realisation of the evolution of the LSS, the projected mass distributions of consecutive lens planes are correlated. In order to decorrelate them, the projection is carried out along a randomly chosen one of the three coordinate axes; in addition, the origin of the coordinate system in each lens plane is translated by a random vector and the lens plane is rotated by a random angle. In this way, the projected mass distributions of consecutive lens planes are as independent as possible, given the restriction of only a single realisation of the 3-d matter distribution.

2. On each plane, the shear matrix is computed on a grid by Fourier transforming the projected density and using its Fourier space relation to the shear. The inverse Fourier transform is then used to return to real space.

3. The photons start on a regular grid on the first lens plane. Perturbations along the line–of–sight distort this grid and are computed using the relation between deflection angle and projected density. Once the photon positions have been obtained, the shear matrix is interpolated onto them and the recursion relations are solved for the Jacobian of the mapping from the n-th lens plane to the first plane.

4. Solving the recursion relations up to the source plane yields the Jacobian matrix at these positions. Note that the ray tracing is done backwards from the observer to the source, thus ensuring that all the photons reach the observer. The first lens plane is the image plane and has the unperturbed photon positions. All sources are assumed to be at a redshift of $z_\mathrm{s} = 1$.

There are two kinds of resolution limitations in the ray-tracing simulations. The first reflects the finite size and resolution of the N-body simulations, the second the use of finite grids when computing deflection angles and shear tensors on the lens planes. At the peak redshift of the lensing contribution, both effects give a small scale resolution of order $0.2'$. However, since the lens efficiency is not very sharply peaked, effects at other redshifts also enter. Thus depending on the statistical measure being used, the small scale resolution lies in the range $\sim 0.2' - 0.4'$.

On large scales the finite box-size of the N-body simulations sets the upper limit on the angular scales available. The angular size of our simulation box at $z = 1$ is about $3°$. Thus on scales comparable to $1°$, only a few sample regions are available, leading to large fluctuations across different realisations. I therefore restrict my considerations to apertures with radius $\theta \leq 10'$ using one realisation for each cosmological model. For the τCDM model, I use ten different realisations of the ray tracing simulations (i.e., they differ in the direction of projections, the translation and rotation of the projected

matter distribution in the individual lens planes) to estimate the cosmic variance.

Chapter 3

Cosmic Shear and Halo Abundances: Analytical versus Numerical Results

The aperture mass M_{ap} introduced in the last chapter has been shown in a series of recent publications to be a useful tool for detecting cosmic shear. As already pointed out in Sec. 2.2.4 quantitative analytical predictions of M_{ap} have been based on a number of simplifying assumptions. Here I test the validity of these assumptions and the quality of the analytic approximations, using the ray racing simulations through cosmological density fields (Sec. 2.4) generated from N–body simulations (Sec. 2.3).

Those analytic approximations that take into account the non–linear evolution of the matter distribution are in surprisingly good agreement with numerical results. This is true for the root mean square (*rms*) value of M_{ap} and also for the halo abundance, while the predictions for the skewness, based on quasi–linear theory, show a large discrepancy with numerical results.

In addition I verify that the probability distribution function of the aperture mass obtained from the simulated shear maps decreases exponentially for M_{ap} values much larger than the *rms*. Finally, I give values for the kurtosis which allow to make an estimate of the cosmic variance for the *rms* of M_{ap} .

Given the good overall agreement I conclude that comparisons between observed distributions of M_{ap} values and theoretical results provide a powerful tool for testing cosmological models.

3.1 Application of M_{ap} to Simulated Shear Maps

For each of the shear maps generated as described in the last chapter, a 2-dim. "M_{ap} map" is created by simulating "observations" of M_{ap} as a function of position on the 2-dim. shear maps. For the number density of the background sources a value of $n = 30$ arcmin^{-2} is assumed, while the dispersion of the intrinsic ellipticity distribution is chosen as $\sigma_\epsilon = 0.2$. The probability distribution function of M_{ap} (PDF) and some of its moments are then calculated for each M_{ap} map and compared to the analytical model.
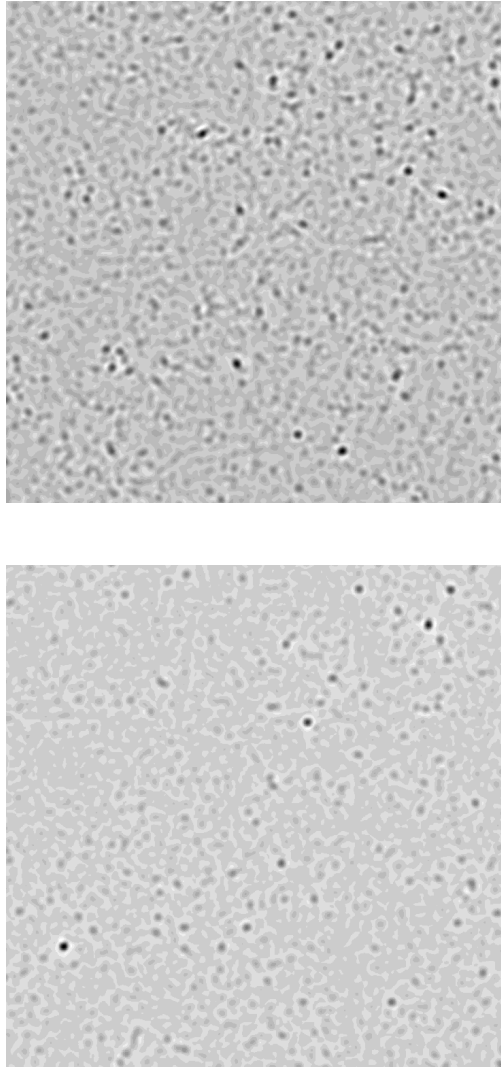
Figure 3.1: The 2-dimensional distribution of M_{ap} for a standard CDM (SCDM, upper panel) and an open model (OCDM, $\Omega_{\mathrm{m}} = 0.3$, lower panel), with parameters given in Table 1. The field size in both panels is $2°$.

It is most instructive to consider two different sets of simulated maps: in the first, I neglect noise from the intrinsic ellipticity distribution of the background sources and compute M_{ap} directly from the shear values on the grid according to Eq. (2.23). This is done either in the limit of weak lensing, i.e., (2.23) is used directly, or γ_{t} is replaced in (2.23) by the reduced shear g_{t}, which is the quantity estimated from the observable galaxy ellipticities.

In the second set of simulations ellipticities of background galaxies are introduced according to the distribution function (2.29). The ellipticities add noise to M_{ap}.

The noise-free results are the ones best compared to the analytic results, whereas those accounting for intrinsic ellipticities yield a more realistic description of the observational situation. In the following the term "without noise" will refer to the first set of M_{ap} simulations, while the term "with noise" will be used for the second one.

As an illustrative example, the 2-dimensional distribution of M_{ap} for a standard CDM (SCDM) and an open model (OCDM) is shown in Fig. 3.1. In both cases high peaks in these maps correspond to haloes in the intervening matter distribution. It is possible to construct a shear-limited sample of haloes from these maps and to determine their abundance. This is done explicitly in Chapter 4.

Comparing the two model universes, we see that the M_{ap} maps reflect the different growth of structure in different cosmologies. The M_{ap} map of the OCDM model is dominated by many isolated peaks which correspond to already collapsed dark matter haloes. The level of background noise coming from matter not yet collapsed is considerably smaller than for the SCDM model in which the structure forms later. The peaks in the SCDM model are less pronounced and isolated than in the open model.

3.2   The PDF of M_{ap} and its Moments

Once the 2-dimensional distribution of M_{ap} is computed, it is straightforward to determine the one-point probability distribution function (PDF) of M_{ap} and its moments. The PDF contains the cosmological information. The lower order moments like *rms* value and skewness can be derived analytically under simplifying assumptions, but the PDF itself cannot be calculated. Therefore, ray tracing simulations provide the only tool for testing the precision of the analytical calculations.
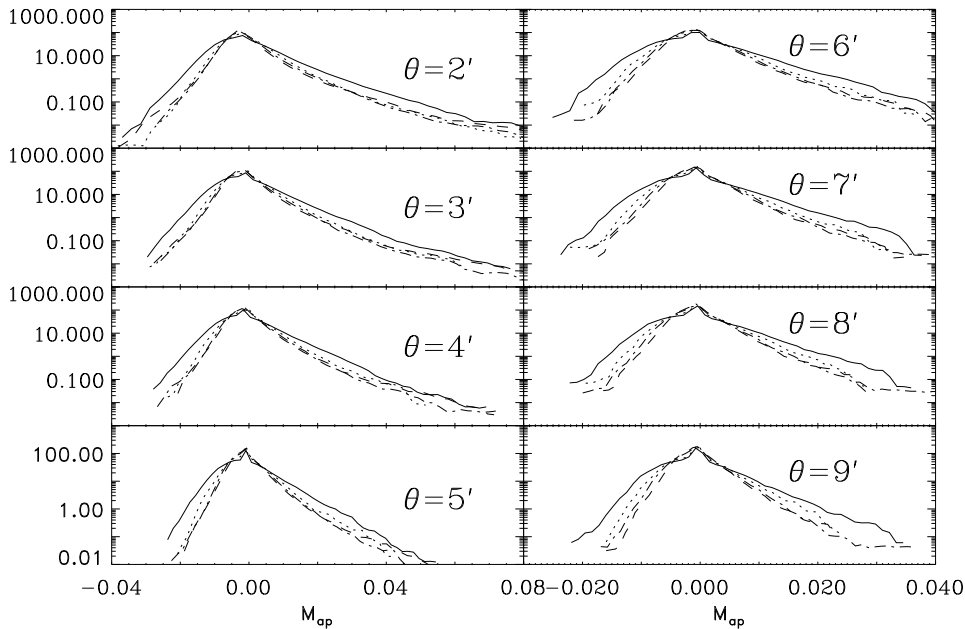

Figure 3.2: The normalised PDF of M_{ap} for different filter scales θ and cosmologies: SCDM (solid line), τCDM (dotted line), OCDM (dashed line) and ΛCDM (dashed-dotted line). The histograms are obtained from M_{ap} maps without noise. Note the different scales on the horizontal axis.
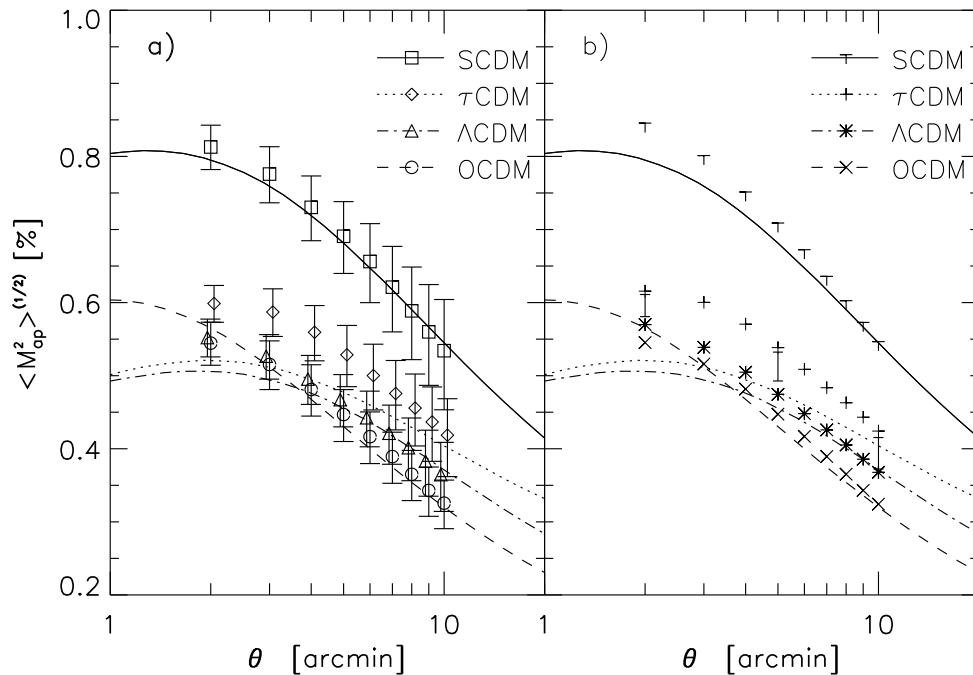
Figure 3.3: The *rms* value of M_{ap} computed with the filter (2.35) versus filter scale θ for different cosmologies. Lines refer to analytic values of $\langle M_{\mathrm{ap}}^2 \rangle^{1/2}$ from SvWJK, while symbols refer to *rms* values obtained from simulations without noise, using γ_{t} (left panel) and g_{t} (right panel). The error bars in the left panel are determined from (3.5) and (3.4). The symbols for OCDM and τCDM are slightly offset along the θ-axis for better display. The error bars in the right panel show the standard errors from 10 realisations of the τCDM model for 2,5, and 10 arcmin. They are centred on the arithmetic mean (*not* on the realisation plotted).

The qualitative features of the PDF for different filter scales θ and for the four different cosmologies (Table 2.1) can be studied in Fig. 3.2. The first point to note is that the non-Gaussian features, namely the tail of the PDF at high M_{ap} values, are less pronounced for larger filter scales. This is due to the fact that the smaller filter scales are more sensitive to the already collapsed, non-linear objects. The second feature to note is the exponential decrease of the tail of M_{ap} which was already obtained semi-analytically in KS2. I shall discuss this feature in more detail later in this section.

I now turn to the *rms* value $\langle M_{\mathrm{ap}}^2 \rangle^{1/2}$ of M_{ap}. Fig. 3.3 compares the analytical *rms* value of M_{ap} calculated using the nonlinear power spectrum of Peacock & Dodds (1996) to the *rms* values computed from the PDFs without noise for γ_{t} (left panel) and g_{t} (right panel). The comparison of the latter shows that the difference between shear and reduced shear is negligible even on filter scales as small as $\theta \sim 2$ arcmin corresponding to the highly nonlinear regime of the mass distribution.

In the left panel of Fig. 3.3, there is an excellent agreement between the analytic predictions and the *rms* values computed from simulations for the SCDM model. There is also good agreement for the ΛCDM and OCDM models, especially for the larger

apertures. The notable exception is the τCDM model, for which the simulations for small filters deviate by a larger factor from the theoretical predictions.

When interpreting this difference between analytical calculation and simulation in the τCDM model, one has to keep in mind that the numerical results of Fig. 3.3 are based on a single realisation. As the cosmic variance is relatively large, it is possible that the large deviation is due to the special choice of the realisation. This interpretation is supported by the fact that the mean for the 10 realisations is considerable lower than for the single realisation plotted. Furthermore, the field sizes of the simulated fields used are too small to represent a characteristic region of the universe.
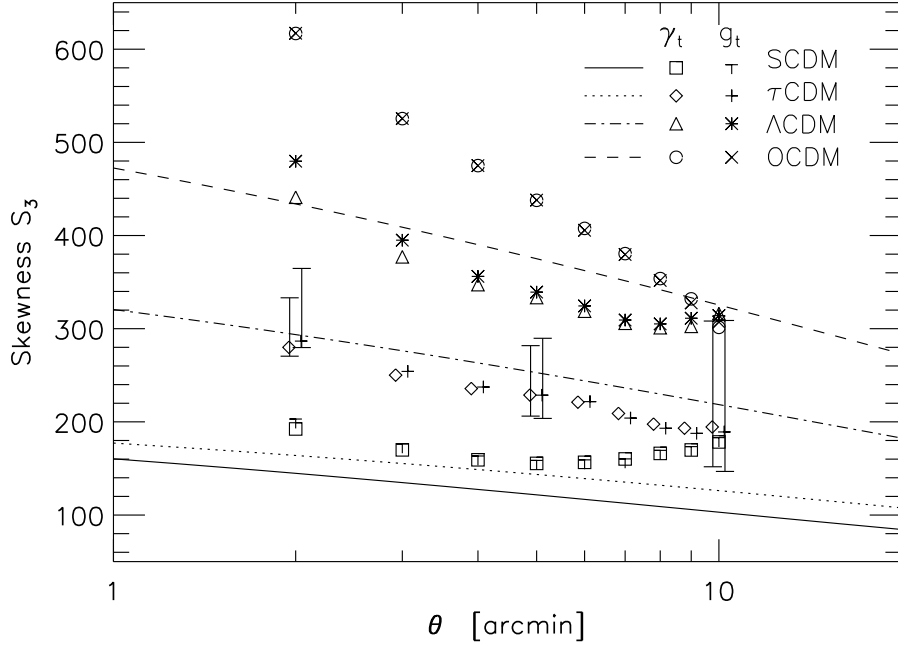
Figure 3.4: The skewness S_3 of the PDF of M_{ap} as defined in (3.1) as a function of filter scale θ for the same cosmological models as in Fig. 3.2. The analytical skewness (lines) from quasi-linear theory is compared to the skewness obtained from the PDF for both, the tangential shear γ_t and the reduced shear g_t. Errors on the τCDM model are like in Fig. 3.3.

The next higher moment of the PDF is the skewness, which is defined as

$$S_3(\theta) := \frac{\langle M_{ap}^3 \rangle}{\langle M_{ap}^2 \rangle^2}, \tag{3.1}$$

for which a similar analysis can be performed as for the *rms* value of M_{ap}. As pointed out by Bernardeau et al. (1997), van Waerbeke et al. (1999), and JSW, the skewness defined in analogy to (3.1) using a top-hat filter is a very sensitive probe of the cosmic density parameter Ω_0.

The dependence of the skewness on filter scale θ is displayed in Fig. 3.4. Again, I compare the skewness computed from the PDF obtained from the ray tracing simulations without noise, both using γ_t and g_t, to the skewness of M_{ap} obtained using

quasi-linear theory (SvWJK). The error bars on the skewness for the τCDM model for 2, 5, and 10 arcmin are derived from the 10 different realisations and are centred on their arithmetic mean.

Again, the differences between the skewness obtained from simulations with γ_t and g_t are small, though slightly larger than for the dispersion, owing to the larger contribution from high-κ regions to the skewness. This difference, which is of order a few percent at most, has been predicted to be small in the Appendix of SvWJK.

When comparing the skewness as determined from second-order perturbation theory for the density evolution to that obtained from simulations (either computed with γ_t or g_t) we see that the former underpredicts the skewness by factors of up to 2. This failure of quasi-linear theory for the prediction of higher-order moments has been demonstrated previously (Jain & Seljak 1997; Gaztanaga & Bernardeau 1998). As the skewness is only determined on scales below 10 arcmin, we are in a regime where the density contrast is non-linear already. The skewness as calculated by Hui (1999) using the so-called hyper-extended perturbation theory (Scoccimarro & Frieman 1999) may provide a more accurate analytical prediction of S_3 than that from second-order perturbation theory.

Another point to note is the increase of the skewness towards smaller filter scales. Generally speaking, such a behaviour is expected, as the non-linear structure growth becomes more and more important for small filter scales. This increase is described insufficiently by quasi-linear theory: for the two EdS universes and even for the Λ model on large filter scales above 5 arcmin, this increase (not the absolute value!) is predicted satisfactorily, but the slope for the open model is larger than analytic values on all scales displayed. This discrepancy between fully non-linear simulations and quasi-linear theory can be attributed to the fact that the open model is much more dominated by already collapsed, non-linear objects than all other models.

The highest moment I consider explicitly is the kurtosis S_4

$$S_4(\Theta) := \frac{\langle M_{\mathrm{ap}}^4 \rangle}{\langle M_{\mathrm{ap}}^2 \rangle^2} - 3. \tag{3.2}$$

The kurtosis is not only important by itself, but also for the determination of the error of the *rms* value of M_{ap}, as will be discussed. As for the skewness, the kurtosis for the noise-free simulations for both γ_t and g_t is plotted, and the scatter for τCDM is determined from the 10 realisations. No analytic result for S_4 is available; however, using third-order perturbation theory, Bernardeau (1998) has calculated the kurtosis for a top-hat filter.

The difference between γ_t and g_t becomes important for the kurtosis, at least for the smaller filter scales, since it is even more dominated by the non-Gaussian tail of the PDF than the skewness. The large error bars on the kurtosis are mainly due to large cosmic variance in combination with the small fields used; thus, the current simulations are unable to provide an accurate determination of S_4.

I now turn to the error bars on the *rms* values of M_{ap} in Fig. 3.3. In the right panel, they were estimated as the standard deviation from 10 different realisations for the τCDM model. The error bars in the left panel were calculated as follows:

As shown in SvWJK, an unbiased estimator of $\langle M_{\mathrm{ap}}^2 \rangle$ from a single aperture is
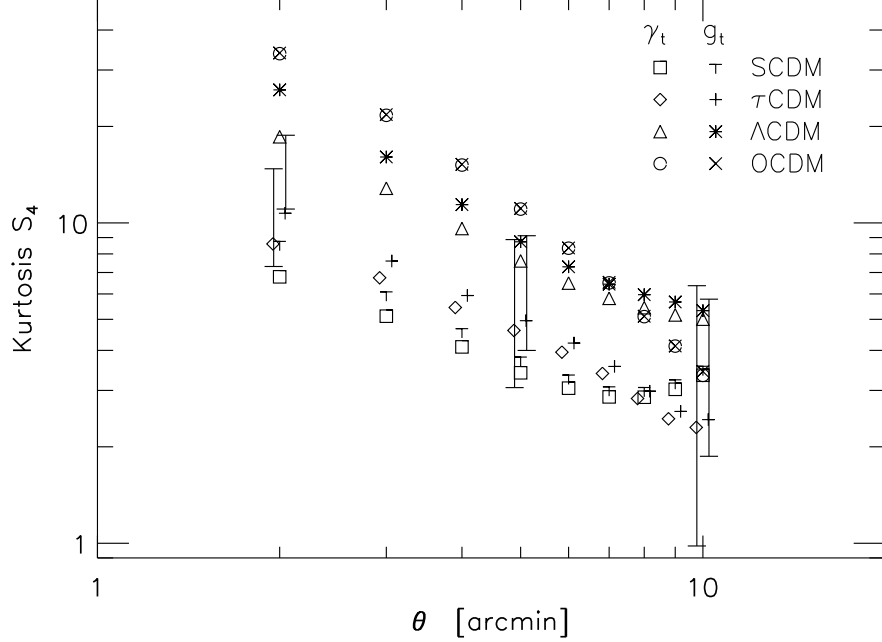
Figure 3.5: The kurtosis S_4 of the PDF of M_{ap} [Eq. (3.2)] as a function of filter scale θ, for the four cosmological models. The kurtosis is derived from the PDF for both, the tangential shear γ_{t} and the reduced shear g_{t}. Errors on the τCDM model are like in Fig. 3.3. No analytic estimate of the kurtosis has been calculated.

given by

$$M = \frac{(\pi\theta^2)^2}{N(N-1)} \sum_{i,j \neq i}^{N} Q_i\, Q_j\, \epsilon_{\mathrm{t}i}\, \epsilon_{\mathrm{t}j} \;, \tag{3.3}$$

where N is the number of galaxies in the aperture, and Q_i is the value of the weight function Q for the i-th galaxy. The dispersion of this estimator is (SvWJK)

$$\sigma^2(M) \approx S_4 \langle M_{\mathrm{ap}}^2 \rangle^2 + \left(\frac{6\sigma_\epsilon^2}{5\sqrt{2}N} + \sqrt{2}\langle M_{\mathrm{ap}}^2 \rangle^2 \right)^2 \;, \tag{3.4}$$

where the two terms in parenthesis correspond to the noise from the intrinsic ellipticity distribution, and the Gaussian cosmic variance, respectively, whereas the term involving S_4 is the excess cosmic variance due to non-Gaussianity. For a collection of N_{f} independent apertures, all containing the same number of galaxy images, an unbiased estimator for $\langle M_{\mathrm{ap}}^2 \rangle$ is the mean $\mathcal{M}$ of M over these apertures, and the dispersion is

$$\sigma\left(\mathcal{M}\right) = \frac{\sigma(M)}{\sqrt{N_{\mathrm{f}}}}. \tag{3.5}$$

Note that this result does not assume that the density field is Gaussian. If one had a collection of N_{f} fields widely separated on the sky, they would be statistically independent, so that $N_{\mathrm{f}} = N$. In the opposite situation where a coherent area on the

sky is available, one can lay down apertures on that field, but they will not be statistically independent. However, as was shown in SvWJK, the M_{ap} values of two apertures which touch each other (i.e., with separation twice their radii), are almost uncorrelated. Whereas the fact that the two aperture masses in these two apertures are uncorrelated does *not* imply that they are *independent* (which would mean that the joint probability distribution for the values of M_{ap} would factorize) – as would be the case for Gaussian fields – I assume the statistical independence for estimating the effective number of fields N_{f} entering (3.5). Thus, the error bars in the left panel of Fig. 3.3 are obtained from (3.5), assuming that the number of independent apertures is $N_{\mathrm{f}} = [\Theta/(2\theta)]^2$, where Θ is the side length of the simulated shear field.

In contrast, the error bars plotted in the right panel of Fig. 3.3 for the τCDM model at the three different filter scales $\theta = 2, 5, 10$ arcmin are based on 10 different realisations of the ray-tracing simulations and allow one to obtain a rough estimate for the error from cosmic variance. Notice that the error bars are centred on the arithmetic mean of the 10 realisations and *not* on the plotted results from a single realisation.

Comparing the size of the error bars in both panels, we see that both methods give errors of the same order of magnitude even though the errors estimated from the 10 realisations are smaller than the errors from the estimator of $\langle M_{\mathrm{ap}} \rangle$. There are two possible reasons for this: first, the effective number of independent apertures is probably larger than the estimate given above, so that the error bars on the left panel in Fig. 3.3 most likely overestimate the true error. Second, in the calculation of the error bars in the right panel, it was assumed that the 10 realisations are independent; but as argued in Sec. 2.4 it is possible that the realisations are not completely independent. This would lead to an underestimation of the cosmic variance. From Fig. 3.3 these two competing effects cannot be quantified. It should be noted that at least on the largest scale plotted, the contribution of the intrinsic ellipticity distribution to the error (3.5) is completely negligible compared to the cosmic variance.

3.3 Halo Abundances

As already indicated in Sec. 2.2.4, high signal-to-noise peaks of M_{ap} can be identified with dark matter haloes, rendering the construction of a mass-limited (more correctly: shear-limited) sample feasible. Analytically, the halo abundances can be modelled using the Press & Schechter (1974) prediction for the mass- and redshift-dependent halo number density, and the universal density profile of NFW, while in the simulated M_{ap} map all connected regions above the corresponding threshold are counted as haloes. Here I consider haloes with signal-to-noise ratio S larger than 5, i.e., a peak in the M_{ap} map is counted as a halo if $M_{\mathrm{ap}} \geq M_5 \equiv 5\sigma_{\mathrm{c}}(\theta)$.

I consider two differently constructed halo abundances in the following: The first sample is simply $N(> M_5, \theta)$, the number density of haloes with an aperture mass larger than M_5 for a given filter size θ. The second sample selects peaks with $M_{\mathrm{ap}} \geq M_5$ within a connected, cross sectional area of $\pi\xi_{\mathrm{t}}^2$, where ξ_{t} is the corresponding cross section radius; the number density of such peaks is denoted $N(> M_5, > \xi_{\mathrm{t}}, \theta)$. Hence, the size of the peaks in the second sample exceeds the threshold ξ_{t}; these peaks are expected to be more robust with respect to noise coming, e.g., from the intrinsic ellipticity distribution and measurement errors. A fixed value of $\xi_{\mathrm{t}} = 0.6$ arc minutes was used
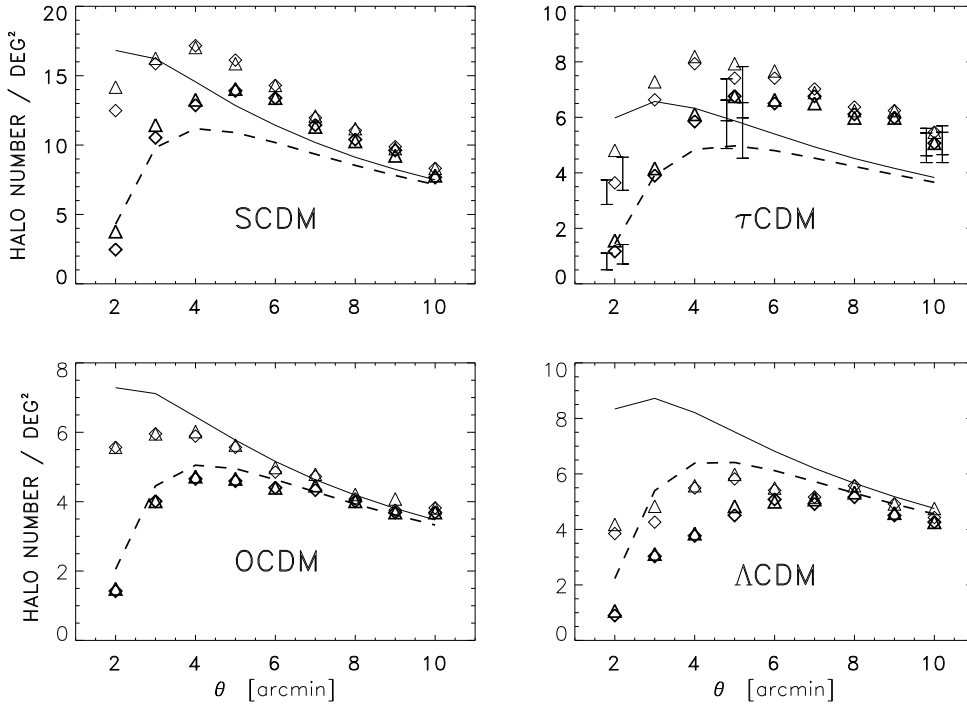
Figure 3.6: The halo number density $N(> M_{\mathrm{ap}}, \theta)$ (thin symbols and thin solid curves) and $N(> M_{\mathrm{ap}}, > 0.6', \theta)$ (thick symbols and thick dashed curves) computed without noise as a function of the filter scale θ for four cosmologies as indicated in the panels. Symbols denote results from the simulations ($\diamond$ from γ_{t}, $\triangle$ from g_{t}) whereas the two line types display the corresponding analytic results from KS1. A signal-to-noise ratio $S > 5$ is used as detection threshold for the haloes. Error bars in the upper right panel display standard errors from 10 realisations for τCDM at 2, 5, and 10 arcmin (errors for γ_{t} offset to the left, errors for g_{t} offset to the right)

corresponding to clusters with a mass of $1.0 \times 10^{15} h^{-1}$ M$_\odot$.

In Fig. 3.6 the number density of the two halo samples as determined from the simulations without noise are compared to the results from the analytic calculation in KS1 over a range of filter scales $2' \leq \theta \leq 10'$. The four panels in Fig. 3.6 refer to the four cosmological models considered. The error bars for the τCDM model at 2,5, and 10 arcmin are again obtained from the 10 different realisations centred on the arithmetic mean of the realisations.

In general, the number counts determined from simulations agree astonishingly well with the analytical results, considering the simplifying assumptions entering the latter. The deviations between simulations and analytical calculation for three of the four cosmologies, namely SCDM, OCDM, and ΛCDM, and especially for the filter scales above 5 arcmin, are less than 10 %. The largest deviation found for these three models is a factor of 2, for the ΛCDM model at smallest filter scale.

The only notable exception is the τCDM model where the deviation remains above 10 % even for the largest filter scales ($\theta = 10$ arcmin). This relatively bad agreement has already been noticed for the *rms* value of M_{ap} and is probably due to the fact that
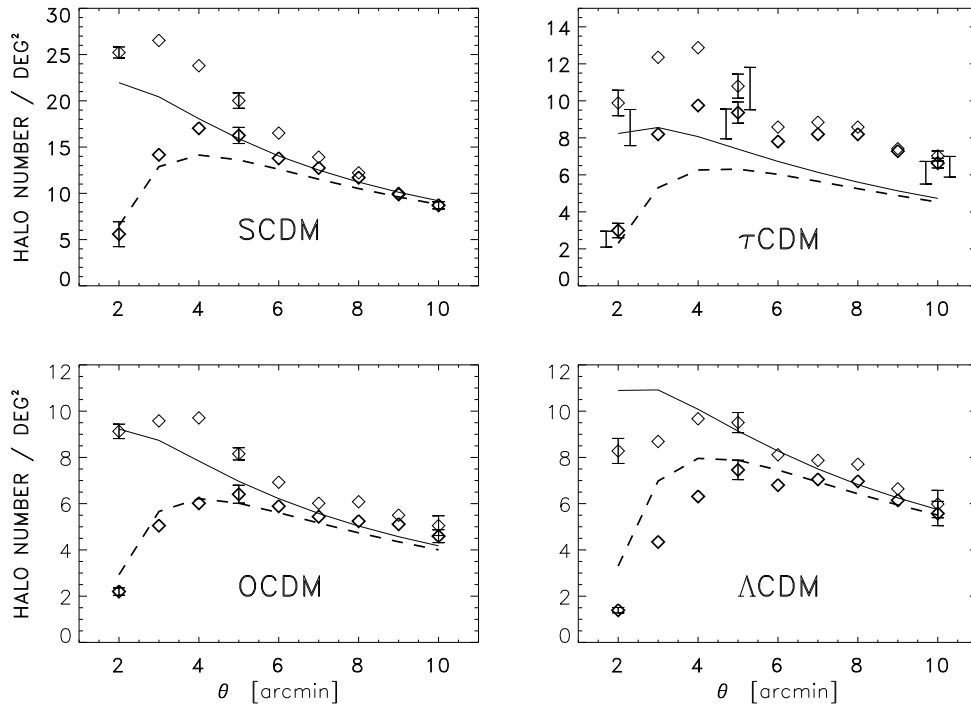
Figure 3.7: The same as Fig. 3.6 but with noise from intrinsic ellipticities of the sources added. For the theoretical model this is done by convolving the values from Fig. 3.6 with a Gaussian, with the dispersion obtained from the intrinsic ellipticity distribution (see KS1). Values from the simulations are denoted by diamonds (thin and thick symbols). Tangential ellipticities now are obtained according to Eq. (2.30). A signal-to-noise ratio $S > 5$ is assumed for the haloes. Error bars centred on the halo abundances are standard errors for 7 realisations of the ellipticity distribution at 2, 5, and 10 arcmin. Error bars in the upper right panel for τCDM are from 10 realisations for 2, 5, and 10 arcmin [error for $N(> M_5, \theta)$ offset to the right, error for $N(> M_5, > \zeta_t, \theta)$ offset to the left].

the realisation plotted is not characteristic for the mean properties of that model, as also indicated by the fact that the halo abundance lies above the mean of all realisations as indicated by the error bars.

The good agreement between analytic estimates and numerical results for the halo number density are surprising, given that (a) Press-Schechter theory does not exactly reproduce the spatial number density of haloes when compared to N-body simulations, and (b) the universal density profile found by NFW has been obtained by spherical averaging, and therefore cannot account for the non-axisymmetry of their projected density. Furthermore, (c) the haloes found from the simulated M_{ap} are expected to be affected by projection effects (Reblinsky & Bartelmann 1999) which are completely neglected in the analytic estimates. Despite these effects which one might suspect to yield significant discrepancies, the analytic estimates are very accurate.

I also investigate the halo abundance in an observationally more realistic situation in Fig. 3.7, including the noise from the intrinsic ellipticity distribution of the back-

ground sources. The plot displays the same quantities as Fig. 3.6 except for the fact that the halo abundances of the two different samples have been determined using the tangential ellipticities in the case of the simulations. The analytic estimates are obtained as in KS1. For all four cosmologies, the error bars are determined using 7 different realisations of the ellipticity distribution of the background sources (2.29). The error bars from the 10 realisations shown for τCDM are slightly sub-Poissonian, as in Fig. 3.6. As expected from the large value of the kurtosis the error coming from the intrinsic ellipticity distribution is much smaller than the error coming from the cosmic variance. On the whole, the number of detected haloes is increased in all cosmologies because, due to the steepness of the Press–Schechter mass function for massive objects, there are more objects just below the threshold than above it. So on average more objects will be lifted above the threshold by noise than brought down below it.

3.4   The Tail of $M_{\rm ap}$

In Fig. 3.8 I compare the PDF for $M_{\rm ap} \geq M_5$ as obtained from analytic calculations (KS2) with that derived from the simulations without noise. The PDF is shown for four filter scales $\theta = 2, 4, 6, 10$ arcmin in the range $M_5 \leq M_{\rm ap} \leq 2M_5$, for which the analytic results predict a nearly exponential behaviour. Indeed, the numerical PDF in the non-Gaussian tail also seems to follow an exponential rather closely, with a slope very similar to the analytic result.

In order to see how much the PDF varies between different realisations, I plot in Fig. 3.9 the PDF for $M_5 \leq M_{\rm ap} \leq 2M_5$ obtained from the 10 realisations in the τCDM model, for 3 filter radii, together with their mean and the corresponding analytic prediction. For the smallest filter scale $\theta = 2'$, all 10 realisations are clearly below the analytic result, whereas for the larger filters, the realisation mean of the PDF agrees very well with the analytic prediction.

Remembering that the analytic predictions were made by assuming that all high values of $M_{\rm ap}$ are coming from regions close to collapsed haloes, in addition to the assumptions used for estimating the number density of $M_{\rm ap}$ peaks (Press-Schechter halo abundance and NFW density profile), this good agreement is somewhat surprising.
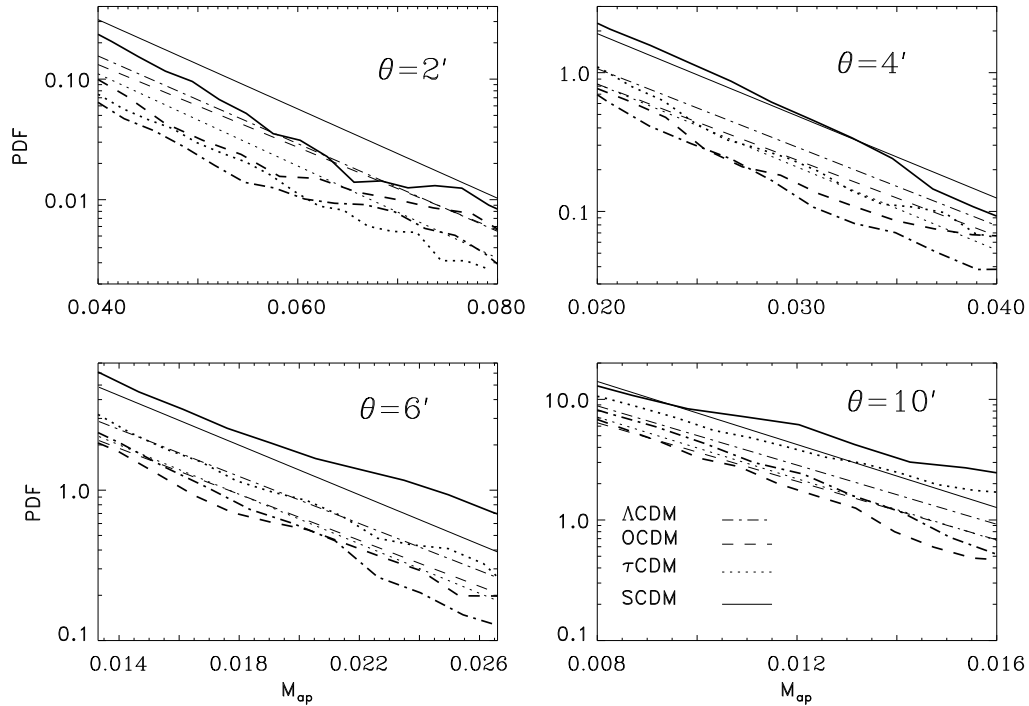
Figure 3.8: The tail of the PDF of M_{ap} for the same cosmologies as indicated in Fig. 3.2 for different filter scales. In each panel we plot the PDF obtained from the analytic estimate in KS2 with thin lines and that from simulations with thick lines. The line types specifies the cosmology: SCDM (solid line), τCDM (dotted line), OCDM (dashed line) and ΛCDM (dashed-dotted line). The M_{ap}-range is $[M_5, 2M_5]$, where $M_5 = S \times 0.016/\theta$ with $S > 5$ for the parameter specified in the text. For computing M_5 I used the fact that the dispersion of M_{ap} for a number density of background sources of $n = 30$ arcmin^{-2} and a dispersion of the intrinsic ellipticity distribution $\sigma_\epsilon = 0.2$ equals $\sigma = 0.016$.
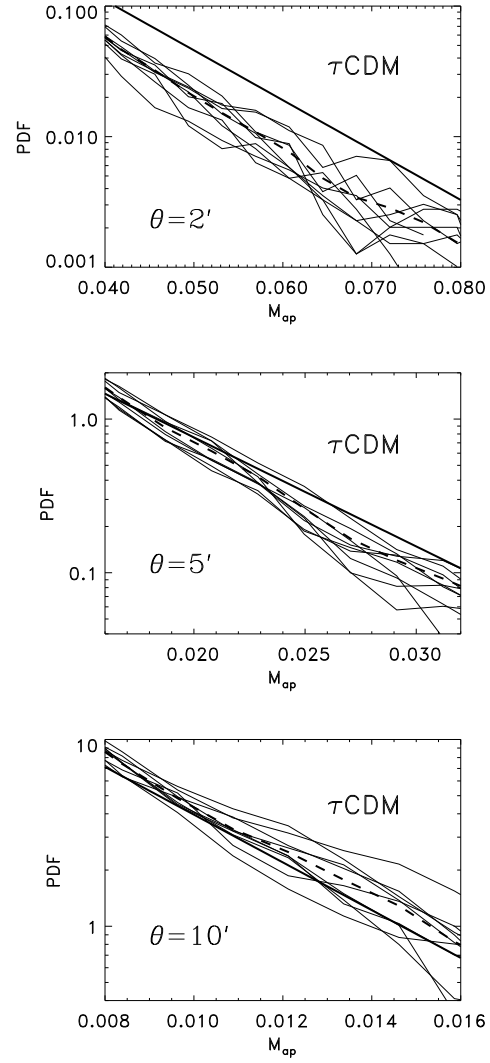
Figure 3.9: The tail of the PDF of M_{ap} for the 10 realisations (thin solid lines) of the τCDM model in comparison to the analytical result (KS2) for the tail of the PDF of M_{ap} (thick solid line). In addition, the mean of the 10 realisations (thick dashed line) is plotted. The different panels are for three filter scales, 2, 5, and 10 arcmin.

Chapter 4

Clusters of Galaxies

In this chapter I investigate S–statistics based on the M_{ap}–statistics concerning its cluster selection properties. One major advantage of the M_{ap}–statistics is that it directly measures mass as opposed to light or X–ray emission based methods. In particular I address the question whether M_{ap}–based selection reduces projection effects prevalent in optically selected cluster samples.

Comparing synthetic cluster catalogues produced from N–body simulation data via the S–statistics to reference catalogues selected by Abell's criterion I find that the S–statistics is more reliable in terms of completeness and spurious detections.

In addition I also investigate the mass estimates in cluster samples selected by both luminosity and weak lensing effects. I conclude that the weak lensing based mass estimates are more accurate than those obtained from the virial theorem.

4.1   Properties and Detection of Clusters of Galaxies

Large galaxy surveys reveal that galaxies are not randomly distributed on the sky. Certain areas on the sky have a noticeably higher or lower than average galaxy density. Galaxies are clustered , quantified by the galaxy–galaxy correlation function, ξ_{gg}, that behaves as

$$\xi_{gg} \simeq (r/5\,h^{-1}\mathrm{Mpc})^{-1.8} \qquad (4.1)$$

for galaxy seperations, $0.1\,h^{-1}\mathrm{Mpc} < r < 20\,h^{-1}\mathrm{Mpc}$. The galaxy–galaxy correlation function measures the probability in excess of random of finding a galaxy at a distance r from another galaxy. Many galaxies are found in small groups of galaxies, and about 10% of all galaxies are found in galaxy clusters, which are gravitationally bound and often virialzed. These galaxy clusters contain from a few tens up to a thousand cluster members. Large clusters of galaxies have typical masses of $5 \times 10^{14}\,\mathrm{M}_{\odot} \simeq 10^{48}\mathrm{gr}$, and typical radii of $1.5\,\mathrm{Mpc} \simeq 5 \times 10^{25}\mathrm{cm}$. In addition, these galaxy surveys find that galaxies form structures like walls and filaments.

There exist several cluster catalogues, the one most widely used being the Abell cluster catalogue (Abell 1958). Briefly, a cluster is classified as an Abell cluster if within the Abell radius of $r_a = 1.5\,h^{-1}\,\mathrm{Mpc}$ from its centre, and after subtraction of the mean background, the number count of galaxies exceeds a certain value n_a. Counting is restricted to the apparent magnitude interval $(m_3, m_3 + 2)$, where m_3 denotes the

apparent magnitude of the third-brightest cluster galaxy. The actual count n_a is used to assign Abell richness classes $\mathcal{R}$. For $\mathcal{R} = 0$, a cluster has to contain at least $n_a = 30$ galaxies, while $\mathcal{R} = 1$ and $\mathcal{R} = 2$ correspond to $n_a \geq 50$ and $n_a \geq 80$, respectively.

The Abell catalogue lists 2712 clusters north of $-20°$ declination away from the galactic plane. This widely used cluster detection and classification scheme does not directly depend on redshift z, even though it has traditionally been used only for fairly shallow cluster surveys. All Abell clusters lie in the redshift range $0.02 \simeq z \simeq 0.2$.

As clusters of galaxies are the largest gravitationally bound objects in the universe and are thought to originate from the highest peaks in the initial density field, they can constrain cosmological models.

The potential problem with using clusters of galaxies for this purpose is that even though they can be easily identified due to their brightness, they are affected by projection effects. Various cosmologically important observables like cluster richness, velocity dispersion, and mass estimates are affected by galaxies projected onto them.

Therefore I compare here the cluster selection by the aperture mass M_{ap} introduced in Chapter 2 with that by the optical Abell criterion. I also present the contamination with intervening matter along the line–of–sight of clusters of galaxies selected by the S–statistics. In addition I investigate the influence of projection effects on mass estimates, comparing the virial theorem based mass estimate to a weak lensing based mass estimate, namely the ζ–statistic. Before doing this I describe the methodology used.

4.2   Methods

4.2.1 N-body Simulation

In order to study the influence of projection effects on cluster surveys selected by optical and gravitational lensing information, I need simulated data allowing to mimic as accurately as possible the selection of clusters and the determination of their properties, for instance their masses. At the same time, the full phase-space information is required to assess the amount of contamination of selected clusters by intervening matter along the line–of–sight (hereafter *los*) and to determine the masses of the clusters with the virial theorem.

For this purpose, I use a large, high-resolution N-body simulation of a standard Cold Dark Matter (SCDM) universe described in Chapter 2 with the parameters given in Table 2.1. I select a simulation box located at a redshift of $z = 0.431$ to achieve high lensing efficiency on sources at redshifts around unity, where I assume the sources to be throughout this chapter.

For the analysis of gravitational lensing effects of the simulated matter distribution, namely the S- and ζ-statistics, the high resolution provided by the N–body simulations is essential. Again, spatial and mass resolution must be distinguished. The spatial resolution is determined by the comoving force softening length, in the case of the SCDM model $l_{soft} = 36\,h^{-1}$ kpc. This limitation is reduced by the high redshift of the simulation box, where the force softening translates to a very small force softening angle, $\theta_{soft} \approx 10''$. The mass resolution, which describes the effect of the finite particle number, is given by the particle mass $m_p = 1.0 \times 10^{10}\,h^{-1}\,M_{\odot}$. The finite mass resolution introduces a white–noise component into the simulations. This is not negligible for

the SCDM model because there a higher proportion of the particles is in voids than for models with lower mean density.

Since I want to evaluate two different methods for detecting clusters *in projection*, I first have to create a sample of *true* 3-dimensional (3-D) clusters or groups from the simulation that will serve as a reference set.

I extract clusters and groups from the 3-D dark-matter simulation with a friends-of-friends algorithm (also called group-finder, cf. Davis et al. 1985). The friends-of-friends algorithm is based on a percolation analysis: It identifies groups and clusters in the simulation box by linking together all particle pairs separated by less than a fraction b of the mean particle separation. Each distinct subset of connected particles is then taken as a group or cluster. I have chosen $b = 0.2$, but the result of group finding does not sensitively depend on the exact choice of b.

Assuming that the clusters and groups found by the group-finder are completely virialised, I compute their virial masses M_{200}, defined as the mass enclosed by a sphere with a radius r_{200} which contains a mean overdensity of 200 times the critical mass density, $\bar{\rho} = 200\,\rho_{\mathrm{crit}}$. For an $\Omega_0 = 1$ universe, this radius approximately separates virialised regions from the infall regions of the haloes (Cole & Lacey 1996). Several reference sets of "true" 3-D clusters are then formed by selecting objects with M_{200} above certain mass thresholds.

In order to be able to apply the S–statistics I first compute the lensing potential

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\,\int_{\mathbb{R}^2}\,\mathrm{d}^2\theta'\,\kappa(\boldsymbol{\theta}')\,\ln\left(|\boldsymbol{\theta} - \boldsymbol{\theta}'|\right) . \qquad (4.2)$$

introduced in (2.12) via the dimensionless surface mass density κ with the Poisson equation (2.14).

Boundary conditions have to be specified when solving Eq. (2.14) numerically. Periodic boundaries are adequate because of the periodicity of the simulation volume.

For numerically computing ψ, the projected particle positions are interpolated on a grid of 2048^2 cells to maintain the high resolution of the N-body simulation. The resulting surface mass density is scaled with the critical surface mass density (2.9) to find the convergence κ. For a numerically stable and efficient method to convert κ to ψ, we use a fast Poisson solver (Swarztrauber 1984). The efficiency of this method rests on the fast Fourier transform (FFT) leading to an asymptotic operation count of $\mathcal{O}(2\,N \log N)$. The algorithm approximates the Laplacian on a grid, transforms to Fourier space, solves the resulting tri-diagonal system of linear equations, and back-transforms to real space. In contrast to other approaches, the approximation is made here by discretising the equations, which can then be solved exactly by a subsequent discrete FFT.

Having determined the deflection potential $\psi(\boldsymbol{\theta})$, the local properties of the lens, such as the surface mass density κ and the complex shear $\gamma \equiv \gamma_1 + \mathrm{i}\,\gamma_2$, can be expressed in terms of second derivatives of $\psi(\boldsymbol{\theta})$, given in (2.17), and (2.14).

As already stated in chapter 2 the parameters α, b, and c of the S–statistics are determined once ν_1 and ν_2 in equation (2.37) are specified; see Schneider (1996). I choose $\nu_1 = 0.05$ and $\nu_2 = 0.8$ (leading to $\alpha = 0.7133$, $b = -18.8875$, and $c = 0.1239$) in order to achieve high signal-to-noise ratios and evaluate the S-statistics for an aperture size of 2 arc minutes. In addition, I assume the number density of background sources

needed to compute the S–statistics is $n = 35$ arc minutes^{-2} and the dispersion of the intrinsic ellipticity distribution is $\sigma_\epsilon = 0.3$.

4.2.2 Construction of Mock Cluster Catalogues by Optical Cluster Selection

Having dark-matter particles only, I need to populate the simulation with galaxies for optical cluster selection. I employ the following scheme.

Galaxy luminosities L are drawn from a Schechter function (Schechter 1976),

$$\phi(L) = N_* \left(L/L_*\right)^{-\alpha} \exp\left(-L/L_*\right) , \tag{4.3}$$

with parameters $L_* = 3.77 \times 10^9 \, L_\odot$ and $\alpha = 1$ taken from the CfA redshift survey by Marzke, Huchra & Geller (1994). The formal divergence for $L \to 0$ in the number-density integral of the luminosity function is avoided by introducing a lower luminosity cut-off $L_0 = 0.1 \, L_*$. For the normalisation of the luminosity distribution, I follow the prescription by Schechter (1976). I calculate a richness estimate by computing the most probable value of the third-brightest absolute magnitude M_3, and then integrate the luminosity distribution from M_3 to $M_3 + 2$. Frenk et al. (1990) showed that this yields the dimension-less normalisation factor $N_* = 60.0$. The normalisation factor determines the amount of luminous galaxies to be introduced into the simulation. The total mass-to-light ratio of the 3-D clusters turns out to be $M/L = 300 \, h \, M_\odot/L_\odot$ on average, in qualitative agreement with observations.

Assuming that mass follows light in the model universe, galaxies inherit positions and velocities from randomly selected dark-matter particles. In this sense, the constructed galaxy sample is unbiased both in number density and velocity.

Transforming luminosities to apparent magnitudes for higher redshifts, I account for the k correction. If the spectral energy distribution varies with frequency ν as a power law with exponent α', the additive k correction is

$$k = -2.5 \left(1 + \alpha'\right) \log_{10}(1 + z) . \tag{4.4}$$

I choose $\alpha' = -1.5$ for the spectral index, which sufficiently well reflects the spectral properties of ordinary galaxies.

Volume-limited cluster catalogues are then obtained after projecting particle positions onto planes along the three orthogonal axes of the simulation box. Groups and clusters in projection are identified with a 2-D version of the friends-of-friends algorithm.

I then apply the optical Abell criterion (Abell 1958) to select galaxy clusters. I also straightforwardly apply Abell's criterion to three-dimensional clusters in order to assess the influence of projection effects on richness-class estimates.

For the background subtraction, I follow Frenk et al. (1990). In order to estimate the background, i.e. the contamination by foreground and background galaxies in the simulation box, I assume that the number of galaxies contributing to the contamination is proportional to the volume projected onto the cluster. In this case, I expect 8 background galaxies within a cylinder of volume $V_{\text{proj}} = \pi r_a^2 l_{\text{box}}$. Therefore, a cluster with richness class $\mathcal{R} = 1$ has to encompass at least 58 galaxies in the appropriate magnitude interval; 8 background galaxies in addition to the 50 genuine cluster members.

Since observed column densities towards galaxy clusters will be considerably larger than assumed here, and since the conditions in realistic observations are less controlled than here, projection effects could even be larger in reality.

4.2.3 Detection of Dark-Matter Concentrations through Weak Gravitational Lensing

The ζ-Statistics

A special choice of the weight function for M_{ap} leads to the ζ-statistics proposed by Kaiser (1995) and first applied by Fahlman et al. (1994). It gives a lower bound to the average surface mass density κ within a circle inside an annulus by measuring the distortions of background galaxy images inside the annulus. The ζ-statistics will be used in Sec. 4.5 for constraining the masses of clusters detected through their S-statistics. The piece-wise constant weight function for the ζ-statistics reads (Schneider 1996)

$$U_\zeta(x) = \begin{cases} \dfrac{1}{\pi x_1^2} & \text{for} \quad 0 \leq x < x_1 \\[2mm] -\dfrac{1}{\pi(R^2 - x_1^2)} & \text{for} \quad x_1 \leq x < R \\[2mm] 0 & \text{for} \quad R \leq x < \infty \end{cases} \qquad (4.5)$$

Inserting this weight function into Eq. (2.23) yields

$$m_\zeta(\boldsymbol{\theta}) \equiv \zeta(r_1, r_2) = \frac{1}{\pi} \frac{r_2^2}{r_2^2 - r_1^2} \int_{r_1}^{r_2} \frac{\mathrm{d}^2 r}{r^2} \gamma_{\mathrm{t}}(\boldsymbol{r}) \;, \qquad (4.6)$$

where $\boldsymbol{r}$ is the distance vector between the point under consideration and $\boldsymbol{\theta}$, and r_1 and $r_2 > r_1$ are the bounding radii of an annulus around $\vec{x}$. It can then be shown that $\zeta(r_1, r_2)$ is related to the mean convergence $\bar{\kappa}(r_1, r_2)$ in the annulus by

$$\zeta(r_1, r_2) = \bar{\kappa}(r_1) - \bar{\kappa}(r_1, r_2) \;, \qquad (4.7)$$

$\bar{\kappa}(r_1)$ being the mean convergence in the circle with radius r_1 around $\boldsymbol{x}$. In other words, the ζ-statistics constrains the average convergence in a circular aperture through the tangential shear in an annulus surrounding the aperture. Since $\bar{\kappa}(r_1, r_2) \geq 0$, $\zeta(r_1, r_2)$ provides a lower bound to the mean surface mass density enclosed by r_1.

As mentioned before, it is possible to use the image ellipticities ϵ of the background galaxies as unbiased estimates of the tangential shear, $\epsilon_{\mathrm{t}} \approx \gamma_{\mathrm{t}}$. Therefore, the integral in (4.6) can be approximated as a discrete sum over galaxy images,

$$\zeta(r_1, r_2) \approx \frac{1}{2} \frac{r_2^2}{N} \sum_{i=1}^{N} \frac{\epsilon_{\mathrm{t}i}}{r_i^2} \;. \qquad (4.8)$$

In this study, I want to obtain a lower bound to the total cluster masses. For a meaningful application of the ζ-statistics, it is important to include the complete cluster into the measurement. This can be achieved following Bartelmann (1995). If I apply the ζ-statistics to a nested set of annuli with radii r_i, $1 < i < n$, then the ζ-statistics for an annulus with $r_i < r_j$ reads

$$\zeta_{i,j} = \bar{\kappa}_i - \bar{\kappa}_{i,j} \;. \qquad (4.9)$$

where $\bar{\kappa}_i \equiv \bar{\kappa}(r_i)$ and $\bar{\kappa}_{ij} \equiv \bar{\kappa}(r_i, r_j)$. On the other hand, the mass M_{ij} in such an annulus is the product of surface mass density times the area,

$$M_{i,j} = A_{i,j}\, \bar{\kappa}_{i,j}\ , \tag{4.10}$$

where the area of the annulus is

$$A_{i,j} = \pi\, (r_j^2 - r_i^2)\ . \tag{4.11}$$

The crucial point is now that the mass contained within a circle of radius R is always the sum of the masses contained in annuli with outer radii $r_j < R$, irrespective of how the area is decomposed into such rings. Keeping this in mind, Eqs. (4.9) and (4.10) can be combined into a system of $(n-1)$ linear equations with n unknowns m_k, where the m_k denote the masses in adjacent rings $M_{k-1,k}$.

   The fact that there is one equation less than the number of unknowns reflects the scaling invariance of the surface mass density κ. Assuming that the outermost annulus does not contain any significant convergence, i.e. $M_{n-1,n} = 0$, I finally arrive at the following set of equations for the masses M_j enclosed by radii r_j:

$$M_j = (1 + \frac{A_{1,j}}{A_1})\, m_1 - A_{1,j}\zeta_{1,j} \tag{4.12}$$

where m_1 is shorthand for

$$m_1 \equiv A_1\, \frac{A_{1,n}\zeta_{1,n} - A_{1,n-1}\zeta_{1,n-1}}{A_{1,n} - A_{1,n-1}}\ . \tag{4.13}$$

   Of course, a lower bound to the total cluster mass could also be obtained by placing an annulus around the entire cluster and applying the ζ-statistics to that annulus only rather than to a set of nested annuli. The present approach has two advantages; first, it yields a profile of κ which allows to assess the location of the outer cluster boundary, for which I found that $R = 1.8 h^{-1}\mathrm{Mpc}$ is an appropriate choice. Second, it uses galaxy ellipticity measurements in all annuli rather than the outermost only, thus reducing the noise. However, the errors in the M_j are correlated at successive radii, making an immediate interpretation of the significance at any given radius less transparent.

4.3   Completeness of Cluster Catalogues

I am now in a position to investigate completeness and homogeneity of cluster catalogues constructed with the S-statistics as opposed to the optical Abell criterion. To this end, I create two different samples of 2-D clusters by applying both methods to simulated data projected along the x-, y-, and z-axes. I then compare these 2-D clusters with the reference set of 3-D clusters.

   To assess the quality of the constructed 2-D catalogues, I use several reference sets of 3-D clusters with different mass ranges. Looking at different mass ranges instead of cluster richness estimates is motivated by two reasons. First, the physical quantities of primary interest are the masses. Furthermore, this kind of comparison is more suitable for the S-statistics, which does not depend on the distribution of luminous galaxies but of the dark matter only. This way of addressing projection effects in cluster catalogues

3-D clusters	$\mathcal{R} > 0$		$\mathcal{R} \geq 0$	
Mass range in $10^{14}\ h^{-1}\ M_\odot$	det.	spur. det.	det.	spur. det.
1.32 – 3.50	53%	50%	66%	82%
1.03 – 3.50	27%	50%	56%	69%
0.82 – 3.50	17%	50%	44%	64%
0.55 – 3.50	13%	25%	36%	40%
0.10 – 3.50	13%	6%	32%	29%

Table 4.1: Completeness and homogeneity of optically selected, synthetic cluster catalogues constructed by Abell's criterion. For five mass ranges of 3-D clusters, the detection rate (det.) and the spurious detection rate (spur. det.) is shown for cluster richness $\mathcal{R} > 0$ and $\mathcal{R} \geq 0$. The detection rate gives the percentage of 3-D clusters correctly identified by Abell's criterion. The spurious detection rate is the fraction of detected clusters which do not correspond to true 3-D clusters within the mass range considered . Further explanation is given in the text.

differs from previous studies (e.g. Cen 1997; van Haarlem et al. 1997), which focused on the influence of projection effects on the richness estimate of Abell catalogues.

The results for the Abell-selected cluster catalogues are summarised in Table 4.1 and Fig. 4.1. The first column of Table 4.1 lists the mass range of the 3-D cluster reference set. The next two columns show the fraction of 3-D clusters correctly detected by Abell's criterion, and the fraction of 2-D objects which do not correspond to 3-D clusters within the chosen mass range, respectively. The fraction of detected clusters is given with respect to the 3-D clusters in the given mass range, while the spurious detections are given relative to the total number of 2-D detections. The last two columns show the same information for a larger sample of Abell clusters also including clusters of richness class $\mathcal{R} = 0$. Figure 4.1 illustrates the information of Table 4.1 as a histogram.

The first mass range considered in Table 4.1, $(1.32 - 3.5) \times 10^{14}\ h^{-1}\ M_\odot$, reflects the masses expected for Abell clusters with richness class $\mathcal{R} > 0$. Looking at absolute numbers, I find that the total number of 2-D clusters with $\mathcal{R} > 0$ in the Abell catalogue is very similar to the number of 3-D objects in this mass range (in Fig. 4.1, the corresponding bars are comparably long). However, as Table 4.1 shows, only 53% of the 3-D clusters from the reference set can be found in the 2-D sample of $\mathcal{R} > 0$ clusters. On the other hand, a high percentage (50%) of the 2-D Abell clusters does not correspond to a true 3-D object. This means that not only half of the 3-D clusters in this mass range are missed by Abell's criterion, but also a lot of spurious 2-D objects are found. This is due to two competing effects occurring in projection. Intrinsically rich clusters may disappear in the background, while the richness class of poor clusters can be enhanced by small groups and field galaxies collected along the line–of–sight. In the above case these two effects approximately cancel, so that the total numbers are approximately correct.

These results are consistent with the findings of Frenk et al. (1990) for projection effects in CDM-like universes with different biasing parameters b. Comparing Tab. 2 in their paper with the results for Abell $\mathcal{R} = 1$ clusters, I find similar projection effects for the model universe and their CDM-like universes with biasing parameters between $b =$
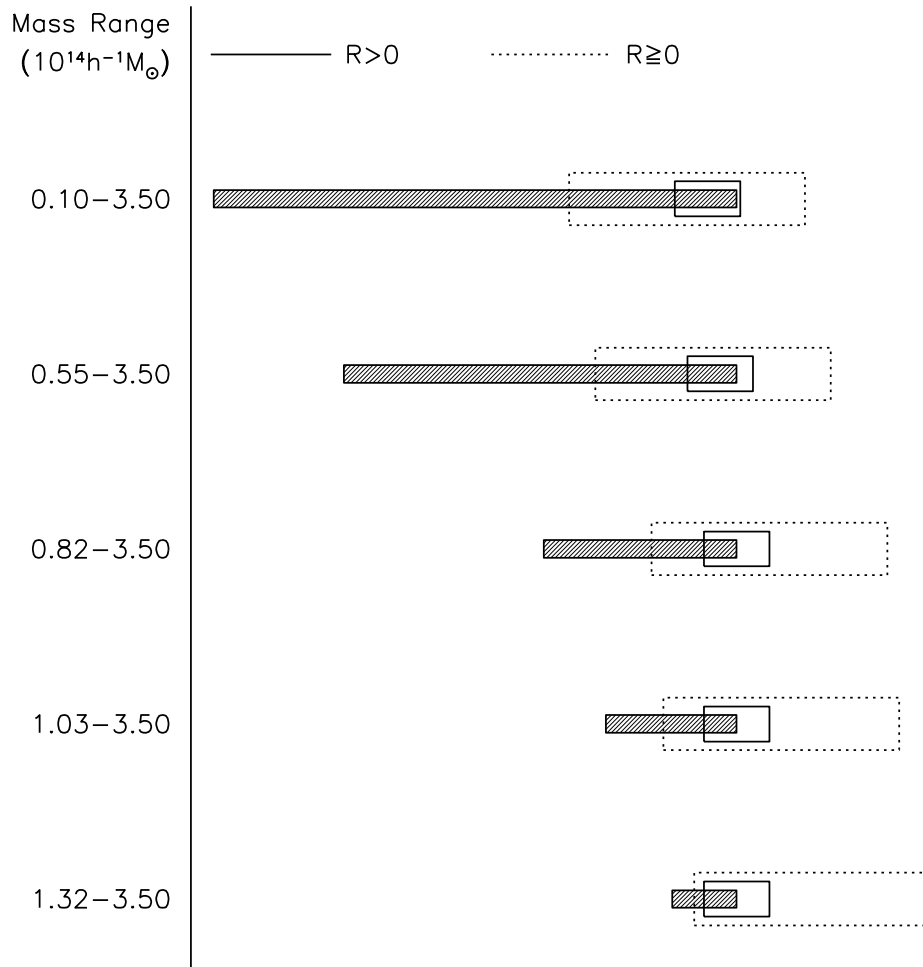
Figure 4.1: Performance of Abell's criterion in identifying clusters in different mass ranges (cf. Table 4.1). The hatched bars illustrate the number of 3-D clusters for each mass range, the solid and dotted bars the number of Abell clusters with $\mathcal{R} > 0$ and $\mathcal{R} \geq 0$, respectively. The overlap between hatched and empty bars illustrates the fraction of correct detections, and the fraction of spurious detections by Abell's criterion corresponds to the rest of the empty bars. The fraction of correctly identified 3-D clusters and the fraction of spurious detections both increase for more massive cluster samples.

2.5 and $b = 1.6$. A direct comparison is difficult because of the different normalisations of their model universe and mine. Furthermore, the redshift dependence of the biasing parameter b is not known, further complicating a detailed comparison.

For 2-D galaxy clusters with richness class $\mathcal{R} = 0$, the fraction of detected 3-D clusters is slightly increased from 53% to 66%. At the same time, the number of spurious 2-D clusters, i.e. clusters which cannot be linked to 3-D objects in the reference set, is increased by more than 30%. A detailed analysis of the line–of–sight structure of these clusters reveals that this large number of spurious detections is partly due to the additive projection of poorer groups corresponding to lower-mass 3-D objects. This increase in the number of both detections and spurious detections reflects the enhance-

ment or reduction of cluster richness classes due to projection.

Extending the reference set of 3-D clusters to lower masses substantiates this assumption. The number of spurious detections declines quite steeply from over 80% for the mass range of $(1.32 - 3.5) \times 10^{14}\,h^{-1}\,M_\odot$ to below 30% for a lower mass threshold of $0.1 \times 10^{14}\,h^{-1}\,M_\odot$, which is more than one order of magnitude smaller than the lower mass threshold for $\mathcal{R} > 0$ clusters. Therefore, many of the 2-D clusters detected by Abell's criterion do indeed correspond to true 3-D objects, but in very different mass ranges. This clearly indicates that for the model universe a change of the richness estimate due to projection is likely. But still the number of truly spurious detections, i.e. detections of 2-D objects which cannot be connected with any 3-D object, remains quite high even in the broadest mass range.

Turning to the performance of the S-statistics in constructing a complete and homogeneous catalogue, Table 4.2 and Fig. 4.2 display the results of the S-statistics in a manner analogous to Table 4.1 and Fig. 4.1 for Abell's criterion. Again, the first column contains the mass range of the investigated 3-D reference set, while the following columns display the percentage of detected 3-D clusters and of spuriously detected 2-D objects above a certain S-value. The analysis is performed for objects detected above different signal-to-noise thresholds.

3-D clusters	$S \geq 5$		$S \geq 4.5$		$S \geq 4$		$S \geq 3.5$		$S \geq 3$	
Mass range in $10^{14}\,h^{-1}\,M_\odot$	det.	spur. det.	det.	spur. det.	det.	spur. det.	det.	spur. det.	det.	spur. det.
$1.32 - 3.50$	60%	50%	60%	68%	100%	65%	100%	80%		
$1.03 - 3.50$	43%	27%	60%	37%	76%	46%	93%	63%		
$0.82 - 3.50$	33%	11%	48%	24%	64%	32%	91%	47%		
$0.55 - 3.50$	18%	6%	28%	10%	37%	21%	61%	28%	81%	32%
$0.10 - 3.50$	14%	6%	22%	7%	30%	16%	46%	29%	90%	22%

Table 4.2: Completeness and homogeneity of a catalogue constructed with the S-statistics based on a weak gravitational lensing analysis of distorted images of background sources. The percentage of detections (det.) and spurious detections (spur. det.) is given for different S-values. For further explanation, see Table 4.1 and the text.
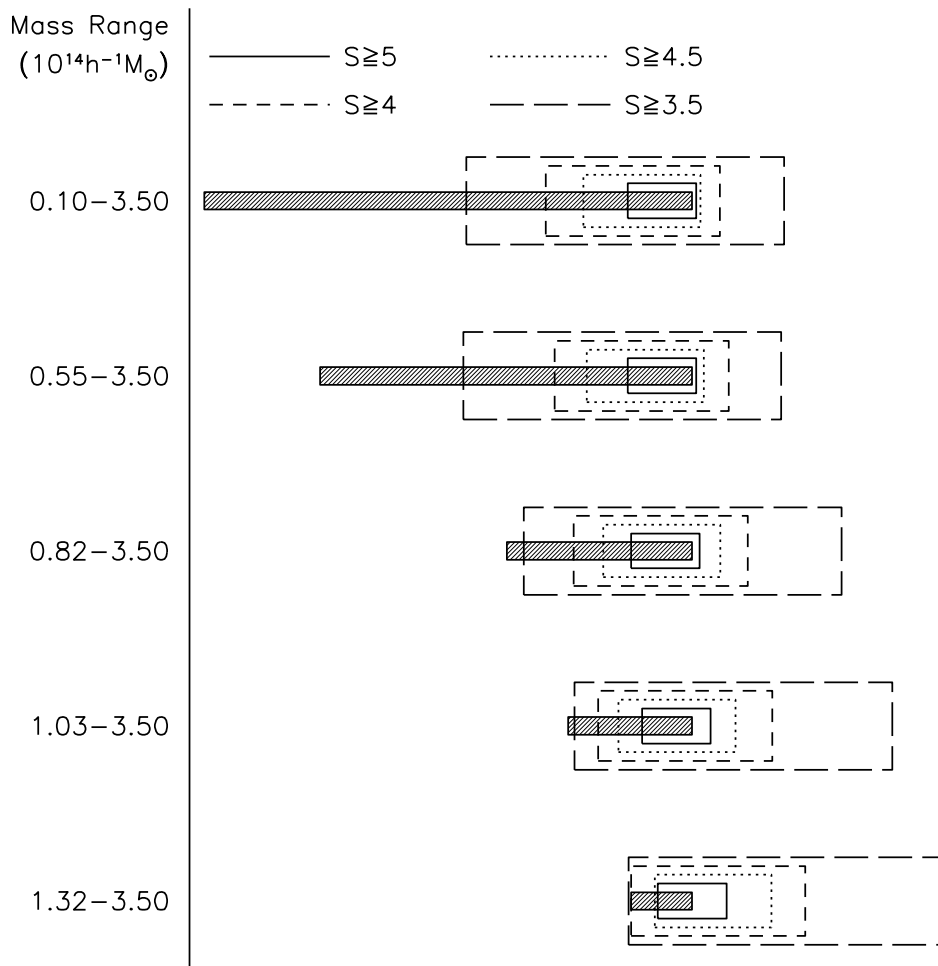
Figure 4.2: Performance of the S-statistics in identifying clusters in different mass ranges (cf. Table 4.2). See the caption of Fig. 4.1 for the meaning of the overlapping and non-overlapping parts of the bars. The different line types are for different S thresholds, as indicated in the plot. With increasing S, the fraction of spurious detections and the completeness are both reduced.

The first S-threshold investigated in detail is $S \geq 5$. This value has been advocated in the literature, e.g. by Schneider & Kneib (1998), as a signal-to-noise ratio promising significant detections. In comparison to the optical Abell criterion, the S-statistics has a similar detection rate for Abell $\mathcal{R} > 0$ like objects (53% with Abell's criterion compared to 60% with the S-statistics). The number of spuriously detected objects (50%) is identical to that for Abell's criterion in the highest mass range.

The differences between the two methods show up when detections and spurious detections at mass ranges with a lower mass threshold are considered. Looking at the detection rate of spurious 2-D objects, I see a much steeper decline as in the Abell case. For a lower mass threshold of $1.03 \times 10^{14} \, h^{-1} \, M_{\odot}$, only 27% of the S-detected clusters do not correspond to 3-D clusters of the reference set, whereas more than half of the Abell clusters in that mass range are spurious detections. For an even lower mass
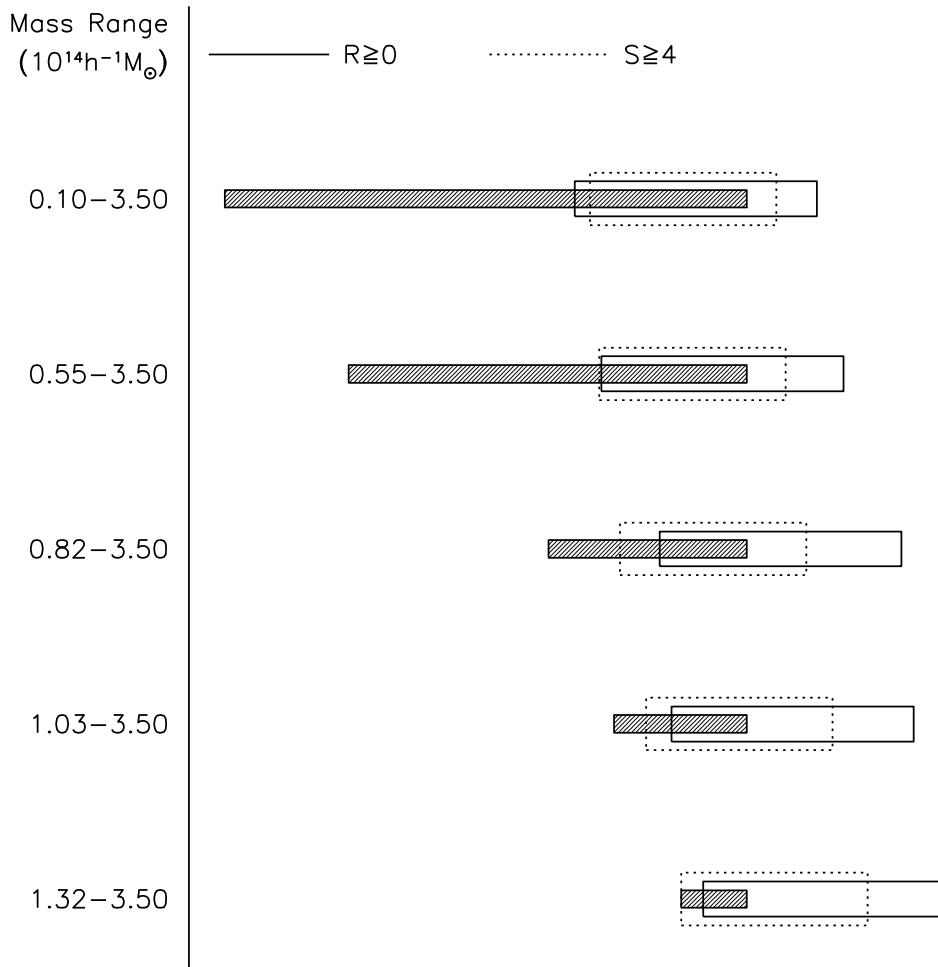
Figure 4.3: Comparison of the performances of the S-statistics and Abell's criterion in identifying clusters in different mass ranges. See the caption of Fig. 4.1 for the meaning of the overlapping and non-overlapping parts of the bars. The solid and dotted bars are for Abell-clusters with $\mathcal{R} \geq 0$ and S-selected clusters with $S \geq 4$, respectively. Evidently, the S-statistics performs better in all mass ranges in terms of spurious detections, and completeness is generally also larger.

threshold of $0.82 \times 10^{14}\, h^{-1}\, M_\odot$, the rate of spurious detections falls to only 11%, which clearly indicates that a large number of suspected spurious detections in reality corresponds to 3-dimensional matter concentrations of lower mass.

   If I reduce the S-threshold for significant detections to e.g. $S \geq 4.5$ or even below, the detection rate of 3-D clusters increases, which means that the S-cluster catalogue becomes more complete. However, the trade-off for the completeness is a higher number of spurious detections, belonging to lower-mass 3-D mass concentrations. For an S-threshold of 4.0, I am able to construct a catalogue containing *all* massive Abell-like clusters at the expense of also detecting many less massive 3-D objects. Figure 4.3 compares the performance of Abell's criterion with $\mathcal{R} \geq 0$ and the S-statistics with $S \geq 4$. It shows that S-selected cluster samples contain fewer spurious detections and are generally more complete than Abell-selected samples.

A crucial point in the application of the S-statistics is the identification of peaks in the S-map. Following Schneider (1996), I use a circular aperture for the S-statistics, which leads to an increased sensitivity for round objects. However, some of the S-maps for the simulation data show extended, non-circular areas with significant S-signals. Several of these structures contain more than one peak coming from within a plateau of high S (see Fig. 4.4). It is important to properly categorise these structures as belonging to a *single* 3-D object and to not count them twice. An example of such a situation is shown in Fig. 4.4. The contour plot shows a blow-up of the S-map around two peaks which almost overlap in the lower-resolution S-map of the whole simulation box (see the mark in Fig. 4.5). These two peaks in the S-map correspond to one of the most massive clusters in the simulation. This is reflected in the high $S = 7.6$ of the higher peak, while the second peak has $S = 3.6$. On the other hand, the sample also contains examples of 3-D clusters projected onto each other showing only a single featureless peak. Therefore, I conclude that the morphological information contained in the S-map is low. Both cases of S-signals will be discussed in more detail in Sec. 4.4.
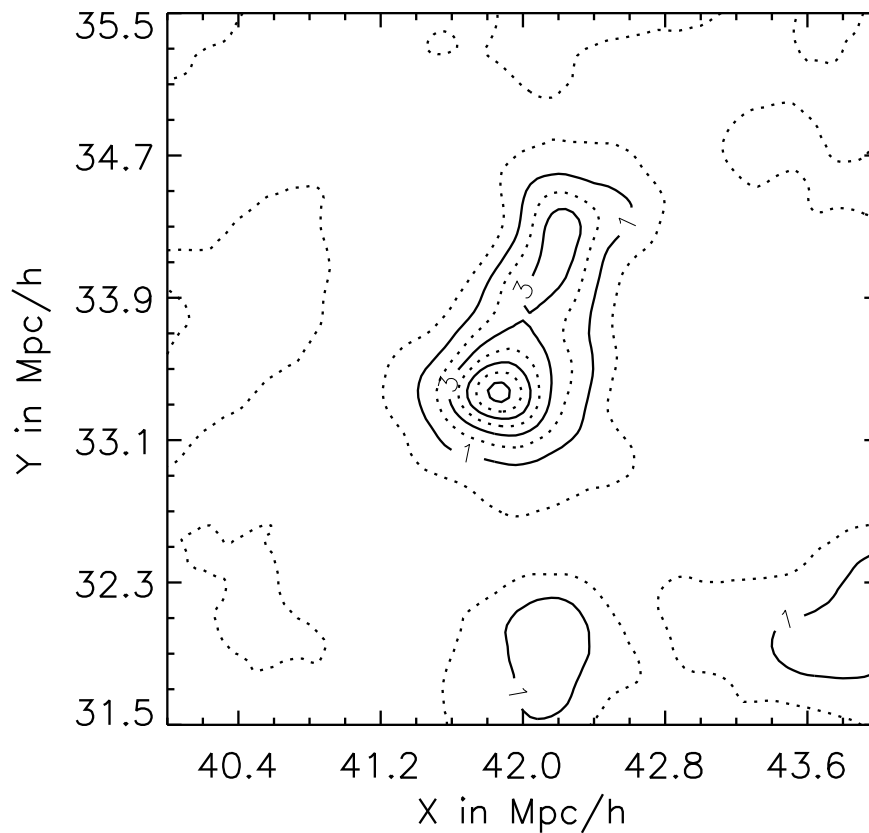

Figure 4.4: S-map for the rectangular section in Fig. 4.5 showing the double-peak structure of the map in more detail. Contours are spaced by $\Delta S = 1$. The structure features two maxima, but corresponds to a single 3-D cluster.

Summarising the quantitative results from both cluster detection methods, I can

say that the S-statistics leads to better results than Abell's criterion. The catalogues constructed with the S-statistics are more complete and suffer less from spurious detections, at least in the sense that most peaks correspond to true 3-D objects. The S-statistics evidently produces fewer truly spurious detections than Abell's criterion. However, I note that it is not possible to obtain a complete catalogue by counting only peaks with a high signal-to-noise value $S \geq 5$. There is no strict correlation between the height of a signal in the S-statistics and the mass of identified 3-dimensional objects.
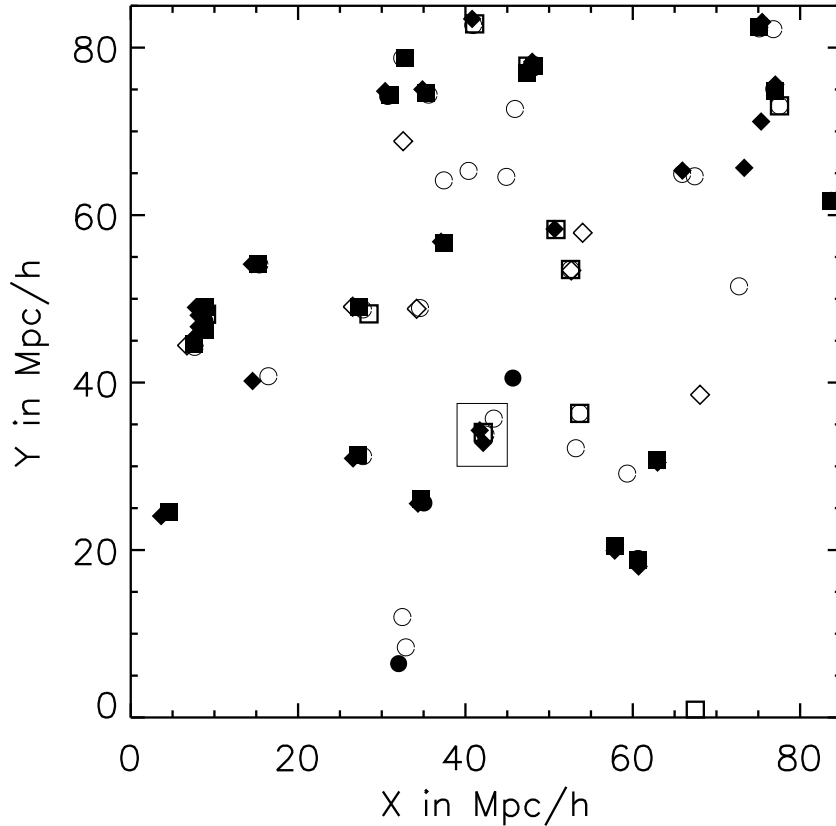
Figure 4.5: Projection of true 3-D clusters/groups (squares) along the z-axis onto the full xy-plane of the simulation box, and comparison with Abell clusters identified in projection (2-D-Abell clusters, circles), and mass-selected clusters identified with the S-statistics (diamonds). 3-D clusters with masses in the range of $(0.10\text{-}3.5) \times 10^{14} \, h^{-1} \, M_{\odot}$, 2-D clusters with $\mathcal{R} \geq 0$ (Abell), and $S \geq 3.0$ (M_{ap}) are shown. Filled symbols refer to clusters with masses $\geq 0.55 \times 10^{14} \, h^{-1} \, M_{\odot}$, $\mathcal{R} > 0$ and $S \geq 3.5$. (The offset between close symbols is deliberate to facilitate reading; it does not reflect the accuracy of measuring cluster positions.)

4.4   Structure of Representative Clusters

To achieve a deeper understanding of projection effects, I study the structure of archetypical clusters or groups along the line–of–sight. The main emphasis in this section will be on clusters selected by the S-statistics. A detailed discussion of the structure of Abell selected clusters along the line–of–sight for both real space and velocity space can be found, e.g., in Cen (1997) and van Haarlem et al. (1997).

Even though this section will concentrate on clusters detected by the S-statistics, in some cases also Abell-selected clusters will be discussed if the S-selected clusters also satisfy Abell's criterion. As Fig. 4.5 shows for one of the three projection directions, this is the case for a lot of S-selected objects, i.e. there is considerable overlap between the two selection methods. Abell's criterion detects the visible light from galaxies, while the S-statistics is sensitive to the underlying distribution of dark matter, making it possible to construct a "mass-selected" sample of clusters, as opposed to "flux-limited" samples which are obtained by observing luminous galaxies. Since this study is performed under the supposition that both selection methods detect the same physical objects, I have to assume that luminous galaxies are good tracers of the dark matter distribution. This is secured by the assumption used to populate the simulation that light follows mass.

For a more detailed analysis, the S-selected clusters will be subdivided into three classes according to the S-threshold employed. The first class considered are clusters detected with $S \geq 5$, the next class contains clusters with $5 > S \geq 4$, and the last class clusters or groups with $4 > S \geq 3$. The division into classes according to signal-to-noise values allows an investigation of systematic differences of projection effects in the different classes.

It is expected from theory that higher-mass clusters lead to larger values in the S-map. Such a trend was found in Sec. 4.3, but there is no sharp correlation between the masses of detected 3-D clusters and the threshold imposed on the S-statistics. This can be explained in terms of the intervening matter along the line–of–sight. Lower-mass 3-D objects are more prone to projection effects in the sense that the intervening matter makes up a more substantial fraction of their mass. Therefore, projection effects become more important for clusters with lower S. In the case of clusters with intrinsically lower masses, less intervening matter is needed to alter the signal in the S-map.

4.4.1 S-Statistics: $S \geq 5$

As discussed in Sec. 4.3, there is a good correspondence between $S \geq 5$ clusters and massive 3-D-clusters. Investigating the line–of–sight structure of these 2-D-clusters, I can generally state that nearly all of them show a high, pronounced peak in the position histogram at the position of the true 3-D cluster. Even though the amount of contamination with intervening matter in this group is only moderate, some velocity histograms deviate significantly from a Gaussian shape.

The cluster given in Sec. 4.3 as an example for a cluster exhibiting a double-peaked S-map (see Fig. 4.4) clearly belongs into this class, since the main peak has $S = 7.3$, while the nearby second peak has only $S = 3.6$. The structure along the line–of–sight in both real space and velocity space is shown in Fig. 4.6. The position histogram is characterised by a dominant peak at the position of the corresponding 3-D cluster which

has a mass of $M_{3-\mathrm{D}} = 3.4 \times 10^{14}\, h^{-1}\, M_\odot$. The amount of dark matter along the line–of–sight is moderate with only two small clumps $40\, h^{-1}$ Mpc and $65\, h^{-1}$ Mpc behind the main clump, which both have masses smaller than $10^{13}\, h^{-1}\, M_\odot$. Although the 3-D cluster has Abell richness class $\mathcal{R} = 0$, the projected cluster satisfies the 2-D Abell criterion for $\mathcal{R} = 1$, indicating an inflation of richness class. The velocity dispersion $\sigma_{3-\mathrm{D}} = 775\ \mathrm{km\,s^{-1}}$ of the 3-D cluster is reduced to $\sigma_{2-D} = 676\ \mathrm{km\,s^{-1}}$ in projection, hinting at an asymmetric velocity ellipsoid of the 3-D cluster.
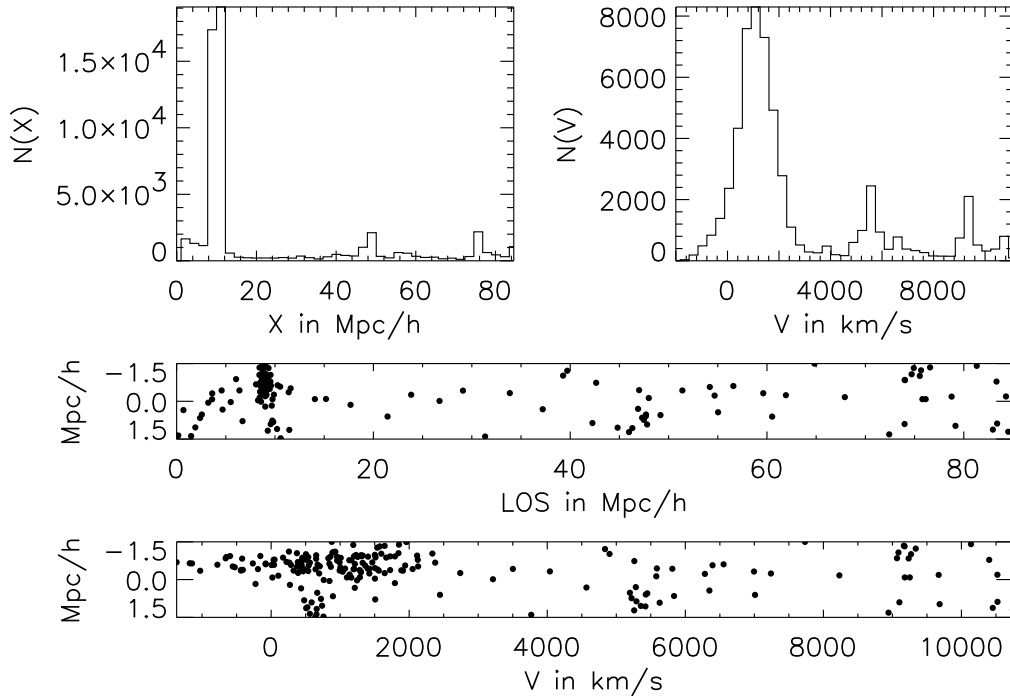
Figure 4.6: Structure of a large 3-D cluster with moderate contamination, detected as a double peak with $S = 7.3$ and $S = 3.6$ (cf. Fig. 4.4). It is also an Abell $\mathcal{R} = 1$ cluster, both in projection and in 3-D. The upper left panel displays the histogram of the dark matter distribution in real space along the line–of–sight, and the upper right panel shows the corresponding velocity histogram. The two lower panels show the dark matter distribution along a cylinder of radius $1.5\, h^{-1}$ Mpc in real space (along the line–of–sight) and in velocity space. For better display in both lower panels, a random fraction of the dark matter particles is omitted.

Although the cluster is only moderately affected by projection effects, the *los* velocity histogram strongly deviates from a Gaussian, which is also true for the velocity distribution of the 3-D-cluster alone. The deviation from Gaussianity, as measured by higher order moments of the distribution like the skewness $\mathcal{S}$ and the curtosis $\mathcal{K}$, is $\mathcal{S}_{3-\mathrm{D}} = -0.62$ and $\mathcal{K}_{3-\mathrm{D}} = 0.47$ for the 3-D-cluster as opposed to $\mathcal{S}_{2-\mathrm{D}} = 1.57$ and $\mathcal{K}_{2-\mathrm{D}} = 3.31$ in projection. The substantial increase of skewness and curtosis in projection emphasises the influence of projection effects in velocity space. Together with the decrease of the velocity dispersion and the increase of richness class in projection, this hints at the presence of non-virialised sub-clumps in the vicinity of the main clus-

ter. Yet, this detection corresponds to a very massive 3-D cluster. More examples are given in Figs. A.1 and A.2 in the Appendix.

4.4.2 S-**Statistics:** $5 > S \geq 4$

The next sample of S-detected clusters has lower signal-to-noise; but, as has been shown in Sec. 4.3, the mass-selected cluster catalogue becomes more complete if these clusters are included. Most of them correspond to true 3-D clusters. Even though these detections are significant and contain some massive clusters, their amount of contamination in relation to their main 3-D cluster is generally larger than for clusters detected with larger S. One typical example of this class is shown in Fig. 4.7, two more are given in Figs. A.3 and A.4 in the Appendix.
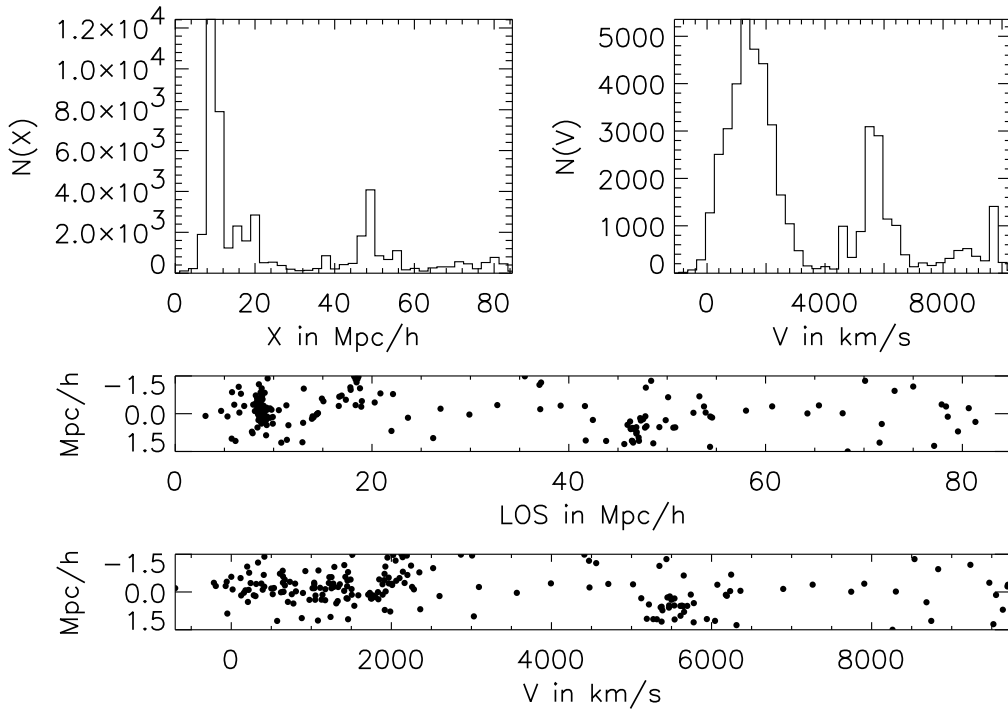
Figure 4.7: Structure of a moderately contaminated 3-D cluster with $S = 4.7$. See the caption of Fig. 4.6 for a description of the panels.

The cluster in Fig. 4.7 is a high-mass cluster with $M_{3-\mathrm{D}} = 1.5 \times 10^{14}\, h^{-1}\, M_{\odot}$, a clump nearby, and a second mass clump $30\, h^{-1}$ Mpc away. It is detected at $S = 4.7$. The velocity dispersion of the projected cluster is broadened from $\sigma_{3-\mathrm{D}} = 697\,\mathrm{km\,s^{-1}}$ to $\sigma_{2-\mathrm{D}} = 865\,\mathrm{km\,s^{-1}}$, and the projected velocity distribution has a bimodal shape. The higher-order moments of the velocity distribution indicate this through skewness and curtosis in projection ($\mathcal{S}_{2-\mathrm{D}} = 1.85$ and $\mathcal{K}_{2-\mathrm{D}} = 3.27$), compared to 3-D ($\mathcal{S}_{3-\mathrm{D}} = 0.18$ and $\mathcal{K}_{3-\mathrm{D}} = -0.40$). The cluster satisfies Abell's criterion with richness class $\mathcal{R} = 0$, both in projection and in 3-D.

4.4.3 S-Statistics: $4 > S \geq 3$

The last class considered contains clusters with S between 3 and 4. This class is mainly discussed for reasons of completeness. Clusters identified with such signal-to-noise correspond to lower-mass objects making projection effects along the line–of–sight more important. This can be seen looking at the three examples in Fig. 4.8 and in Figs. A.5 and A.6.
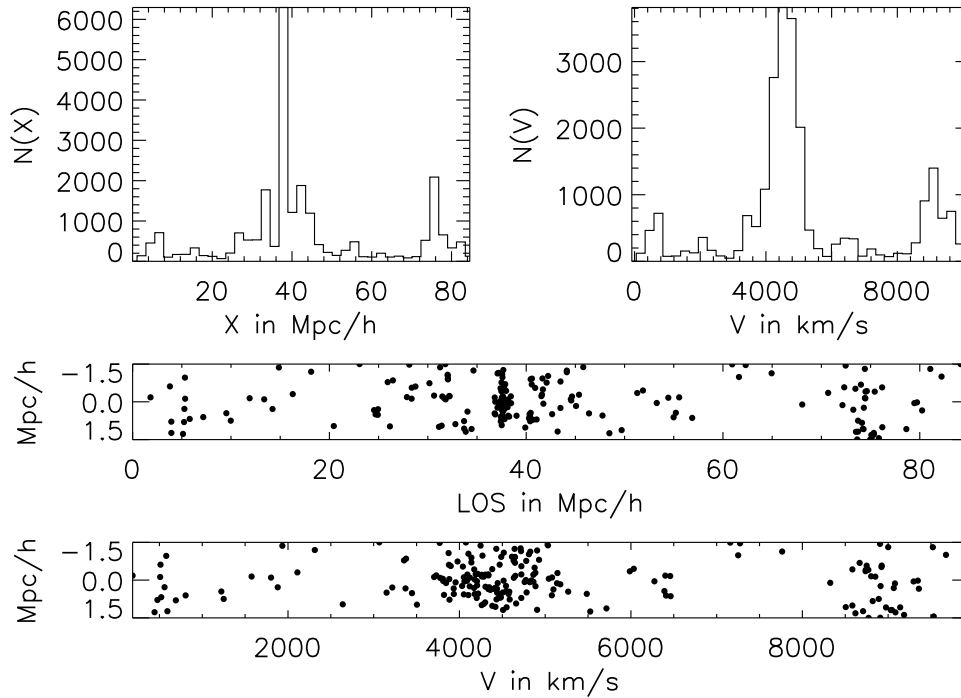
Figure 4.8: Line–of–sight structure of a highly contaminated cluster. This cluster is detected at $S = 3.2$. See the caption of Fig. 4.6 for a description of the panels.

The peak with $S = 3.2$ corresponding to the cluster shown in Fig. 4.8 is due to a 3-D cluster with a rather low mass only, $M_{3-\mathrm{D}} = 5.9 \times 10^{13} \, h^{-1} \, M_{\odot}$. Compared to the low 3-D cluster mass, there is a high amount of intervening matter with several mass clumps along the line–of–sight. Because of this substantial contamination, the projected velocity dispersion is overestimated; $\sigma_{2-\mathrm{D}} = 669 \, \mathrm{km \, s^{-1}}$ compared to $\sigma_{3-\mathrm{D}} = 364 \, \mathrm{km \, s^{-1}}$ for the 3-D cluster. The velocity histogram has a second peak at the high-velocity tail of the distribution. The higher-order moments of the velocity distribution of the 3-D cluster ($\mathcal{S}_{3-\mathrm{D}} = -0.05$ and $\mathcal{K}_{3-\mathrm{D}} = -0.07$) change by a large amount when looking at the projected velocity distribution ($\mathcal{S}_{2-\mathrm{D}} = 1.12$ and $\mathcal{K}_{2-\mathrm{D}} = 2.43$). This reflects the large influence the intervening matter exerts on observation.

4.5   Mass Estimates

The previous two sections put emphasis on the completeness of catalogues constructed with two different selection criteria, Abell's criterion and the S-statistics based on weak

gravitational lensing. Furthermore, I investigated the structure of some detected clusters along the line–of–sight. Of course, both kinds of information are important when deriving statistical information from such catalogues. A third very important test for cosmological theories are the different mass estimates and their relationship with each other. A mass estimate closely related to the optical selection derives from the virial theorem. As a gravitational-lensing based mass estimate, I choose the ζ-statistics, which is closely related to the S-statistics as demonstrated in Sec. 4.2.3.

Under the assumption that clusters of galaxies are bound self-gravitating systems in dynamical equilibrium, the total cluster mass can be estimated via the virial theorem (Binney & Tremaine 1987; Sarazin 1986),

$$M_{\text{tot}} = \frac{R_{\text{G}} \langle v^2 \rangle}{G} \, , \tag{4.14}$$

where R_{G} is the gravitational radius of the cluster relating the system's mass to its potential energy. Gunn & Gott (1972) showed that this radius is approximately given by r_{200}, the radius of a sphere containing an overdensity of $200 \, \rho_{\text{crit}}$. For the clusters of my study, a radius of $R_{\text{G}} = 0.75 \, h^{-1}$ Mpc is a good approximation. Observationally, only the *los* velocity dispersion $\sigma_{\parallel}$ can be measured. Assuming isotropic orbits, the two quantities are related by $\langle v^2 \rangle = 3 \sigma_{\parallel}^2$.

When calculating the radial velocities from simulated data, the Hubble flow has to be added to the peculiar velocities of the simulation. Since the simulation data are at high redshift, the dependence on redshift of the cosmological parameters also has to be taken into account. Therefore, the radial velocities are given by

$$v_{\parallel} = \dot{a}(z) \, x_{\parallel} + a(z) \, \dot{x}_{\parallel} = a(t) \left[H(z) \, x_{\parallel} + \dot{x}_{\parallel} \right] , \tag{4.15}$$

with the expansion factor

$$a(z) = (1 + z)^{-1} \tag{4.16}$$

and the Hubble parameter

$$H(z) = H_0 \, (1 + z)^{3/2} \, . \tag{4.17}$$

Apart from the validity of the virial theorem, the virial mass estimate depends solely on a correct estimate of the velocity dispersion $\sigma_{\parallel}$. Since the velocity dispersion is very sensitive to field galaxies and small sub-clumps projected onto the main clusters, it is important to remove these from the sample. I convolve the velocity histogram with a $4000 \, \text{km s}^{-1}$ wide top-hat filter to reject interlopers, i.e. galaxies with relative velocities greater than $4000 \, \text{km s}^{-1}$ from the peak of the convolved histogram are removed. I further employ the so-called 3σ-clipping procedure proposed by Yahil & Vidal (1977) which has widely been applied to observational samples. It can be summarised as follows:

1. compute the mean radial velocity;

2. remove the galaxy which deviates most from the mean of the sample and redetermine the mean without this galaxy;

3. if the removed galaxy deviates from the new mean by more than 3σ, it is removed from the sample;

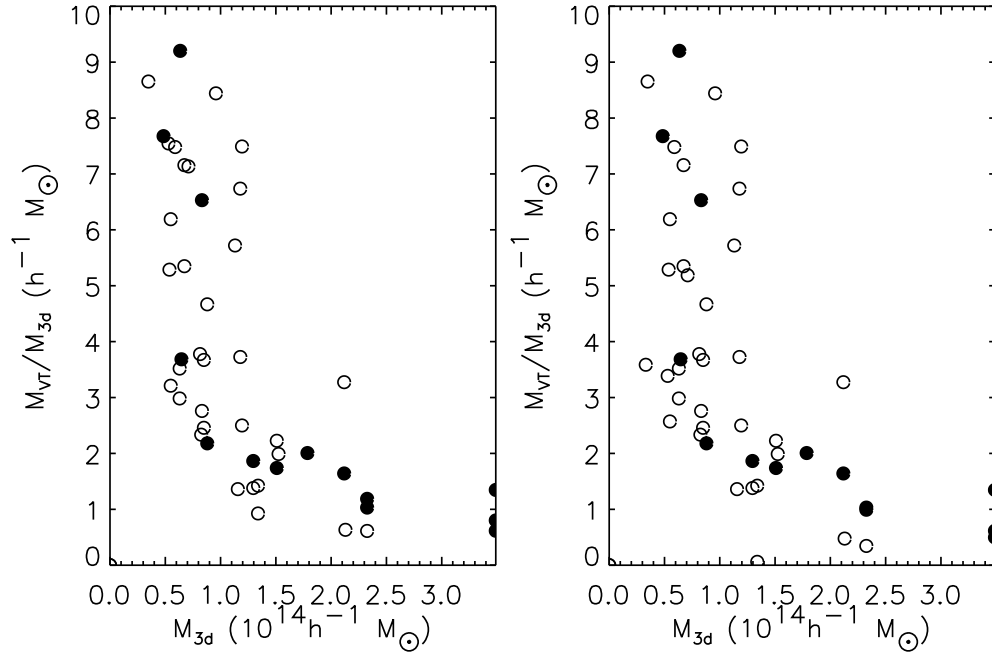4. repeat the procedure until the last tested galaxy remains in the sample.

Figure 4.9: Comparison of the virial mass estimate M_{VT} to the true cluster mass $M_{3-\mathrm{D}}$. Both panels of this figure display the ratio $M_{\mathrm{VT}}/M_{3-\mathrm{D}}$ as a function of the true cluster mass $M_{3-\mathrm{D}}$. In the left panel, the velocity dispersions before 3σ clipping are used to compute the virial-theorem based mass estimate, while in the right panel, the velocity dispersions after 3σ clipping are used. In both panels, the open circles refer to $\mathcal{R} = 0$ clusters and the filled circles to $\mathcal{R} = 1$ clusters.

Figure 4.9 displays the correlation of the virial mass estimate with the true mass of the corresponding 3-D cluster. The left and right panels show the ratio of the virial mass with the 3-D cluster mass as a function of the 3-D mass before and after 3σ clipping, respectively.

The first thing to notice is that the masses of clusters with richness class $\mathcal{R} > 0$ are less severely overestimated than masses of clusters with lower richness. This holds true for the mass estimates before and after 3σ-clipping for both the mean and the median, as can be seen in Table 4.3. The second thing readily seen in Fig. 4.9 and Table 4.3 is the large dispersion of the underlying distribution. This dispersion is smaller for clusters with higher richness class than for clusters with the lowest richness class $\mathcal{R} = 0$ considered. I also note that this dispersion is hardly affected by the 3σ-clipping procedure. The only effect of the clipping procedure is to reduce the average of the estimated cluster masses irrespective of the richness class. A third trend to be seen in Fig. 4.9 is that the overestimation of the mass is generally more severe for 3-D clusters with lower

Table 4.3: Comparison of statistical parameters for the ratio $M_{\mathrm{VT}}/M_{3-\mathrm{D}}$ and $M_\zeta/M_{3-\mathrm{D}}$ for different sub-samples of masses estimated via the velocity dispersion and the ζ-statistics.

sample	mean	median	standard deviation
complete optical sample before 3σ clipping	4.90	3.52	1.91
Abell cluster $\mathcal{R} = 1$ before 3σ clipping	3.84	1.86	1.10
Abell cluster $\mathcal{R} = 0$ before 3σ clipping	5.44	3.73	2.67
complete optical sample after 3σ clipping	4.53	3.39	2.01
Abell cluster $\mathcal{R} = 1$ after 3σ clipping	3.58	1.86	1.18
Abell cluster $\mathcal{R} = 0$ after 3σ clipping	5.01	3.67	2.80
complete lensing sample	1.27	1.05	0.34
$S \geq 5$	1.23	1.13	0.34
$5 \geq S \geq 4$	1.32	1.02	0.31

mass. For the most massive clusters in the sample ($M_{3-D} \lesssim 2.0 \times 10^{14}\, h^{-1}\, M_\odot$), the continuation of this trend in some cases leads to an *underestimation* of the masses, as can be seen in the right-hand side of each panel in Fig. 4.9. The 3σ clipping procedure fails to correct for the mass overestimates. When comparing Fig. 4.9 and Table 4.3 to Fig. 15 of Cen (1997), one has to keep in mind the different selection procedure for clusters or groups of galaxies in both studies, but on the whole the results are consistent.

The behaviour described above can largely be attributed to the influence of projection effects on the velocity dispersion. Generally, the inclusion of field galaxies and unvirialised sub-clumps broadens the distribution and leads to distributions which deviate significantly from Gaussian shape, as illustrated by the examples in Sec. 4.4. The clipping procedure is successful when the amount of contamination along the line–of–sight is low or moderate, but the algorithm fails to remove larger sub-clumps projected onto the main cluster which can significantly broaden the distribution, sometimes even making it bimodal. In some cases it is possible that the clipping procedure removes galaxies belonging to the 3-D cluster, thus contributing to an underestimation of the mass.

Even though I expect from the studies of Frenk et al. (1990) and van Haarlem et al. (1997) that the high-velocity tail of the velocity distribution is severely overestimated, the effect on the mass estimate is most pronounced for galaxy clusters with lower mass. This is due to the fact that they are more easily overestimated with respect to their true dispersion.

The ζ-statistics as compared to Abell's criterion leads to smaller overestimates of the 3-D cluster masses as shown in Fig. 4.10. Interpreting the quantitative results of the ζ-statistics mass estimate, one has to keep in mind two competing effects: On the one hand, the ζ-statistics, like every gravitational-lensing based method, measures all the mass along the line–of–sight to the cluster; on the other hand, it gives a lower bound to the cluster mass. In combination, these two competing effects lead to fairly moderate
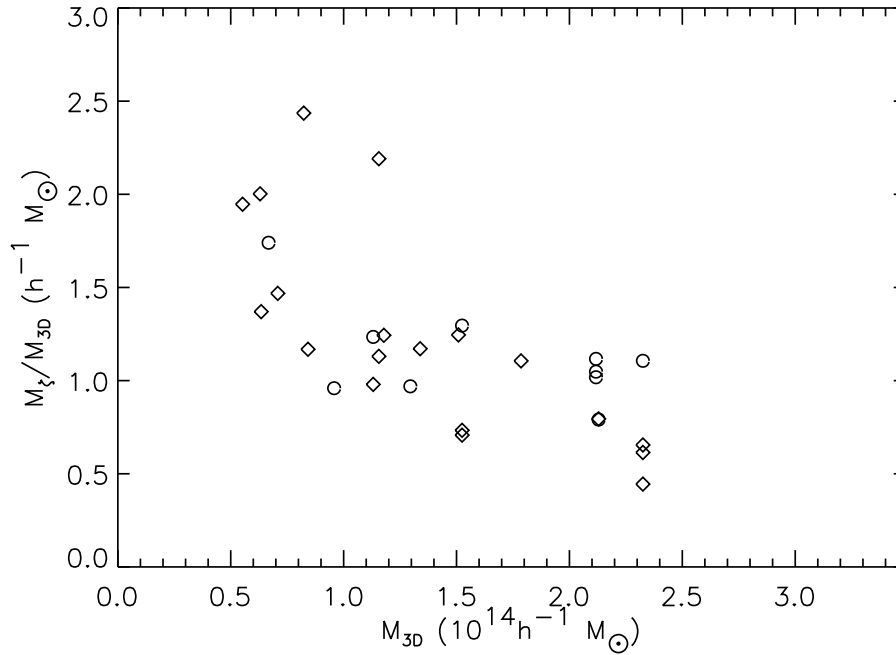
Figure 4.10: Comparison of the ζ-statistics mass estimate to the true cluster mass $M_{3-\mathrm{D}}$. Diamonds and circles refer to clusters identified with a $S \geq 5$ and $S \geq 4$, respectively. Clusters detected at lower S are excluded because of their high contamination.

mass overestimates, as can be seen in Table 4.3. This also explains the difference to the lensing mass estimates given in the paper by Cen (1997). There, all masses along the line–of–sight are added up under the assumption of a perfect lensing reconstruction method with an otherwise calibrated mass-sheet degeneracy.

The other interesting feature in Fig. 4.10 and Table 4.3 is the low dispersion of the underlying distribution. This dispersion does not depend sensitively on the S-value at which the clusters are detected. (Clusters detected with $S < 4$ where excluded here because of their large contamination.) The dispersion is typically less than a third of the dispersion in the Abell samples.

As for clusters detected with Abell's criterion, masses of small 3-D clusters are more strongly overestimated than for more massive 3-D clusters. This is due to the fact that the proportion of contaminating matter to the 3-D cluster mass is higher for less massive 3-D objects than for the extremely massive objects. For the intermediate-mass objects, the fact that the ζ-statistics only gives lower bounds to the masses partially outweighs this effect. There, the masses are even slightly underestimated.

Investigating the relationship between velocity-based mass estimates and the gravitational lensing based ζ-statistics in Fig. 4.11, I see that the ζ-statistics gives on average smaller estimates of the 3-D cluster masses than the virial theorem. Again I stress that this is due to the fact that the ζ-statistics is derived under the assumption of an empty outer annulus, restricting it to estimate lower bounds to the masses. The dispersion between the ratio of ζ-statistics mass to virial mass is large, which is mainly due to the large underlying dispersion in the virial mass estimate.

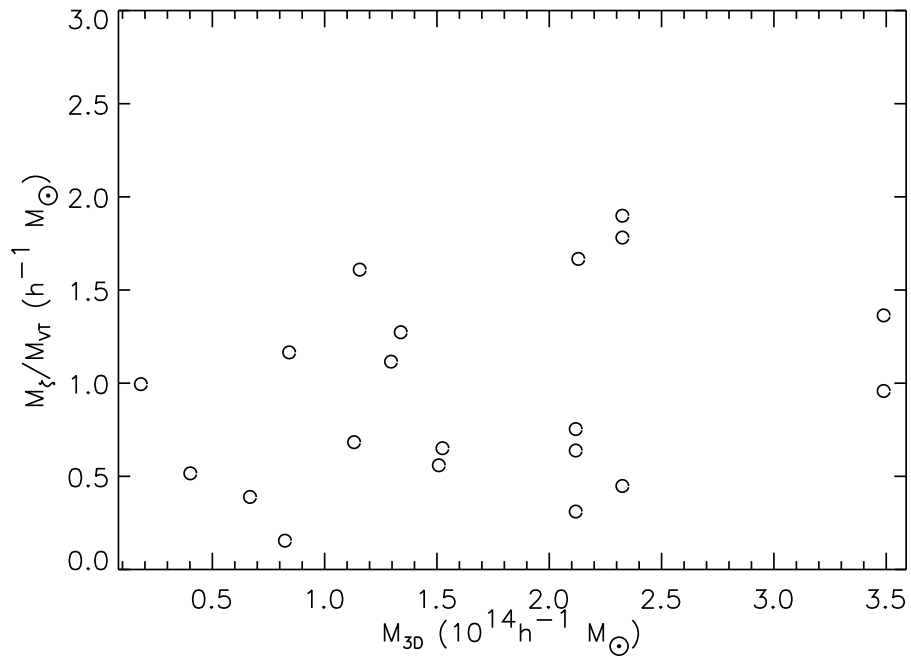Figure 4.11: Comparison of the ζ-statistics mass estimate to the virial mass estimate M_{VT}. Only such clusters are included which are detected with both methods, Abell's criterion and S-statistics. Only clusters with $S > 4$ are considered.

Chapter 5

Cluster Deprojection

In the following I propose an algorithm for the deprojection of observed cluster images. This algorithm is based on the Richardson-Lucy algorithm for the rectification of observed distributions, and it combines multiple sets of observable data from clusters of galaxies — weak lensing (lensing potential ψ), X–ray (X–ray surface brightness S_x), and Sunyaev–Zel'dovich (temperature decrement ΔT_{SZ}) data — to obtain information on the structure along the line–of–sight, namely the 3-dim. gravitational potential φ. The combination of multiple data sets allows one to exploit the different dependences of the various observable distributions on the gravitational potential along the line–of–sight (los).

For deriving my *multiple–data Richardson-Lucy deprojection algorithm* I first describe the mathematical background of the general Richardson-Lucy (hereafter RL) approach, then I specify a geometrical model for the cluster that is suitable for deriving a RL-type deprojection equation for the gravitational potential φ, and finally I show how the three above–mentioned observable distributions can be incorporated into this deprojection procedure. In addition, I discuss strategies for implementing the algorithm into computer programs, study their respective numerical stability, and assess their performance by applying them to synthetic cluster data from gas–dynamical simulations. In the end, I give an outlook on the suitability of the algorithm for practical applications to true observational data.

5.1   Richardson–Lucy Algorithm

The question of how to deproject observed cluster images is a prime example for so-called *inverse problems*, which often arise in astronomy. Inverse problems in astronomical research reflect the fact that it is in general not possible to directly measure the quantities of interest due to the large distances between observers and studied objects. Furthermore, the theoretical understanding of the relevant physical phenomena is often so limited or the problem is so complex, that it is not possible to derive appropriate models from first principles.

Most inverse problems in astronomy can be cast into the form of a Fredholm integral equation of the first kind,

$$\psi(x) = \int \phi(\zeta) P(x|\zeta) \, \mathrm{d}\zeta, \qquad (5.1)$$

where $\phi(\zeta)$ is the function of interest, $\psi(x)$ is the function accessible through observation, and the integral kernel $P(x|\zeta)$ reflects the measurement process. In general, ϕ and ψ represent probability density functions, which implies that they and the kernel $P(x|\zeta)$ obey the normalization and non-negativity constraints:

$$\int \psi(x)\,\mathrm{d}x \;\; = \;\; 1, \qquad \text{and} \qquad \psi(x) \geq 0, \tag{5.2a}$$

$$\int \phi(\zeta)\,\mathrm{d}\zeta \;\; = \;\; 1, \qquad \text{and} \qquad \phi(\zeta) \geq 0, \tag{5.2b}$$

$$\int P(x|\zeta)\,\mathrm{d}\zeta \;\; = \;\; 1, \qquad \text{and} \qquad P(x|\zeta) \geq 0, \tag{5.2c}$$

where $P(x|\zeta)\mathrm{d}x$ is the probability – presumed to be known – that x' will fall in the interval $(x, x + \mathrm{d}x)$ when it is known that $\zeta' = \zeta$.

The naive approach, to solve the integral equation (5.1) directly, in most cases leads to very poor results. The reason for this is that the observed distribution $\psi(x)$ is not exactly known, instead it is a measured quantity $\tilde{\psi}$ having sampling errors. Often the estimates $\tilde{\psi}$ for ψ are not continuous and also violate the non–negativity constraint (5.2a). For solving the integral equation directly, this is a very unfavourable condition as the function $\psi(x)$ is generally smoother than $\phi(\zeta)$ due to the folding with the kernel $P(x|\zeta)$. Thus, solving for $\phi(\zeta)$ greatly magnifies the statistical fluctuations present in $\tilde{\psi}(x)$, which is an observed quantity, thus being hampered by measurement errors.

Richardson (1972) and Lucy (1974, 1994) recognized that the problem is statistical in nature and proposed an iterative inversion algorithm. Their algorithm can be derived via Bayes' theorem for conditional probabilities: Let $Q(\zeta|x)\mathrm{d}\zeta$ be the ("inverse") probability that $\zeta' \in (\zeta, \zeta + \mathrm{d}\zeta)$ when it is known that the measured quantity is $x = x'$. The probability that $x' \in (x, x + \mathrm{d}x)$ and $\zeta' \in (\zeta, \zeta + \mathrm{d}\zeta)$ is then given by $\psi(x)\mathrm{d}x \times Q(\zeta|x)\mathrm{d}\zeta$. On the other hand for $\zeta' \in (\zeta, \zeta + \mathrm{d}\zeta)$ and $x' \in (x, x + \mathrm{d}x)$ to occur, the probability is $\phi(\zeta)\mathrm{d}\zeta P(x|\zeta)\mathrm{d}x$. Equating these two expressions leads to

$$Q(\zeta|x)\psi(x) = P(x|\zeta)\phi(\zeta), \tag{5.3}$$

which can be solved for $Q(\zeta|x)$,

$$Q(\zeta|x) = \frac{\phi(\zeta)P(x|\zeta)}{\psi(x)} = \frac{\phi(\zeta)P(x|\zeta)}{\int \phi(\zeta)P(x|\zeta)\mathrm{d}\zeta}, \tag{5.4}$$

where we have used Eq. (5.1) to substitute for ψ. Notice that Eq. (5.4) is simply *Bayes' theorem* for conditional probabilities. Going back to Eq. (5.3), we can also integrate over x, employ the normalization condition Eq. (5.2c) for the kernel $P(x|\zeta)$ and finally obtain

$$\phi(\zeta) = \int \psi(x)Q(\zeta|x)\,\mathrm{d}x, \tag{5.5}$$

which formally resembles the inverse of the integral equation (5.1), if $Q(\zeta|x)$ is identified with the reciprocal kernel.

Equation (5.5) may not be used to calculate the desired quantity $\phi(\zeta)$, since the reciprocal kernel $Q(\zeta|x)$ itself is a functional of ϕ through Eq. (5.4). Since, however,

the integral kernel $P(x|\zeta)$ is assumed to be known, Bayes' theorem (5.4) and the identity (5.5) suggest an iterative solution of Eq. (5.1) for $\phi(\zeta)$: Starting with an intial guess ϕ_0 for ϕ, an estimate for $Q(\zeta|x)$ is calculated via Eq. (5.4), and this is in turn combined with the observed sample $\tilde{\psi}(x)$ to obtain an improved estimate for the theoretical distribution $\phi(\zeta)$ via Eq. (5.5).

The iterative algorithm thus consists of the following two steps:

$$\psi_n(x) = \int \phi_n(\zeta)P(x|\zeta)\mathrm{d}\zeta, \tag{5.6a}$$

$$\phi_{n+1}(\zeta) = \phi_n(\zeta)\int \frac{\tilde{\psi}(x)}{\psi_n(x)}P(x|\zeta)\mathrm{d}x. \tag{5.6b}$$

The iteratively constructed functions ϕ_n satisfy the constraints (5.2b): From Eq. (5.6b) it follows that $\phi_n \leq 0$ if $\phi_0 \leq 0$. The normalization constraint is fulfilled, as one can proof by integrating Eq. (5.6b) with respect to ζ and using the normalizations of the probabilities $Q_n(\zeta|x)\mathrm{d}\zeta$ and $\tilde{\psi}_n\mathrm{d}x$.

Ideally we would like an iterative algorithm to converge to the exact solution, and from Eq. (5.6) it can be seen that the above scheme converges if $\psi_n(x)$ is sufficiently close to $\tilde{\psi}(x)$ for all points x, except for those in a set of zero measure. However, this inherent convergence criterion is much to strong for practical purposes where $\tilde{\psi}(x)$ may be contaminated by non-negligible measurement errors. Looking again at the two coupled equations (5.6a) and (5.6b), we see that deviations of $\tilde{\psi}/\psi_n$ from unity on a length scale large compared to that of $P(x|\zeta)$ are removed in essentially one iteration, whereas deviations on a small length scale are mostly averaged out when folded with $P(x|\zeta)$, and result only in small corrections to ϕ_n. Thus the algorithm has the property of first fitting the large–scale differences between the given initial guess and the true solution, while it fits the small–scale fluctuations only in later iteration steps. Under the reasonable assumption that the small scale fluctuations are more likely to be caused by statistical errors in $\tilde{\psi}$, this behaviour of the RL algorithm is indeed highly desirable. It ensures that the RL procedure very quickely results in an approximate solution in which most of the significant information in the observed $\tilde{\psi}$ is already recovered. But one has to keep in mind that there is *no obvious* convergence criterion for recovering the large–scale fluctuations only; in general it is not easy to say when one starts to fit small–scale statistical fluctuations. Thus, for the algorithm to work, one has to know and to control the errors in the measured data very well.

It is possible to establish a relationship between the iterative algorithm introduced above and a corresponding maximum–likelihood method. This was done by Lucy in 1974, of which I here only quote the main result. It can be shown that when the integrals in Eq. (5.6) are approximated by sums, i.e. $\phi(x) \rightarrow \phi_j$, and when only the fraction ϵ of the correction to ϕ_n is actually applied, the scheme converges for $n \rightarrow \infty$ to a solution of the corresponding maximum likelihood (ML) problem, provided that ϵ is sufficiently small. Lucy also demonstrated that the ML solution is, in general, unique, and that the direct solution is identical to the corresponding ML-problem, if it does not violate the constraint $\phi_j \leq 0$.
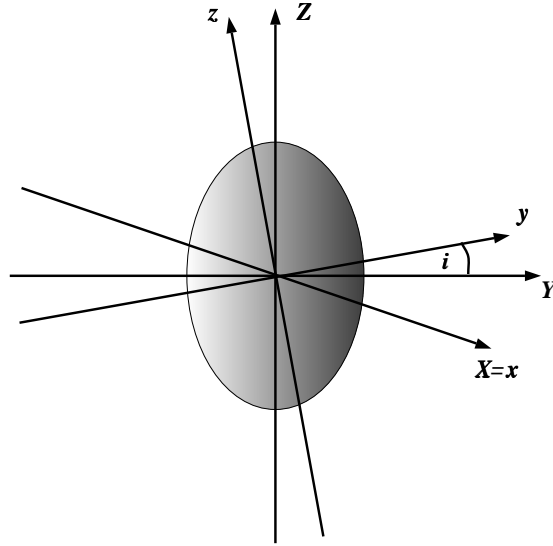
Figure 5.1: Display of the relation between cluster coordinates (X, Y, Z) and observer's coordinates (x, y, z) assuming axial symmetry.

5.2 Richardson-Lucy Algorithm: The Axisymmetric Case

Consider a cluster of galaxies covered by a system of cartesian coordinates (X, Y, Z). We are interested in recovering the distribution of some physical quantity $\phi(X, Y, Z)$, which we assume to have axial symmetry with respect to the Z-axis of the cluster coordinate system:

$$\phi(X, Y, Z) \equiv \phi(R, Z) \qquad \text{with} \qquad R^2 = X^2 + Y^2. \qquad (5.7)$$

Furthermore, we assume that the projection of ϕ is observed as some quantity ψ, where the observer's coordinate system (x, y, z) is inclined by an angle i, and in which $x = X$ and z is the line–of–sight (see Fig.5.1). The transformation between cluster coordinates (X, Y, Z) and observer's coordinates (x, y, z) is thus given by

$$X = x, \qquad (5.8a)$$
$$Y = y \cos(i) - z \sin(i), \qquad (5.8b)$$
$$Z = y \sin(i) + z \cos(i), \qquad (5.8c)$$

and the observed projection ψ is given as the los integral over ϕ,

$$\psi(x, y) = \int \mathrm{d}z \, \phi(R(x, y, z), Z(x, y, z)). \qquad (5.9)$$

If we want to apply the RL-algorithm to recover ϕ from ψ, we first have to bring the fundamental los integral (5.9) into the form of Eq. (5.1),

$$\psi(x, y) = \int \mathrm{d}R \int \mathrm{d}Z \, P(x, y | R, Z) \phi(R, Z). \qquad (5.10)$$

where $P(x,y|R,Z)$ has to be normalized to unity. In order to derive the kernel $P(x,y|R,Z)$ we start with the following identity for ϕ,

$$\phi(R,Z) = \int_0^\infty dR'^2 \int_{-\infty}^\infty dZ' \quad \phi(R',Z')\delta(R^2 - R'^2)\delta(Z - Z'), \tag{5.11}$$

which we insert into Eq. (5.9) and obtain

$$\psi(x,y) = \int_{-\infty}^\infty dz \int_0^\infty dR'^2 \quad \delta[R^2(x,y,z) - R'^2]$$

$$\times \int_{-\infty}^\infty dZ'\phi(R',Z')\delta[Z(x,y,z) - Z']. \tag{5.12}$$

Next, we simplify $\delta(Z - Z')$ by substituting the coordinate transformation (5.8c) for Z and applying a computational rule for δ-functions [1]. Here we have

$$f(z_s) = y\sin i + z_s \cos i - Z' \overset{!}{=} 0 \quad \Rightarrow \quad z_s = \frac{Z' - y\sin i}{\cos i}, \tag{5.13}$$

provided that $\cos i \neq 0$. Here and in the following we can restrict the discussion to the inclination angles in the range of $0 \leq |i| < \pi/2$, leading to non-zero, positive values for $\cos i$. Together with $f'(z) = \cos i$ and the corresponding substitution (5.7), (5.8) for $R(x,y,z)$ we arrive at

$$\psi(x,y) = \int_0^\infty dR'^2 \int_{-\infty}^\infty dZ' \int_{-\infty}^\infty dz \quad \frac{\phi(R',Z')}{\cos i}$$

$$\times \delta\left[x^2 + (y\cos i - z\cos i)^2 - R'^2\right] \delta\left[z - \frac{Z' - y\sin i}{\cos i}\right]. \tag{5.14}$$

Finally, integration of the second δ-function over z yields

$$\psi(x,y) = \pi \int_0^\infty dR'^2 \int_{-\infty}^\infty dZ' \, \phi(R',Z')$$

$$\times \frac{\delta\left[(y\cos i - (Z' - y\sin i)\tan i)^2 - (R'^2 - x^2)\right]}{\pi \cos i}, \tag{5.15}$$

which is in the desired form of Eq. (5.10). We can identify the kernel

$$P(x,y|R,Z) = \frac{\delta[(\frac{y}{\cos i} - Z\tan(i))^2 - (R^2 - x^2)]}{\pi \cos i}, \tag{5.16}$$

and find that it is properly normalized:

$$\int_{-\infty}^\infty dx \int_{-\infty}^\infty dy \, P(x,y|R,Z) = 1. \tag{5.17}$$

Having the explicit expression (5.16) for the kernel $P(x,y|R,Z)$, we could now start to apply the RL-scheme for recovering $\phi(R,Z)$ from an observed $\tilde{\psi}(x,y)$ by

---

[1] $\delta(f(x)) = \sum_s \frac{\delta(x - x_s)}{|f'(x_s)|}$, where $f' = \frac{df(x)}{dx}$. The function $f(x)$ is a real, differentiable function with $f(x_s) = 0$. The roots x_s are real and simple.

means of Eqs. (5.6 a,b). However, since the probability kernel $P(x, y|R, Z)$ contains a δ-function, which is notoriously difficult to deal with in the context of discretized grid-data, I have chosen to reformulate the main Eqs. (5.6 a,b) for this special axisymetric case.

For the first integral of the iterative scheme the formulation (5.15) of the integral using the probability kernel $P(x, y|R, Z)$ is not necessary. Instead, this integral can be evaluated as a simple integral along the line–of–sight,

$$\psi_n(x, y) = \int_{-\infty}^{\infty} dz \, \phi_n(R(x, y, z), Z(x, y, z)), \tag{5.18}$$

where the coordinate transformation (5.8c) is used to evaluate the integration along z. The integral (5.18) is analytically equivalent to the integral (5.15), but not numerically. The approach using the probability kernel $P(x, y|R, Z)$ as in the integral (5.6a), which was used in Binney et al. (1990), involves the δ–functions. For my purpose, I found the approach given in (5.15) to be numerically extremely unstable, and therefore I employed the direct approach stated in (5.18).

The second step (5.6b) in the iterative RL-scheme reads in our case

$$\frac{\phi_{n+1}(R, Z)}{\phi_n(R, Z)} = \int dx \int dy \, \frac{\tilde{\psi}(x, y)}{\psi_n(x, y)} \, P(x, y|R, Z). \tag{5.19}$$

For this second integral the evaluation of the probability kernel $P(x, y|R, Z)$ cannot be avoided. However, it is possible to eliminate the δ-function by again applying the same rule as before to $\delta(f(y))$ and by subsequently integrating over y.

For the quadratic equation

$$f(y) = \left(\frac{y}{\cos i} - Z \tan i\right)^2 - (R^2 - x^2) \stackrel{!}{=} 0 \tag{5.20}$$

we find the two roots

$$y_\pm = \cos i \left(Z \tan i \pm (R^2 - x^2)^{1/2}\right), \tag{5.21}$$

which for a fixed pair of cluster coordinates (R, Z) describe an ellipse in the observer's sky. Since

$$f'(y) = \frac{2}{\cos i}\left(\frac{y}{\cos i} - Z \tan i\right) \quad \Rightarrow \quad f'(y_\pm) = \pm\frac{2}{\cos i}\sqrt{R^2 - x^2}, \tag{5.22}$$

we obtain after integration over y

$$\frac{\phi_{n+1}(R, Z)}{\phi_n(R, Z)} = \frac{1}{2\pi} \int dx \, \frac{1}{\sqrt{R^2 - x^2}} \left[\frac{\tilde{\psi}(x, y_+)}{\psi_n(x, y_+)} + \frac{\tilde{\psi}(x, y_-)}{\psi_n(x, y_-)}\right]. \tag{5.23}$$

The integration over x in Eq. (5.23) is limited to a finite range $[-R, +R]$. Thus it is convenient to introduce a new variable t via

$$x(t) = R \cos t \quad \Rightarrow \quad x(-\pi) = -R \quad \text{and} \quad x(0) = R. \tag{5.24}$$

Utilizing that

$$\sqrt{R^2 - x^2} = R\sqrt{1 - \cos^2 t} = -R \sin t \qquad \text{for} \quad t \in [-\pi, 0], \tag{5.25}$$

we arrive at our final result, namely a formulation for the second integral in the RL-scheme without δ-functions:

$$\frac{\phi_{n+1}(R,Z)}{\phi_n(R,Z)} = \frac{1}{2\pi} \sum_{\pm} \int_{-\pi}^{0} \mathrm{d}t \, \frac{\tilde{\psi}(R\cos t, Z\sin i \pm R\sin t\cos i)}{\psi(R\cos t, Z\sin i \pm R\sin t\cos i)}. \tag{5.26}$$

Please note that the integral (5.26) describes a full ellipse on the sky. The difference between Eq. (5.26) and the Appendix of Binney et al. (1990) is, that Eq. (5.26) uses the full information of the ellipse, while Binney et al. (1990), use only half of it. This formulation has the advantage that no assumptions on the data are made, i.e. the latter are not assumed to be symmetrical along any of the projection axes.

5.3 Observables Sensitive to the LOS–Structure

As already stated in the introduction of this chapter, the overall goal is to combine multiple data sets within the Richardson–Lucy algorithm to deproject cluster images. For observed distributions — denoted as ψ in the previous sections — I now discuss concrete observables, namely the weak lensing potential, the X–ray surface brightness and the SZ temperature decrement. For employing the Richardson–Lucy algorithm, it is important to connect the observables to the theoretical distribution ϕ. I choose the gravitational potential φ as theoretical distribution. In principle it is possible to choose the density ρ as theoretical distribution, but the gravitational potential φ possesses better symmetry properties than the density ρ. Therefore substructure has less impact on the potential than on the density better fullfilling the symmetry assumptions made deriving the kernel $P(x, y | R, Z)$.

5.3.1 Lensing Potential

Hereafter, ψ shall exclusively denote the lensing potential introduced in Chapter 2. In Eq. (2.13) the dependence of the lensing potential on the 3-dim. gravitiational potential φ is given as the los integral

$$\psi(x, y) \propto \int_{-\infty}^{\infty} \varphi(R, Z) \mathrm{d}z. \tag{5.27}$$

With current observational techniques the lensing potential of clusters can be determined up to a radius of ≤ 1.5 Mpc from the center.

5.3.2 X–ray Emissivity

When X–ray telescopes became available after 1966, it was discovered that clusters of galaxies are powerful X–ray emitters. They have luminosities in the range of $(10^{43} - 10^{45})$ erg s^{-1}, making them the most luminous X–ray emitters in the sky. The source of X-ray emission in clusters is extended rather than point–like, and the X–ray spectra are best explained by thermal *bremsstrahlung* (free–free radiation) from the hot, dilute plasma with temperatures in the range $(10^7 - 10^8)$ K and densities of $\approx 10^{-4} - 10^{-3}$ particles per cm^3.

For the present purpose it is sufficient to include continuum emission only. Semi-classical derivations of free–free emission can be found in standard textbooks, e.g. in Rybicki & Lightman (1979) and in Shu (1991). The emissivity at a frequency ν associated with electrons accelerated by ions of charge Z in a plasma with temperature T is given by

$$j_\nu = \frac{2^5 \pi e^6}{3 m_e c^3} \left(\frac{2\pi}{3 m_e k} \right)^{1/2} Z^2 n_e n_i g_{ff}(Z, T, \nu) T^{-1/2} \exp\left(-\frac{h\nu}{kT} \right), \qquad (5.28)$$

where n_i and n_e are the number densities of ions and electrons, respectively. The Gaunt factor $g_{ff}(Z, T, \nu)$ corrects for quantum–mechanical effects and for the effect of distant collisions. It is a slowly varying function of frequency and temperature, and can be set to unity for nearly all frequencies and temperatures of practical interest. For a completely ionized gas mixture with a mass ratio of 75% hydrogen and 25% helium, i.e. a gas with a mean mass per particle $\bar{m} = 10^{-24}$ g, the thermal bremsstrahlung at position $\boldsymbol{x}$ in the energy range $E_{\mathrm{a}} \leq E \leq E_{\mathrm{b}}$ is

$$j_{\mathrm{x}}(\boldsymbol{x}; E_{\mathrm{a}}, E_{\mathrm{b}}) = 5.53 \times 10^{-24} \ \mathrm{erg} \ \mathrm{cm}^{-3} \mathrm{s}^{-1}$$

$$\times \left(\frac{kT}{\mathrm{keV}} \right)^{1/2} \left(\frac{n_e}{\mathrm{cm}^{-3}} \right)^2 \left[\exp\left(-\frac{E_{\mathrm{a}}}{kT} \right) - \exp\left(-\frac{E_{\mathrm{b}}}{kT} \right) \right], \quad (5.29)$$

where the electron density n_e in this case is given by

$$n_e = 0.52 \frac{\rho_{\mathrm{gas}}}{\bar{m}}. \qquad (5.30)$$

The observable X–ray surface brightness received at the 2–dim. position $\boldsymbol{\zeta}$ is the line–of–sight integral of the X–ray emissivity j_{x},

$$S_{\mathrm{x}}(\boldsymbol{\zeta}; E_{\mathrm{a}}, E_{\mathrm{b}}) = \frac{1}{4\pi(1 + z)^3} \int j_{\mathrm{x}}(R, Z; E_{\mathrm{a}}, E_{\mathrm{b}}) \mathrm{d}l, \qquad (5.31)$$

where the factor $(1 + z)^3$ accounts for the redshifting of the photons and the ratio between luminosity distance and angular diameter distance.

Assuming a hydrostatic gas distribution, it is possible to relate the observed X–ray surface brightness S_{x} to the 3-dim. gravitational potential φ by the Euler equation

$$\nabla P = -\rho_{\mathrm{gas}} \nabla \varphi(R, Z), \qquad (5.32)$$

where the gas pressure P obeys the ideal equation of state

$$P = P_0 \frac{\rho_{\mathrm{gas}}}{\rho_{\mathrm{gas},0}}. \qquad (5.33)$$

Therefore we obtain for a hydrostatic gas distribution a dependence on the potential φ of the form

$$\rho_{\mathrm{gas}}(R, Z) = \rho_{\mathrm{gas},0} \times \exp\left[-\frac{\bar{m}}{kT} \Big(\varphi(R, Z) - \varphi_0 \Big) \right]. \qquad (5.34)$$

Since the hydrostatic gas distribution is isothermal, the temperature T is independent of position, and we thus arrive at the following relationship between observed X–ray surface brightness $S_{\mathrm{x}}(x, y)$ and the 3-dim. gravitational potential $\varphi(R, Z)$:

$$S_{\mathrm{x}}(x, y) \propto \int_{-\infty}^{\infty} \exp\left[-2\varphi'(R, Z) \right] \mathrm{d}z. \qquad (5.35)$$

5.3.3 Sunyaev–Zel'dovich Effect

The Compton scattering of thermal electrons in clusters of galaxies on the cosmic microwave background (CMB) radiation field, which is called the Sunyaev–Zel'dovich effect (Sunyaev & Zel'dovich, 1970, 1972, 1980), is one of the most important astrophysical processes in a low–energy environment, where only small energy transfers occur, with observable consequences. In essence, the Sunyaev–Zel'dovich (hereafter SZ) effect is a perturbation of the spectrum of the cosmic microwave background as its photons pass through the hot gas of clusters of galaxies. One important property of the SZ-effect is, that for a given line integral of pressure through the cluster gas, the effect does *not* depend on redshift. The SZ–effect is a very important cosmological probe, which can be used to study the evolution and structure of the Universe. In conjunction with information at other wavelengths, particularly X–rays, it can be used to estimate the Hubble constant (Silk & White 1978; Birkinshaw 1999).

The frequency shift leads to an apparent deficit in intensity at low frequencies of the CMB spectrum, and an increase at higher frequencies. Here I assume that the temperature decrement $\Delta T_{\rm SZ}$ at certain frequencies can be measured. The temperature decrement as a function of redshift, expressed in terms of the R–J brightness temperature ($T_{\rm RJ} = (\lambda^2/2k)I$), is given as

$$\Delta T_{\rm SZ} = \left(\frac{x}{e^x - 1} \right)^2 e^x \left(x \coth \frac{x}{2} - 4 \right) y \, T_{\rm CMB}, \tag{5.36}$$

where y is the Comptonization parameter

$$y = \int \left(\frac{kT_e}{m_e c^2} \right) (\sigma_T n_e) \, {\rm d}l, \tag{5.37}$$

and $x = h\nu/kT_{\rm CMB}$. The first term of the componization parameter y describes the effect on a single electron, while the second term gives the probability. As opposed to the X–ray case which depends on $n_e^2 T^{1/2}$, the SZ–effect depends on $n_e T \propto \rho T$, which is proportional to the pressure P, thus providing additional constraints on the cluster. As in the X–ray case a hydrostatic gas distribution is assumed. The mixture of of 75% hydrogen and 25% again leads to an electron density $n_e = 0.52 \, \rho_{\rm gas}/\bar{m}$. Therefore the temperature decrement depends on the gravitational potential in the following way

$$\Delta T_{\rm SZ}(x, y) \propto \int_{-\infty}^{\infty} \exp\left(-\varphi'(R, Z) \right) {\rm d}z. \tag{5.38}$$

It is worthwile noting that for both, the X-ray surface brightness $S_{\rm x}$ and the SZ–temperature decrement $\Delta T_{\rm SZ}$, the dependance on the quantity of interest, the 3-dim. gravitational potential φ, is exponential, requiring great care in the numerical implementation.

5.3.4 Multiple–Data Richardson–Lucy Deprojection

Now I propose the multiple data Richardson–Lucy deprojection (MDRL) algorithm. The combination of different data sets can be achieved in three separate steps. The first

step of the algorithm is to compute the three los integrals

$$\psi_n(x, y) \;\; = \;\; \int_{-\infty}^{\infty} \mathrm{d}z \varphi_n(R(x, y, z), Z(x, y, z)), \tag{5.39a}$$

$$S_{\mathrm{x},n}(x, y) \;\; = \;\; \int_{-\infty}^{\infty} \mathrm{d}z \exp\left(-2\varphi_n(R(x, y, z), Z(x, y, z))\right), \tag{5.39b}$$

$$\Delta T_{\mathrm{SZ},n}(x, y) \;\; = \;\; \int_{-\infty}^{\infty} \mathrm{d}z \exp\left(-\varphi_n(R(x, y, z), Z(x, y, z))\right), \tag{5.39c}$$

resulting in iterated input data ψ_n, S_{x}, and ΔT_{SZ}. In the second step we define the three integrals

$$F_n(R, Z) \;\; = \;\; \int f_n(x, y) P(x, y | R, Z) \, \mathrm{d}x \mathrm{d}y, \tag{5.40a}$$

$$G_n(R, Z) \;\; = \;\; \int g_n(x, y) P(x, y | R, Z) \, \mathrm{d}x \mathrm{d}y, \tag{5.40b}$$

$$H_n(R, Z) \;\; = \;\; \int h_n(x, y) P(x, y | R, Z) \, \mathrm{d}x \mathrm{d}y, \tag{5.40c}$$

where f_n, g_n and h_n are defined as

$$f_n = \frac{\tilde{\psi}}{\psi_n} \quad ; \quad g_n = \frac{\tilde{S}_x}{S_{x,n}} \quad ; \quad h_n = \frac{\tilde{\Delta T}_{\mathrm{SZ}}}{\Delta T_{\mathrm{SZ},n}}. \tag{5.41}$$

Here $\tilde{\psi}$, $\tilde{S}_{\mathrm{x}}$, and $\tilde{\Delta T}_{\mathrm{SZ}}$ denote the observed input distributions.

In the second step above the three integrals – for the lensing potential ψ the integral (5.40a), for the X–ray surface brightness S_{x} integral (5.40b), and for the SZ temperature decrement ΔT_{SZ} the integral (5.40c) – correspond to the integral (5.6b) of Richardson–Lucy's iterative inversion algorithm, while the equations (5.39a) – (5.39c) correspond to the integral (5.6a) of the inversion algorithm.

In order for the multiple–data Richardson–Lucy algorithm to work the results of the integrations have to be combined after every iteration step as

$$\varphi_{n+1} = \varphi_n \left[\alpha F_n + \beta \left(1 - \frac{1}{2} \ln G_n \right) + \gamma \left(1 - \ln H_n \right) \right]. \tag{5.42}$$

α, β, and γ are weighting factors with $\alpha + \beta + \gamma = 1$, which can be used to determine the relative weight put on the respective input data. By means of these weighting factors, it is possible to recover the case of having just one set of measured data.

As already mentioned in section (5.3) for the X–ray case and the SZ–case the dependence on the gravitational potential φ is exponential. Thus it is necessary for these two cases to take the logarithm of the integrals (5.40b) and (5.40c) in order to reconstruct the gravitational potential. The factor $1/2$ for the X–ray integral G_n arises due to the fact that the X–ray surface brightness depends on the electron density n_e^2 and not n_e.

Equation (5.42) completes the multiple–data Richardson–Lucy deprojection algorithm, thus constituting the third step. In terms of the iterative inversion algorithm of section 5.1 this step corresponds to the evaluation of the full equation (5.6b).

5.4 Implementation of the Multiple–Data Richardson–Lucy Deprojection

The implementation of the program follows closely the formulation of the MDRL–algorithm given in the last section. For the algorithm, two different types of integrals need to be evaluated. After reading the observed data and assuming an initial guess for the gravitational potential the iteration cycle is entered. The first step in the iteration cycle is to integrate the gravitational potential φ according to Eqs. (5.39a) – (5.39c), to obtain the iterated ψ_n, $S_{x,n}$, and $\Delta T_{SZ,n}$. Now the results of this *first* integration can be used to compute the integrals (5.40a) – (5.40c) that constitute the *second* integration. The last step of the iteration cycle is to compute the new estimate for the gravitational potential φ_{n+1} from Eq. (5.42). This new, improved estimate of the gravitational potential is used to reenter the iteration cycle. The whole algorithm is stopped after a few iterations; usually $7 - 8$ iterations lead to satisfactory results for the potential.

Both, the input data ψ_0, $S_{x,0}$, $\Delta T_{SZ,0}$ and the gravitational potential φ are represented as discretized data on a rectangular grid. The observed data is assumed to cover a *finite* data field $(x_{\min}, x_{\max}) \times (y_{\min}, y_{\max})$, and the reconstructed potential thus covers a corresponding field $(R_{\min}, R_{\max}) \times (Z_{\min}, Z_{\max})$, in cluster coordinates (R, Z). For the necessary first and second integrations I found the most stable solution to be a finite difference scheme (NAG, Mark 17, Routine D01GAF), where the integral between two successive points is calculated using a four–point formula centered on the interval concerned. For the first and the last intervals four–point forward and backward difference formulae, respectively, are used.

For the first integrals of the iterative scheme (5.39a) – (5.39c) I implemented a standard line–of–sight integral in the formulation of Eq. (5.18). As expected, preliminary tests with the equivalent formulation (5.15) via the integral kernel $P(x, y | R, Z)$ as in Binney et al. (1990), did not lead to satisfactory results.

I evaluate the line–of–sight integral by first rotating the discretized gravitational potential before computing an integral along the z–axis.

The second integration (5.40a) – (5.40c) yielding F_n, G_n, and H_n is performed on an ellipse as in formula (5.26), that was derived for the axisymetric case in Sec 5.2. For easier reference we explicitly write down the corresponding integral for the lensing case:

$$F_n(R, Z) = \frac{1}{2\pi} \sum_{\pm} \int_{-\pi}^{0} dt \, \frac{\tilde{\psi}(R \cos t, Z \sin i \pm R \sin t \cos i)}{\psi(R \cos t, Z \sin i \pm R \sin t \cos i)}. \qquad (5.43)$$

From Eq. (5.43) we see that for each given point in the (R, Z) plane the fraction $\tilde{\psi}/\psi$ has to be integrated over an ellipse

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} R \cos t \\ Z \sin i & \pm R \cos i \sin t \end{pmatrix} \qquad t \in [0, 2\pi] \qquad (5.44)$$

that is shifted along the y axis by $Z \sin i$ and contracted in the y-direction by a factor of $\cos i$. Due to the finite range of values for x and y we are faced with the problem that some parts of the ellipse, and thus of our integration path, may not be covered by the input data. This is most likely the case if *both* R and $|Z|$ take medium to large values.

Even worse, if *either* R or Z becomes very large, then the full ellipse will be outside the input range for x and y.

Therefore one either has to exclude all ellipses with such large (R, Z)–coordinates or one has to define appropriate "boundary" conditions, i.e. values for the ratio $\tilde{\psi}/\psi_n$, $\tilde{S}_{\mathrm{x}}/S_{\mathrm{x},n}$ and $\tilde{\Delta T}_{\mathrm{SZ}}/\Delta T_{\mathrm{SZ},n}$ have to be specified for points (x, y) outside of the data field $(x_{\min}, x_{\max}) \times (y_{\min}, y_{\max})$. I tested in some detail the following three possible choices:

1. Assume a smooth expansion of the values of the ratio *observed / iterated* data sets, e.g. for $x \geq x_{\max}$ and $y \geq y_{\max}$ $F_n(x, y) = F_n(x_{\max}, y_{\max})$, $G_n(x, y) = G_n(x_{\max}, y_{\max})$, and $H_n(x, y) = H_n(x_{\max}, y_{\max})$.

2. Assume a perfect reconstruction, i.e. the ratio *observed / iterated* data sets equals unity: $F_n(x, y) = G_n(x, y) = H_n(x, y) = 1$ for $x \geq x_{\max}$ or $x \leq x_{\min}$ and $y \geq y_{\max}$ or $y \leq y_{\min}$.

3. Assume rotational symmetry for the values of the ratios. This can be achieved by computing the distance to the center of the data field and averaging over all values for the corresponding circle which are inside the data field. However this method is computationally relatively expensive, and the results obtained are poor.

Experimenting with these three different boundary conditions and various sizes of the input field, I found that this boundary problem has no significant influence on the quality of the reconstruction achieved for the central part of the potential. A typical example of this problem is given in Fig. 5.2, where a lensing based reconstruction of the gravitational potential from a gas–dynamical simulation (see next section) is shown for the three different boundary conditions listed above. It is obvious from Fig. 5.2
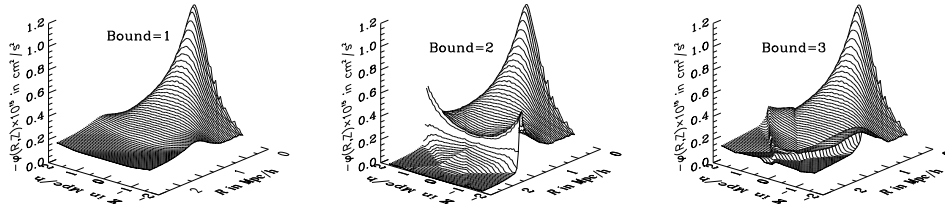

Figure 5.2: Gravitational potential of a cluster from a gas–dynamical simulation reconstructed from lensing data using different boundary conditions in the second integration of the axisymmetric RL algorithm. The enumeration of the boundary conditions is as in Sec. 5.4.

that the conditions No.2 and 3 introduce unphysical numerical artefacts for large values of R that have nothing to do with the true potential. Furthermore, condition No.3 is computationally relatively expensive. Thus I have decided to exclusively use boundary condition number No.1 in the remainder of this chapter.

Taking the logarithm for the X–ray and the SZ–case in Eq. (5.42), which is due to the exponential dependence in the two cases on the gravitational potential, is numerically a very unstable operation. Small deviations in computing the ratio $\tilde{S}_{\mathrm{x}}/S_{\mathrm{x},n}$ and $\tilde{\Delta T}_{\mathrm{SZ}}/\Delta T_{\mathrm{SZ},n}$ are magnified by taking the logarithm, thus preventing convergence of

the algorithm. In order to avoid this problem I employed a cut–off criterion for the ratio. This is done at the expense of a slower convergence, but as the algorithm converges very quickely and is computationally inexpensive this does not pose a serious problem.

5.5 Deprojection of Cluster Images from Gas–Dynamical Simulations

Having put the Multiple–Data Richardson-Lucy algorithm together, the next important step is to explore how this algorithm behaves, when it is applied to observed data. In particular, we have to assess the key property of the MDRL algorithm, namely the quality of the reconstructed gravitational potential achievable for a given set of input data. Therefore we have to apply the method first to input data for which we know the true gravitational potential. For this purpose we use clusters from gas–dynamical simulations kindly provided by Klaus Dolag, to construct observed images for the lensing potential ψ, the X-ray surface brightness S_x, and the SZ-temperature decrement ΔT_{SZ}, reconstruct the gravitational potential φ^{MDRL}, and compare it to the true grav. potential φ calculated directly from the simulation data. The gas–dynamical simulations include information on the dark–matter distribution, the gas distribution and the temperature of the cluster. They were created using a GRAPE-MSPH code that combines the gravitational interaction of the dark matter component with the hydrodynamics of a gaseous component. In addition, the code includes the magnetohydrodynamic equations following the evolution of the magnetic fields. A detailed description of the program and the simulations can be found in Dolag et al. (1999).

The cluster simulations were run using a COBE–normalized CDM power spectrum with a Hubble parameter $H_0 = 50$ km s^{-1} Mpc^{-1}, and $\Omega_0 = 1.0$, $\Omega_\Lambda = 0.0$. The virial mass of the cluster used is $M_{200} \approx 2.4 \times 10^{15} h^{-1}$ M$_\odot$, which resides in a volume of roughly $V \approx (5 h^{-1}$ Mpc$)^3$. The simulations contain approximately 5×10^4 dark–matter particles and also the same number of gas particles. The dark–matter particles have a mass of $m_{dm} \approx 1.6 \times 10^{11} h^{-1}$ M$_\odot = 20 \times m_{gas}$. The masses of the DM and gas particles provide an estimate for the resolution limit of the simulations. For the purpose of mimicking "observed data sets" within current observational limits, the above resolution is completely sufficient.

The gas and DM distributions of single clusters from the simulations are then used to compute the true gravitational potential of the cluster, from which then the observed lensing (Eq. (5.27)), X-ray (Eq. (5.35)), and SZ-data (Eq. (5.38)) are deduced, which in turn serve as input for the MDRL algorithm. Figure 5.3 shows three typical input sets created from the gas–dynamical cluster simulation of a single, very massive sample cluster. From the left to the right the lensing potential ψ, the X-ray surface brightness S_X in the energy band 2 keV to 12 keV, and the temperature decrement ΔT_{SZ} at an assumed frequency of 10 GHz are displayed.

Before turning to the *multiple* data RL-reconstruction of the potential φ, we have to investigate how well the algorithm works for each of the three different types of input data separately. This means we first consider only the case where φ is deprojected from either lensing, X-ray, or SZ-data alone, and then compare the findings.

We start by looking at different initial guesses for the gravitational potential φ_0, which are used in the first iteration cycle. Ideally the algorithm should not depend on
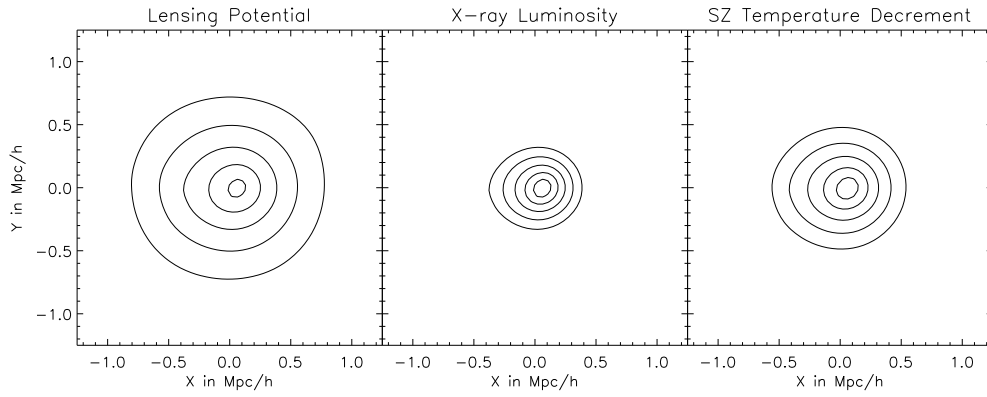
Figure 5.3: Input data sets created from the cluster simulation data. Left panel: lensing potential ψ, contours at $(-4, -3.5, -3, -2.5, -2) \times 10^{-8}$. Middle panel: X-ray surface brightness S_{x}, contours at $(5.75, 5.8, 5.85, 5.9, 5.95) \times 10^{-5}$. Right panel: SZ-temperature decrement ΔT_{SZ}, contours at $(-1.78, -1.77, -1.76, -1.75, -1.74) \times 10^{-2}$.

the choice of the initial guess, therefore two extreme cases for the initial guess are tested and the results are shown in Fig. 5.4. On the left of Fig. 5.4 I use a NFW potential as initial guess, which resembles the original profile rather closely, especially concerning the curvature of the potential. On the right of Fig. 5.4 I use a plane as initial guess, which makes only minimal assumptions about the potential. Fig. 5.4 shows the reconstructed gravitational potential φ after 8 iterations for lensing data only, but the results for X–ray and Sunyaev–Zel'dovich data are very similar. Comparing the reconstructed potentials in the lower panels obtained from these completely different initial guesses we clearly see that both initial guesses lead to qualitatively very similar results. The main difference is that the potential reconstructed with the NFW profile as initial guess is steeper in the inner part as opposed to the potential obtained from the plane as initial guess. This can be attributed to the fact that the smallest ellipses with (R, Z) coordinates close to zero are not taken into account due to numerical reasons. This is due to the fact that the finite difference formula requireses at least four points on the ellipse. This explains the differences in the inner part of the reconstructed potentials. In addition, the potential reconstructed from the NFW profile shows less "artefacts" for large (R, Z)–coordinates. Even though the "bump" for large Z–coordinates and R– coordinates is less pronounced for the potential reconstructed from the NFW profile than for the one reconstructed from the plane, the behaviour for these large values of (R, Z) is introduced by the fact that the data field used for the integration is finite. As large (R, Z)–coordinates constitute the boundaries these differences are not relevant for assessing the quality of the reconstruction. For the reconstruction the behaviour in the central $1.5h^{-1}$ Mpc of the cluster is much more important. In this sense the differences found for the two choices of initial guesses are negligible.

The evolution of the reconstructed gravitational potential φ_n with the number of iteration steps n is exemplified in Fig. 5.5. Again, the discussion is confined to the lensing case with an inclination angle of $i = 30°$, and 8 iteration steps. The potential obtained from the X–ray and SZ case evolves in a qualitatively similar way. The initial guess in this case is a plane shown in the upper left panel, while the lower right panel shows the original potential obtained directly from the simulated cluster. Fig.
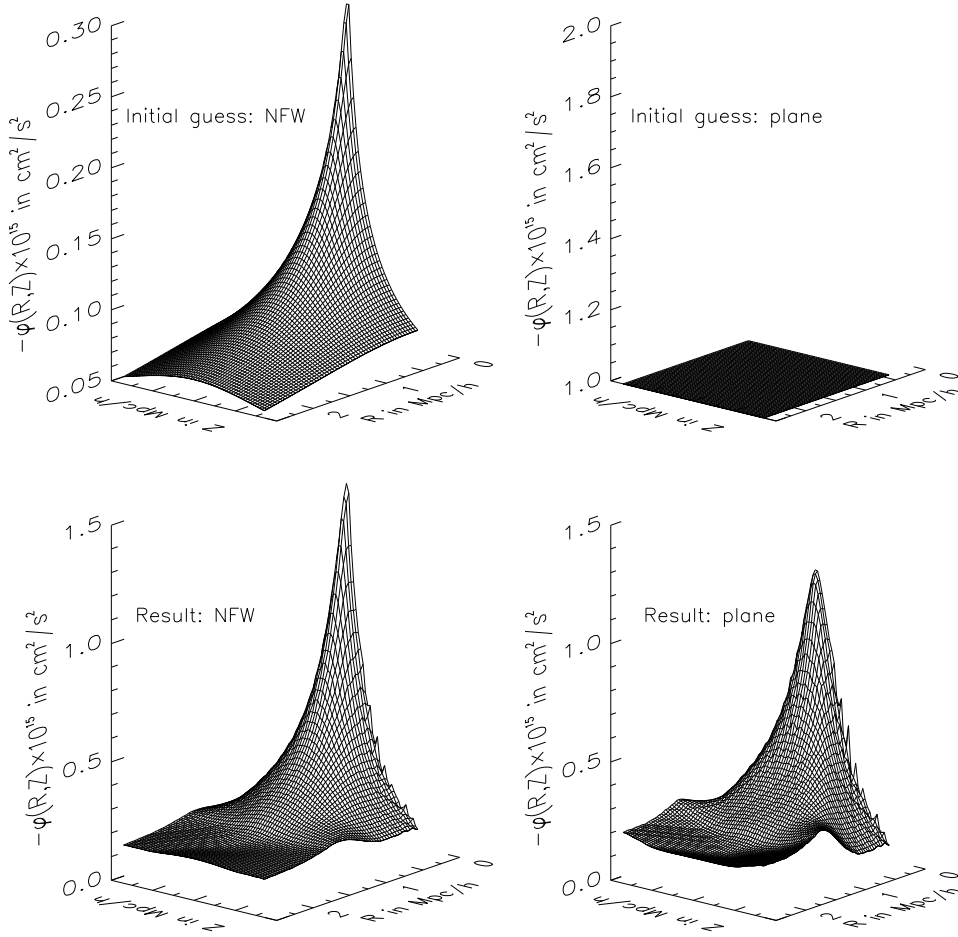
Figure 5.4: Comparison of two different initial guesses φ_0 (upper panels) for the recon-
struction of the gravitational potential φ. The lower two panels show the reconstructed
potential φ^{MDRL} from lensing data after 8 iterations for an inclination angle of $i = 30°$.

5.5 demonstrates that the algorithm converges extremely fast, even for an initial guess
making only minimal a priori assumptions about the cluster potential. Already after
the second iteration the potential is in the correct order of magnitude and has acquired
the characteristic features of the true cluster potential. In addition, we notice that φ_n is
hardly altered in the last two steps, indicating that the algorithm has converged in the
sense that most of the large–scale information is recovered. The two main differences
between the reconstructed potential from the last step, φ_8, and the true cluster potential
φ is the presense of two "dents" in φ_8 at $(R = 1h^{-1}$ Mpc , $Z = \pm 2h^{-1}$ Mpc), and
several small "wiggles" at the flanks of the potential for $R = 0$. I found that the size of
the "dents" can be correlated with the finite range of the "observed data", hinting again
at the inherent problems with the finiteness of the boundaries as discussed in Sec. 5.4.
The "wiggles" reflecting the property of the algorithm to fit small scale fluctuations last,
are in this case probably caused by numerical discretization effects, and thus reflect an
unwanted property of the algorithm. This numerical noise can be suppressed by using
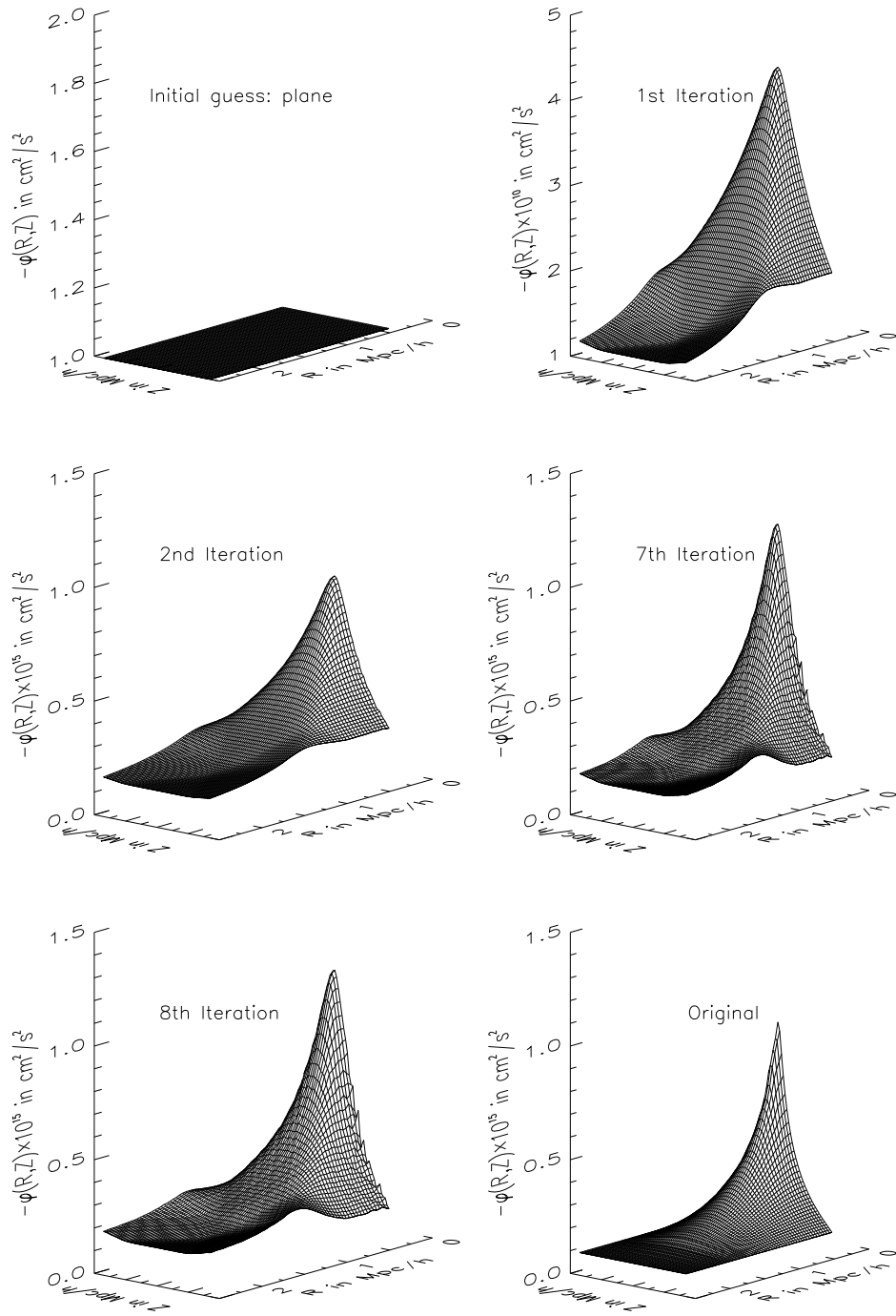a smoothing procedure after every iteration step.

Figure 5.5: Display of the iterated gravitational potential φ_n after different iteration steps n. The reconstruction is performed for the lensing potential ψ. The upper left panel shows the initial guess φ_0, while the lower right panel displays the original cluster potential φ.

In order to gain a better understanding on how the algorithm converges for the three different types of input data, it is instructive to look at the integrals F_n (5.40a) to H_n (5.40c). For fixed (R, Z) these integrals determine the multiplicative factors that advance φ_n to φ_{n+1} via Eq. (5.42), and, as already mentioned, good convergence requires that these integrals approach unity. The values of F_n, G_n, and H_n after the last step of Fig. 5.5 are plotted in Fig. 5.6. We see that the convergence after 8 iterations is al-
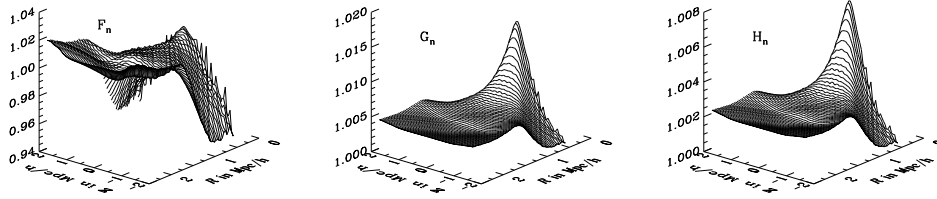

Figure 5.6: The integrals F_n (lensing; 5.40a), G_n (X-ray; 5.40b), and H_n (SZ; 5.40c) from the second step of the reconstruction algorithm. The same cluster data as in Fig. 5.5 was used ($i = 30°$; 8 iterations).

ready excellent over the full range of (R, Z) values for all integrals. The X-ray and SZ integrals G_n and H_n both overestimate unity by the same amount and feature a similar shape, reflecting the fact that they have a similar dependence on the potential φ, whereas the lensing integral F_n deviates more strongly from unity to both larger and smaller values, indicating a different convergence behaviour. The main difference is the fact that the integral F_n for small values of R and large values of Z is below unity thus, lowering the potential in this range.

We finish the discussion of the reconstruction from single data sources by comparing in Fig. 5.7 the true, original gravitational potential φ to the results of the reconstructions using the lensing potential, the X–ray surface brightness and the SZ temperature decrement alone as input data for the MDRL algorithm. By this the amount of information on the 3-dim. structure can be determined that is already present in each of the single data sets. Looking at the surface plots of Fig. 5.7 one sees that all three types of input data give qualitatively very similar results. The lensing reconstruction is superimposed by numerical noise, which is also present in the X–ray and SZ case, albeit much less pronounced.

A more detailed comparison of the three reconstructed and the reference potential is possible if surface cuts such as in Fig. 5.8 are studied. For the cuts through the central part of the cluster we generally see a good agreement of the three reconstructions with the original potential. The agreement becomes worse for larger radial coordinates as displayed in the right panel of Fig. 5.8 for the cut through $\varphi(R = 1/2R_{\max}, Z)$. We also notice that the difference between the X–ray and the SZ case is negligible, reflecting their very similar dependence on the gravitational potential.

Compared with the lensing potential ψ both, the X–ray and the SZ data give a very good reconstruction of the inner parts of the potential $(R, |Z| \leq 0.5h^{-1}$ Mpc) which is especially true for the cut along the Z–axis, where the match is nearly perfect. Apart from the fact that the numerical noise in the lensing potential is more pronouced, we also see that the potential is shifted slightly with respect to the original potential and the
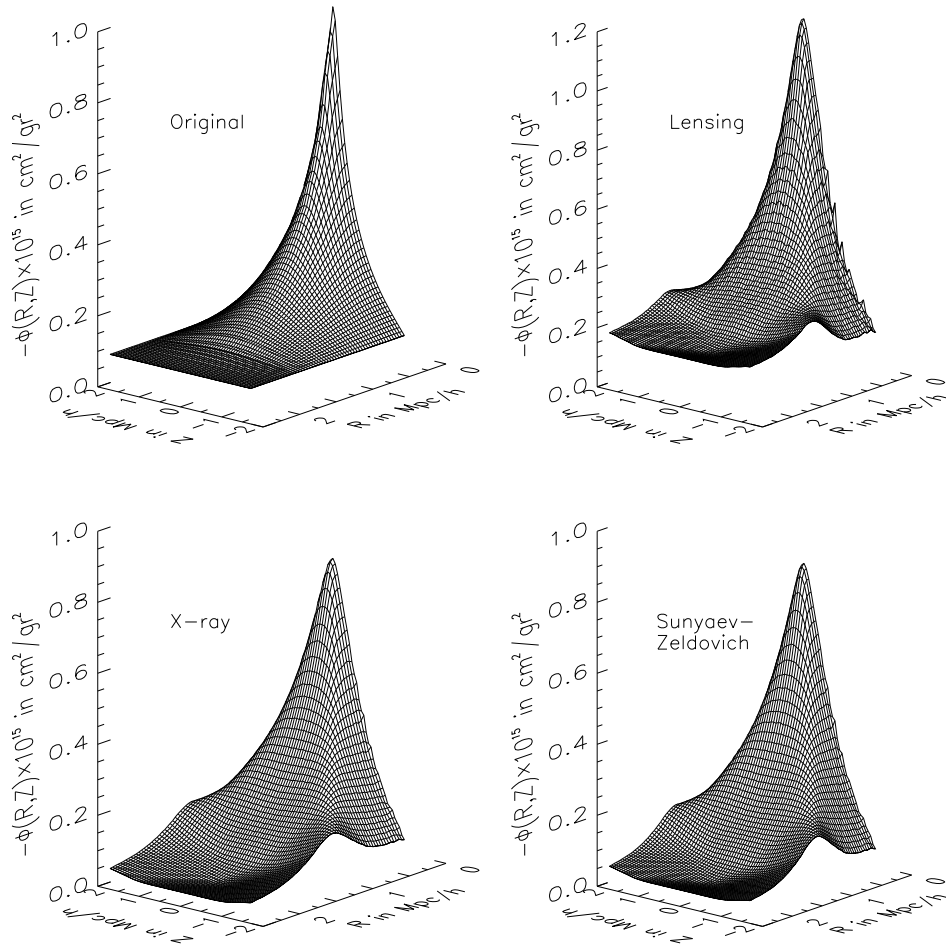
Figure 5.7: Comparison of the true, original gravitational potential φ (upper left panel) to reconstructions obtained by the lensing potential ψ alone (upper right panel), by the X-ray surface brightness S_x alone (lower left panel), and by the SZ temperature decrement ΔT_{SZ} (lower left panel) alone. The potential is shown in cluster coordinates (R, Z).

reconstructions obtained from X-ray and SZ data. This shift does not pose a problem as it is possible to align the potential with the other two reconstructions by adding a constant value. More important, in contrast to the X-ray and SZ-case the curvature and the overall shape of the lensing reconstruction is closer to the true potential even for larger radial coordinates $R \approx 1h^{-1}$ Mpc.

Finally, we are in a position to combine all input data sets for a true *multiple–data reconstruction*, thus allowing a better reconstruction of the gravitational potential φ. For the present example I chose to combine all three data sets with a weight factor of $1/3$. The results of the reconstruction are shown in Fig. 5.9. In the upper panel we compare two cuts of the result of the reconstruction obtained after 8 iterations with the original potential and the reconstructions computed for the single data sets, respectively. Especially for the cut $\varphi(R, Z = 0)$ we do see an improvement over the use of just
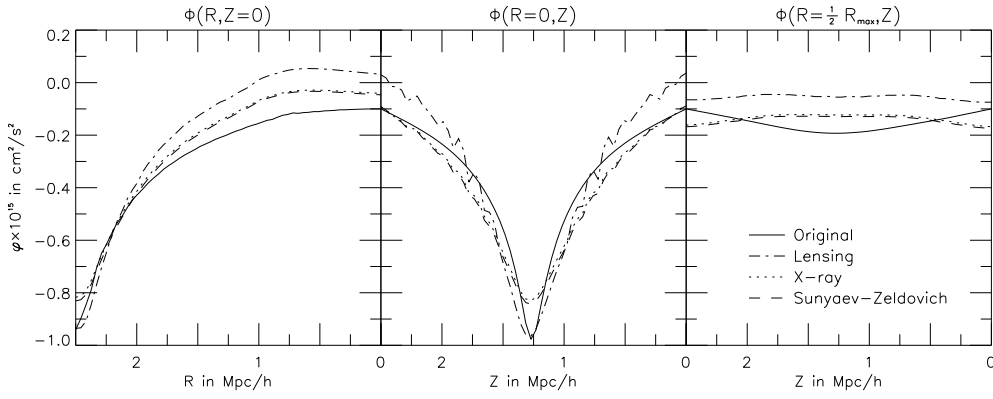
Figure 5.8: Comparison of three different cuts through the true, original gravitational potential φ and the three single-data reconstructions displayed in Fig. 5.7.

one single data set: The combined reconstruction is more reliable even for values of $R > 1h^{-1}$ Mpc. This is a clear improvement over the reconstructions obtained from X–ray and SZ data alone. The shift in the potential present for the reconstruction from lensing data is not present for the reconstruction from the combined data sets. The full surface plots of the original potential and the combined reconstruction demonstrates that the multiple data set reconstruction is able to recover all important features.

At this point it is worthwile to assess the quality of the reconstruction in a quantitative way. For this purpose it is instructive to look at the relative errors between the original gravitational potential φ_{orig} and the reconstructed one φ_{rec}, which is computed as $|\varphi_{\mathrm{orig}} - \varphi_{\mathrm{rec}}|/|\varphi_{\mathrm{orig}}|$. The result for the inner part of the potential, i.e. $R \in (0, 1.0)h^{-1}$ Mpc and $Z \in (-1.0, 1.0)h^{-1}$ Mpc, is displayed in Fig. 5.10. In the different panels of this figure the relative errors between the original and the reconstructed potential for lensing data (upper left panel), X–ray data (upper right panel), and Sunyaev–Zel'dovich data (lower left panel) are shown; in addition, the result for the combination of all three data types is given in the lower right panel.

For all four reconstructions we see that the deviation over large parts of the potential is less than 5%. When looking at the lensing reconstruction in more detail, we note that in this case the zone with an error margin of less than 5% is relatively wide, especially in the Z–direction. As already noted before, both, the X–ray and the Sunyaev–Zel'dovich reconstruction, show very similar features, which is also reflected in Fig. 5.10. Both cases give excellent reconstructions in the center with coordinate values (R, Z) of less than $0.5h^{-1}$ Mpc, but the quality of the reconstruction in the outer parts is not as good as in the lensing case. This confirms the theoretical expectation, that the data from X-ray and SZ measurements, which have their main contributions coming from the cluster core, are less affected by projection effects, nicely complementing the weak lensing data, which is only sensitive to the gravitating matter.

When looking at the results of the combined reconstruction an improvement over the single data reconstructions is obvious. Here the region with error margins of less than 5% is the largest.

Depending on the quality of the data at hand and a priori knowledge of the possibly different levels of noise present in the data, it is possible to adjust the weighting of the different data sets. This is also useful when one is interested in a limited region of the
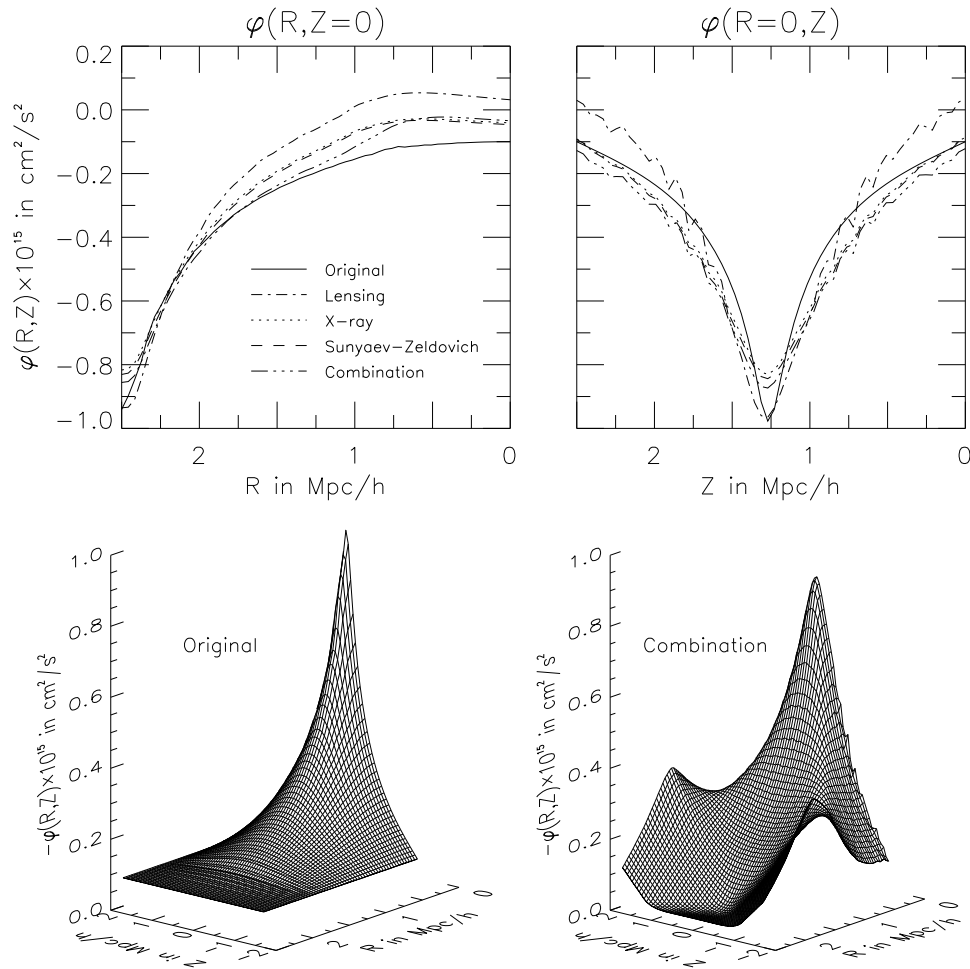
Figure 5.9: Result of the reconstruction obtained by combining all data sets shown in Fig. 5.7, each with a weighting of $1/3$. The upper panel shows two different cuts through the resulting potential comparing the original potential with the results obtained for the combined data set and the results for the single data sets from Fig 5.8. The lower panel compares the original potential to the potential reconstructed with the combined data sets as surface plot.

cluster which might be represented more accurately by a certain observable, like e.g. the cluster center which significantly contributes to the X–ray and SZ data. In summary, the combination of data sets can be expected to give improved results, with the astronomer being able to control the reconstruction process by means of the weight factors.

5.6 χ^2–based Determination of Inclination Angle i

Clearly in addition to a quantitative assessment of the quality of the reconstruction a systematic procedure for determining the input parameters for the MDRL algorithm is highly desirable.
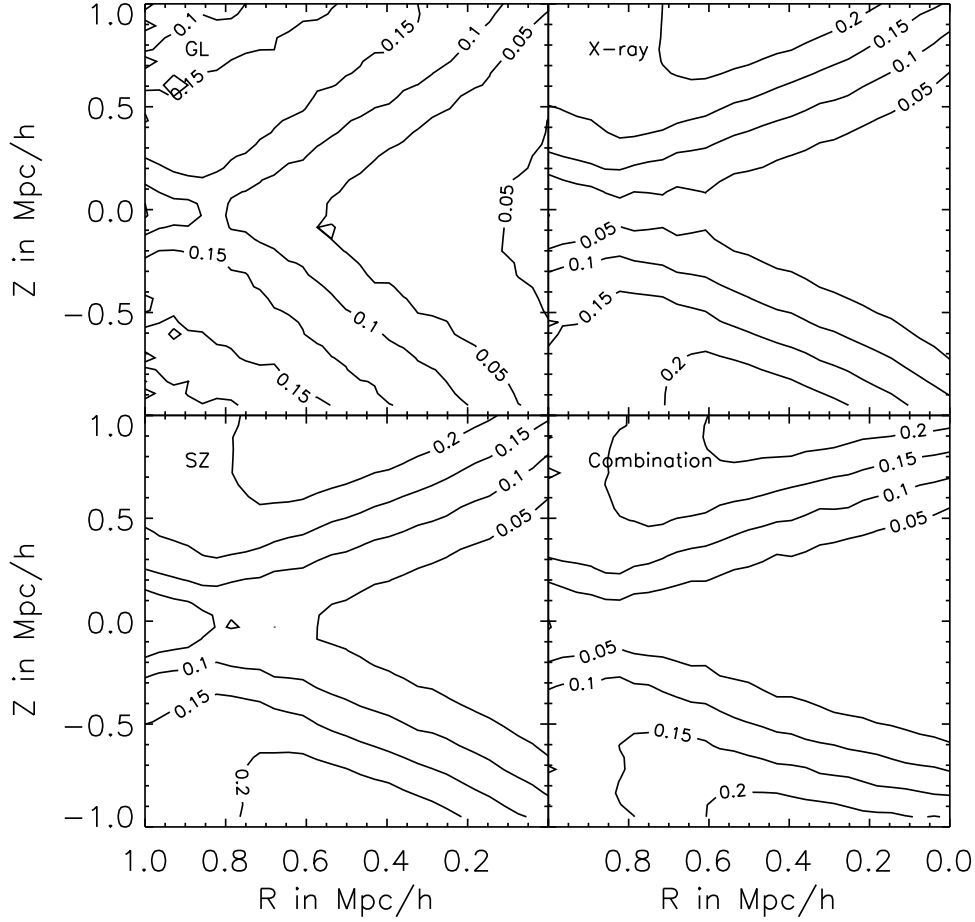
Figure 5.10: The relative error between the original gravitational potential $\varphi(R, Z)$ and the reconstructed potential computed as $|\varphi_{\mathrm{orig}} - \varphi_{\mathrm{rec}}|/|\varphi_{\mathrm{orig}}|$. The central part of the potential is displayed: $R \in (0, 1.0)h^{-1}$ Mpc and $Z \in (-1.0, 1.0)h^{-1}$ Mpc. The reconstructions shown are run with the same parameters as in Fig. 5.8. Upper right panel: lensing data only. Upper left panel: X–ray data only. Lower left panel: Sunyaev–Zel'dovich data only. Lower right panel: Combination of all three data types. Contours mark deviations of $(0.05, 0.1, 0.15, 0.2)$.

As explained in Sec. 5.2 the MDRL-algorithm in its current formulation needs to be provided with the inclination angle i as input parameter and a choice for the weight factors α, β and γ for the three different contributions has to be made. Once all these parameters have been specified, the MDRL algorithmus yields a reconstructed estimate φ_{rec} for the potential, which in turn determines the best estimates ψ, S_{x}, and ΔT_{SZ} for the observed input distributions $\tilde{\psi}$ (lensing potential), $\tilde{S}_{\mathrm{x}}$ (X-ray surface brightness), and $\Delta \tilde{T}_{\mathrm{SZ}}$ (Sunyaev-Zel'dovich temperature decrement).

The idea is now to minimize an appropriate χ^2 function, e.g.

$$\chi^2 = \sum_{i=1}^{N_g} \frac{(\tilde{\psi} - \psi(\vec{\xi_i}))^2}{\sigma_{\mathrm{GL}}^2} + \frac{(\tilde{S}_{\mathrm{X}} - S_{\mathrm{X}}(\vec{\xi_i}))^2}{\sigma_{\mathrm{X}}^2} + \frac{(\Delta\tilde{T}_{\mathrm{SZ}} - \Delta T_{\mathrm{SZ}}(\vec{\xi_i}))^2}{\sigma_{\mathrm{SZ}}^2}, \qquad (5.45)$$

to obtain the best parameters within the framework of the model assumptions underlying the presented MDRL algorithm. Here N_g would the number of grid points, with the i-th grid point given by a vector $\vec{\xi_i}$ in the xy-plane onto which the observables are projected. The desired quantitative assessment of the quality of the reconstruction could then be given in terms of a subsequent goodness-of-fit (GoF) evaluation, which would tell us how likely the reconstruction within this model is for the best set of parameters found before.

This idea is illustrated in the following using the cluster example from Sec. 5.5. We recall that the true 3-dimensional gravitational cluster potential $\varphi(R, Z)$ is observed under an angle of $i = 30°$, and assume that we do not know the proper angle i in advance. For setting up the χ^2 statistics we compute a series of single data and multiple data reconstructions with various values for the angles i and the weight factors. We then use Eq. (5.45) to evaluate the corresponding χ^2 function, and obtain a χ^2 value for every set of input parameters.

The minimum over all the χ^2 values then indicates the best choice for the inclination angle i and the weight factors. According to our qualitative analysis from Sec. 5.5 we would expect that the inclination angle i indeed should be very close to the true value.

Work on the quantitative estimation of the inclination angle and the weight factors α, β and γ using the approach sketched above is currently in progress.

Chapter 6

Conclusions and Outlook

The main objective of this thesis was the investigation of projection effects in clusters of galaxies, which influence the amount of cosmological information that can be extracted with weak-lensing methods. In this context I first studied how the statistics of cosmic shear measurements can be utilized to find dark matter haloes from their shear properties, before turning to the question of the significance of projection effects for cluster catalogues selected through the coherent image distortion patterns imposed on faint galaxies in their background. For both purposes the new weak-lensing-based estimator M_{ap}, recently introduced by Schneider (1996), was used to gain information about clusters of galaxies and the LSS into which clusters of galaxies are embedded. A question closely related to projection effects is the reconstruction of the true, 3-dimensional shape of individual clusters. This can be achieved by a deprojection of the observable, 2-dimensional cluster images through combining lensing-based data with additional information from X–rays and the Sunyaev–Zel'dovich effect. In the following I discuss the results obtained within the present thesis before providing an outlook for future theoretical as well as observational work.

6.1   Discussion of Results

With the goal of investigating the LSS and its evolution in mind, I used ray-tracing simulations through N-body-generated cosmic density distributions to study the statistical properties of the aperture mass M_{ap} as a statistics for cosmic shear measurements and for finding dark matter haloes from their shear properties. In particular, I compared results from these simulations with the available analytic results and found in most cases a very good agreement, except for the skewness which is the least accurate of these predictions. Whereas all other predictions tested here are based on manifestly non-linear results (like the Press-Schechter halo abundance and the Peacock & Dodds (1996) power spectrum), the skewness was estimated analytically by using second-order Eulerian perturbation theory which, on the scales considered, is not very accurate.

 Comparing the results from the ray tracing simulations with analytic studies, I obtain the following main results:

1. The *rms* of M_{ap} is accurately described by analytic results if the fully non-linear prescription of the power spectrum of density fluctuations is used.

2. The statistical error of this *rms* is dominated by cosmic variance, which in turn depends on the kurtosis of $M_{\rm ap}$. This kurtosis turns out to be unexpectedly large even on angular scales of $\sim 10'$, implying the need for many more measurements of $M_{\rm ap}$ than expected for a Gaussian field, for a given accuracy of the estimated projected power spectrum.

3. The skewness is only approximately described by analytic considerations based on second-order perturbation theory.

4. The predicted abundance of dark matter haloes detectable at given statistical significance is very well approximated by the semi-analytic theory which combines the Press-Schechter number density of haloes with the universal density profile of Navarro, Frenk & White.

5. Similarly, the functional form of the probability distribution of $M_{\rm ap}$ for values much higher than the *rms* (i.e., in the non-Gaussian tail) is found to closely follow an exponential form, of similar slope and amplitude as predicted by analytic theory which needs to assume that such high values originate due to collapsed haloes.

Thus, on the whole, I find that the analytical estimates for the statistical properties of $M_{\rm ap}$ are surprisingly accurate, but also that our simulations are not sufficiently large for an accurate estimate of the higher-order statistical measures, owing to the finite size of the simulation box in combination with the large effect of cosmic variance.

As discussed in SvWJK, KS1, KS2, van Waerbeke et al. (1999) and Bartelmann & Schneider (1999), the aperture mass is a useful cosmic shear measure which will eventually allow one to constrain cosmological parameters, completely independent of any assumption on the relation between mass and light. For this purpose, the predictions from cosmology must be known precisely, and my results here indicate that analytic estimates are relatively accurate. Unfortunately, I found a large cosmic variance; e.g., in the estimate of the variance of the *rms* value of $M_{\rm ap}$, the kurtosis enters and it decreases only rather slowly with increasing filter scale.

Having established that the second-order aperture-mass statistics $M_{\rm ap}$ is particularly suitable for investigating cosmic shear and that it is suitable for detecting dark matter haloes, which in this context are defined operationally as something visible as a suffieciently high peak in an S–map, it is also important to look at clusters of galaxies as actual 3–dimensional objects. In the next part of my thesis I therefore investigated, using simulated cluster data, for the first time whether mass-selected galaxy cluster samples constructed with the aperture mass measure $M_{\rm ap}$ are more reliable than samples constructed via Abell's criterion. As mentioned in the introduction, image distortions trace the gravitational tidal field of a lens rather than its mass, and it is in that sense that I speak of "mass-selected" cluster samples. I also compared the performance of cluster mass estimators based on cluster-galaxy kinematics and gravitational lensing. The results can be summarised as follows.

As already found in previous studies, Abell clusters are severely affected by projection effects. This not only concerns the selection of Abell clusters, but also mass estimates based on galaxy kinematics and the virial theorem, indicating that the velocity dispersion is also hampered by projection effects. A second reason for the failure is

the fact that the assumption of dynamical equilibrium is not justified in at least some of the clusters. The projection effects are worse for clusters and groups of lower richness class.

Clusters detected with a high significance S of M_{ap} are less affected by projection effects than typical Abell-selected clusters. Like Abell cluster samples, the mass-selected cluster samples are generally incomplete: Samples of clusters detected above a certain S threshold typically do not encompass all three-dimensional clusters present in the simulation; some clusters have lower S. However, the completeness of the samples can be increased by lowering the S threshold. I therefore investigated the effect of varying the S threshold on the samples. Completeness of $\approx 100\%$ can be achieved for massive three-dimensional cluster samples ($M \gtrsim 10^{14}\, h^{-1}\, M_{\odot}$) by varying $S \gtrsim 4$. Then, the samples also contain a substantial fraction of spurious detections, most of which correspond to real clusters with smaller masses. Generally, there is a trade-off between completeness and the contamination by spurious detections. More complete cluster samples are more heavily contaminated by spurious clusters, and the balance can be adapted choosing the S threshold. It should be noted that the exact thresholds on S depend somewhat on the choice of the weight function entering the definition of S (cf. the discussion in Chapter 3).

While qualitatively the same trend is also observed in Abell-selected cluster samples, the S-statistics generally performs significantly better than Abell's criterion: Higher completeness can typically be achieved with a lower fraction of spurious detections. For instance, cluster samples detected at $S \geq 4$ contain all of the most massive clusters in the simulation and 65% spurious detections, while Abell samples with richness $\mathcal{R} \geq 0$ encompass only about two-thirds of the most massive clusters and 82% spurious detections.

Lensing-based mass estimates are significantly more accurate than mass estimates based on cluster-galaxy kinematics and the virial theorem. Virial masses are typically biased high because line–of–sight velocity distributions are broadened by projection effects. Lensing also adds up mass in front of and behind the clusters, but the bias is less severe. The standard deviation from the true (three-dimensional) mass of the lensing mass estimate is smaller by a factor of three or more than that of the virial mass estimates. It should, however, be noticed that the accuracy of lensing-based mass estimates depends on the depth of the background-galaxy sample and other observational effects. While the mass estimates based on the ζ statistics are accurate to within $\approx 30\%$ in the simulations, they may well be less accurate under realistic observational conditions.

The study underestimates projection effects because of the limited size of the simulation volume. This affects both the optical and the lensing-based cluster selection. Yet it appears that selection of clusters by mass yields more reliable cluster samples than optical cluster selection, and, more importantly, the quality of the samples can be controlled by an objective criterion, namely the signal-to-noise threshold imposed. The study has shown that selection of clusters by means of gravitational lensing techniques can be adapted such that the resulting samples are superior to Abell-selected samples in terms of completeness, spurious detections, and the quality of mass estimates.

After studying the selection of clusters of galaxies and the related projection effects, it was an interesting question whether it is possible to obtain an improved reconstruction of the structure of individual clusters along the line–of–sight by combining observational data from different sources. In this thesis I was able to device an algorithm

based on the Richardson–Lucy deconvolution algorithm that uses lensing, X–ray and Sunyaev–Zel'dovich data to gain information about the 3-dim. structure of a cluster of galaxies. My first implementation of this *multiple–data Richardson–Lucy* MDRL algorithm was applied to synthetic clusters generated in gas–dynamical simulations, and I found it to work stably and efficiently:

Starting from earlier work of Binney et al. (1990) on deprojection of elliptical galaxies from photometric data, an integral kernel for an axisymmetric cluster model was derived and integral expressions for the n-th projected distribution $\psi_n(x, y)$ and the n-th iterative estimate $\phi_n(R, Z)$ to the true potential $\phi(R, Z)$ were obtained, that are suitable for numerical evaluation. In particular, I found that the simple line–of–sight integral for ψ_n is easier to evaluate and numerically more stable as the expression given in Binney et al. (1990), and thus is to be preferred.

It was shown that the three observables of interest, the lensing potential ψ, the X-ray luminosity S_x, and the Sunyeav–Zel'dovich temperature decrement ΔT_{SZ} can all be written as functionals of the gravitational potential $\phi(R, Z)$ assumnig hydrostatic equilibrium and an isothermal gas distribution. The lensing potential ψ by itself is given directly as a simple line–of–sight integral over $\phi(R, Z)$, whereas S_x and ΔT_{SZ} both have an exponential dependence on $\phi(R, Z)$ mediated through the electron density within an isothermal hydrostatic intracluster gas model. S_x is here assumed to result in continuum bremsstrahlung of the gas distribution, and ΔT_{SZ} is due to Compton scattering of CMB photons passing through the hot cluster gas.

In the practical implementation of the MDRL algorithm into a computer program a few important observations were made. The evaluation of the n-th iterative estimate $\phi_n(R, Z)$ requires integrations over ellipses on the projection plane, where the size and the relative location of the ellipse is determined by the current pair of (R, Z) values. In the X-ray and SZ case the logarithmic dependence on the projected quantities requires the use of cut-off criteria to minimize the magnification of small deviations between observed and reconstructed projected data.

A first application of the MDRL method to a cluster from gas–dynamical simulations showed that the approach already works well for single data reconstructions from ψ, S_x, or ΔT_{SZ}. The algorithm is very insensitive to the initial guess provided for ϕ; qualitatively no significant difference in the reconstructed potentials obtained from a very simple constant value function or the realistic NFW model potential chosen as initial guesses were found. As expected theoretically, the X-ray and SZ reconstructions, which have their main contributions coming from the cluster core and are less affected by projection effects, give a better description of the core region, whereas the lensing potential that is only sensitive to the dark matter distribution, but more prone to projection effects, better reproduces the overall shape of the potential. Finally, I found that a combination of all three data sets within a multiple data reconstruction improves upon the single data results: the inner region is described as well as in the X-ray and SZ case, the shift in the lensing reconstruction in this region is completely suppressed, and the description for larger (R, Z) radii lacking in the X-ray and SZ-case is compensated for by the lensing data.

I believe that this multiple data Richardson–Lucy reconstruction method will be a valuable and widely applicable tool.

6.2 Future Prospects

The prospects and challenges can be divided into further studies of constructing and defining dark matter halo samples on the one side, and additional work on the deprojection of individual clusters on the other side. While for the retrieval of the cluster mass function from lensing a lot of theoretical work is needed, the deprojection of cluster images is in a stage where it can, and should be applied to observational data.

My study of weak-lensing-based mass estimates and the semi-analytic approximations given by Kruse & Schneider (1999a,b) raise the question whether a cluster mass function can be obtained from weak lensing which is based on no selection criteria other than the lensing signal itself. Knowledge of the cluster mass function and its evolution provides a wealth of cosmological information. Especially, since lensing is most sensitive to clusters at moderately high redshift ($z \approx 0.4 - 0.6$), a lensing-based cluster mass function would nicely complement information on the mass function obtained from X-ray selected cluster samples. What is more, lensing-based mass estimates do not rely on any assumptions on the composition and physical state of the cluster matter, in contrast to X-ray mass estimates. In addition, lensing-based cluster detections only require sufficiently deep imaging of wide fields in optical or near-infrared wave bands, and detection algorithms can then be applied in a straightforward manner. It can therefore be expected that reliable, mass-selected cluster samples at moderate to high redshifts can be constructed in the near future from upcoming deep, wide-field surveys with a straightforward, well-controlled algorithm, and that the accuracy of cluster mass estimates will generally be substantially improved.

However, before weak-lensing-detected clusters can be used for this type of analysis, extensive numerical studies need to be performed because weak lensing provides information about the projected masses from a very broad redshift range, rendering the inference of the mass of an object a very delicate undertaking. It is therefore extremely important to investigate and model the properties of noise arising in weak lensing. First steps into this direction were made by Jain & van Waerbeke (1999), who included an analytic model for the noise from the intrinsic ellipticities of the background sources. A lot of work remains to be done here; e.g. the redshift distribution of the sources needs to be taken into account, and it will be of great interest to investigate the influence of the clustering of the background sources.

It can be expected that the first successful application of the aperture mass will be the definition of a sample of haloes defined in terms of their lensing properties only, with a first example given by Erben et al. (1999). The combination of cosmic shear information and CMB measurements can be extremely useful, as shown by Hu & Tegmark (1999), increasing the precision of the determination of cosmological parameters substantially over each of the two individual methods. Their study was based solely on the dispersion of cosmic shear, i.e., on second-order statistics. It is to be expected that a similar combination of CMB results with the PDF of $M_{\rm ap}$ will yield even more precise parameter estimates. A detailed study of this combination is expected to be very valuable, but requires a larger grid of cosmological N-body simulations.

A further interesting project is an extension of the investigation of cosmic shear to larger angular scales. Based on the finding that even for the small field sizes I investigated so far the tail of $M_{\rm ap}$ contains information on already collapsed haloes, it will be rewarding to apply maximum-likelihood methods to larger fields, which should allow

one to discriminate between different cosmological models.

Finally, a streamlined and extended implementation of the multiple data Richardson–Lucy cluster reconstruction algorithm is planned, in which the experiences gained in the present work will be incorporated, and which will allow the treatment of true observational data. This revised algorithm will utilise more realistic assumptions about the cluster itself, which becomes possible with the new generation of X-ray telescopes with their higher resolution and more accurate information on the temperature gradient in the intracluster plasma. Furthermore, I plan to add maximum-entropy regularisation to the algorithm, which has the advantage of providing a well-defined convergence criterion.

But first and foremost, the cluster deprojection algorithm shall be applied to observed data sets, thus helping to shed light on such important issues as determinations of cluster mass and baryon fractions, the structure of cluster-galaxy orbits, and the cosmological interpretation of galaxy clusters.

Relevant data will soon become available in substantial quantities. For example, a group around J.P. Kneib (OMP, F) plans to undertake a detailed optical survey of X-ray selected clusters in order to study their mass distribution. This group also proposed to complement the X-ray data with deep multi-colour imaging which allows for weak-lensing analyses of the clusters. The cluster deprojection algorithm will be very well suited for constraining the three-dimensional distribution of the clusters from combined lensing and X-ray data. In addition, the target runs from _Chandra_ will observe some clusters for which weak-lensing analyses already exist, which can then be used in conjunction with the _Chandra_ observations. Dedicated surveys will also produce ample Sunyaev-Zel'dovich data sets in the near future. It is now the time to finally combine all available sorts of cluster data to consistently reconstruct the three-dimensional gravitational potential underlying all of them.

In that respect, it is also of great interest to compare the cluster deprojection with that proposed by Zaroubi et al. (1998), which is based on the Fourier slice theorem, to assess the strengths and weaknesses of both algorithms and to optimise them.

Appendix A

Structure of Further Representative Clusters

Here I give a few more examples of line-of-sight structures of S-selected galaxy clusters here.

A.1 S-Statistics: $S \geq 5$

A second example for a cluster with high S is given in Fig. A.1. The cluster is detected at $S = 9.6$. The particle distribution in real space is broad and dominated by a massive 3-D cluster with a mass of $2.1 \times 10^{14} \, h^{-1} \, M_\odot$. This cluster is detected as an Abell cluster in projection, but the main 3-D cluster by itself already passes the luminosity threshold of a 3-D Abell cluster. In contrast to the first example, the velocity dispersion is hardly affected by projection. The 3-D cluster has a velocity dispersion of $\sigma_{3-\mathrm{D}} = 884 \, \mathrm{km\,s^{-1}}$, while the dispersion of the projected cluster is $\sigma_{2-\mathrm{D}} = 779 \, \mathrm{km\,s^{-1}}$. The higher-order moments indicate a velocity distribution close to Gaussian shape for both the 3-D cluster ($\mathcal{S}_{3-\mathrm{D}} = 0.01$, $\mathcal{K}_{3-\mathrm{D}} = -0.04$) and the projected cluster ($\mathcal{S}_{2-\mathrm{D}} = 0.09$, $\mathcal{K}_{2-\mathrm{D}} = -0.37$). All this reveals a fairly relaxed cluster with low contamination.

Almost all other clusters in this class show similar position and velocity histograms. The only exceptions are the 2-D clusters corresponding to less massive 3-D clusters. For one of these clusters with relatively high $S = 7.9$, the structure is given in Fig. A.2. Even though the position histogram is dominated by a 3-D cluster, the distribution for this cluster is broad, and there is a large amount of intervening matter with at least four smaller clumps with masses of order $M = 10^{13} \, h^{-1} \, M_\odot$. Qualitatively, the *los* velocity histogram looks artificially broadened by these clumps, and in fact the velocity dispersion ($\sigma_{3-\mathrm{D}} = 651 \, \mathrm{km\,s^{-1}}$) is significantly increased in projection ($\sigma_{2-\mathrm{D}} = 860 \, \mathrm{km\,s^{-1}}$). The higher-order moments are also strongly affected by this intervening matter ($\mathcal{S}_{3-\mathrm{D}} = -0.05$ and $\mathcal{K}_{3-\mathrm{D}} = -0.64$ compared to $\mathcal{S}_{2-\mathrm{D}} = 1.19$ and $\mathcal{K}_{2-\mathrm{D}} = 0.87$). This cluster is detected as $\mathcal{R} = 1$ Abell cluster although it corresponds only to a $\mathcal{R} = 0$ cluster in 3-D.
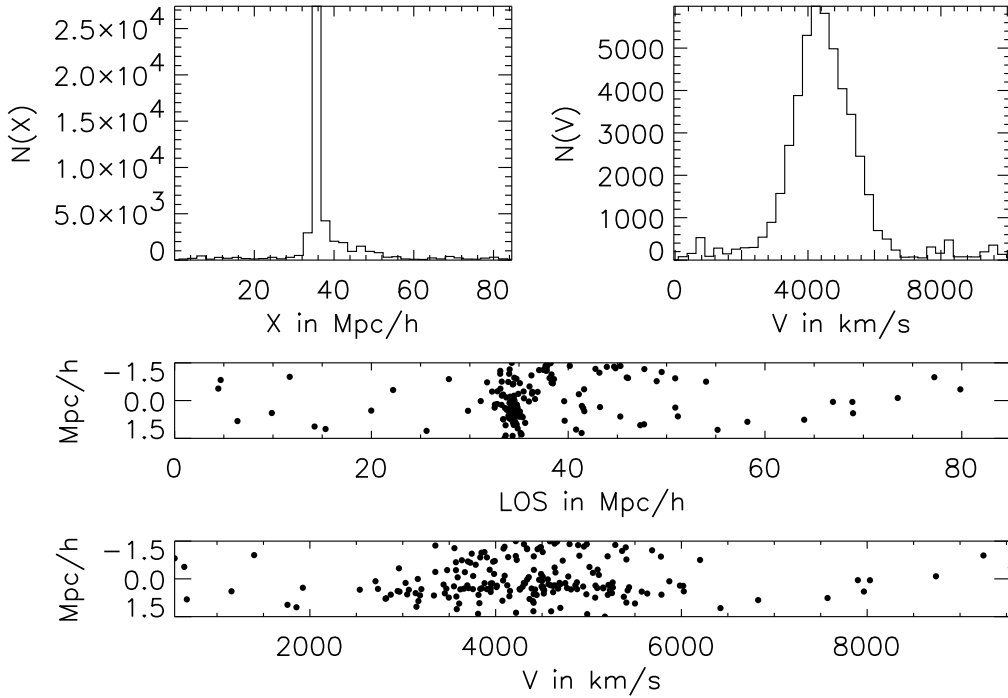
Figure A.1: Structure of a massive 3-D cluster detected with $S = 9.6$. See the caption of Fig. 4.6 for a description of the panels.

A.2 S-Statistics: $5 > S \geq 4$

Figure A.3 shows a cluster with $S = 4.7$. It is apparently only mildly contaminated by a clump $30\,h^{-1}$ Mpc from the main clump, which is a high-mass object with $M_{3-\mathrm{D}} = 2.3 \times 10^{14}\,h^{-1}\,M_\odot$. The projected velocity dispersion is almost unaffected ($\sigma_{2-\mathrm{D}} = 650\,\mathrm{km\,s^{-1}}$ compared to $\sigma_{3-\mathrm{D}} = 635\,\mathrm{km\,s^{-1}}$), and shows a bimodal feature, which is also reflected by the curtosis of the projected cluster, $\mathcal{K}_{2-\mathrm{D}} = 0.51$, while the velocity distribution of the 3-D cluster has a negative curtosis of $\mathcal{K}_{3-\mathrm{D}} = -0.22$. Similarly, the skewness changes from $\mathcal{S}_{2-\mathrm{D}} = -0.44$ to $\mathcal{S}_{3-\mathrm{D}} = -0.03$. The cluster is detected as a 2-D Abell cluster with richness class $\mathcal{R} = 1$, while the richness class of the 3-D cluster is $\mathcal{R} = 0$. Therefore, the richness class is inflated due to projection. Even though this cluster shows some projection effects, the corresponding 3-D cluster is massive and therefore clusters like that should be included in a mass-limited sample.

The last example for this class is shown in Fig. A.4. Here, the S-map has a peak with $S = 4.4$. The position histogram shows a very broad peak with a secondary maximum on top of the main peak. The corresponding 3-D cluster has a high mass, $M_{3-\mathrm{D}} = 3.4 \times 10^{14}\,h^{-1}\,M_\odot$. The projected velocity distribution is only moderately skewed with $\mathcal{S}_{2-\mathrm{D}} = 0.22$ compared to the skewness of the main cluster alone, $\mathcal{S}_{3-\mathrm{D}} = 0.17$. However, the curtosis of the projected peak, $\mathcal{K}_{2-\mathrm{D}} = 0.38$, even changes sign when compared to the 3-D cluster, $\mathcal{K}_{3-\mathrm{D}} = -0.27$. This cluster satisfies Abell's criterion in projection, but the main peak has a lower richness class, $\mathcal{R} = 0$.
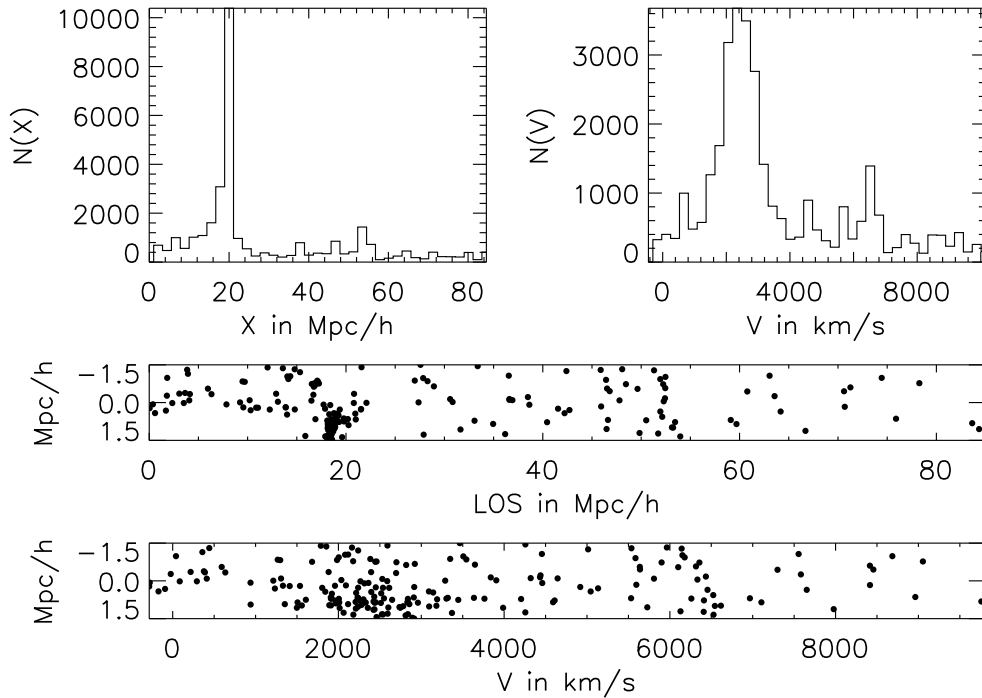
Figure A.2: Structure of a less massive 3-D cluster whose size is increased in projection due to matter concentrations along the line-of-sight. See the caption of Fig. 4.6 for a description of the panels.

A.3 S-Statistics: $4 > S \geq 3$

Another example for a low-S cluster detected at $S = 3.4$ is displayed in Fig. A.5. This detection also corresponds to a 3-D cluster with $M_{3-\mathrm{D}} = 6.6 \times 10^{13} \, h^{-1} \, M_{\odot}$. Again, the velocity distribution of this cluster is largely altered by the considerable amount of intervening matter. The velocity dispersion itself is inflated from $\sigma_{3-\mathrm{D}} = 504 \, \mathrm{km \, s^{-1}}$ to $\sigma_{2-\mathrm{D}} = 1134 \, \mathrm{km \, s^{-1}}$. This is reflected by the curtosis, which changes from $\mathcal{K}_{3-\mathrm{D}} = -0.02$ to $\mathcal{K}_{2-\mathrm{D}} = -0.5$, while the skewness changes from $\mathcal{S}_{3-\mathrm{D}} = 0.18$ to $\mathcal{S}_{2-\mathrm{D}} = -0.94$. Both low-$S$ examples are neither 2-D Abell clusters nor do they pass the selection criteria for Abell clusters in 3-D.

The last example in Fig. A.6 with $S = 3.75$ does not correspond to a 3-D cluster or group with mass exceeding $M_{3-\mathrm{D}} = 10^{13} \, h^{-1} \, M_{\odot}$. Instead, one sees a large amount of contaminating matter and smaller sub-clumps. This material is responsible for the signal in the S map. The velocity distribution is characterised by three peaks with dispersion $\sigma_{2-\mathrm{D}} = 1235 \, \mathrm{km \, s^{-1}}$, skewness $\mathcal{S}_{2-\mathrm{D}} = -0.30$, and curtosis $\mathcal{K}_{2-\mathrm{D}} = -1.39$. Obviously, the contamination along the line-of-sight is large enough to lead to the detection of an Abell cluster with richness class $\mathcal{R} = 0$.
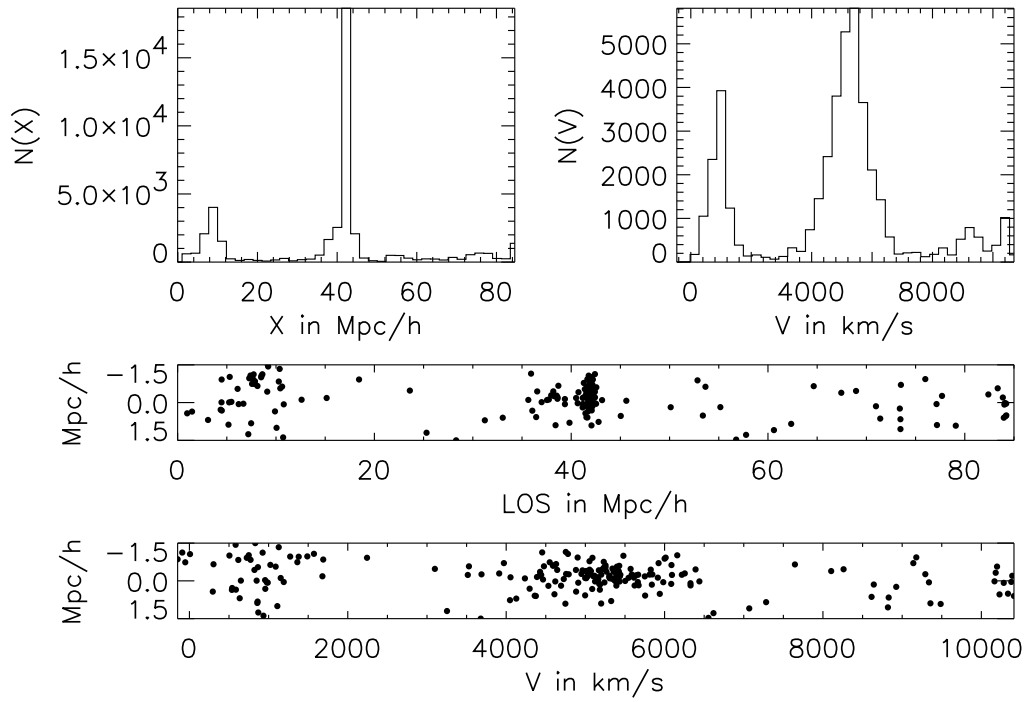
Figure A.3: Structure of a moderately contaminated 3-D cluster with a $S = 4.7$. See the caption of Fig. 4.6 for a description of the panels.
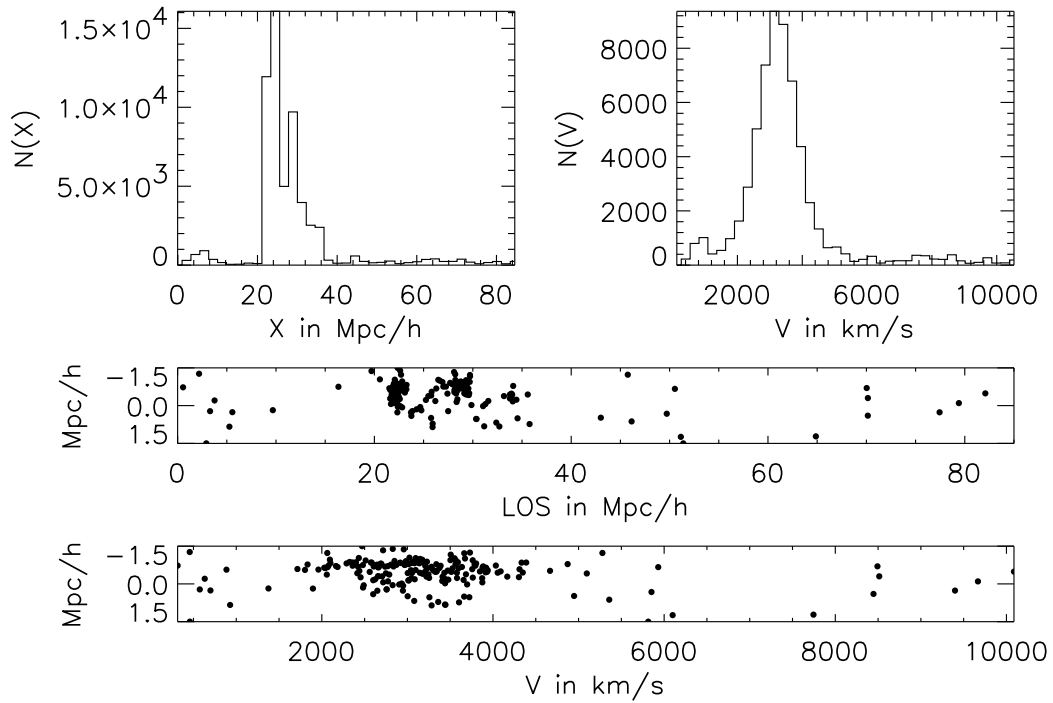

Figure A.4: Structure of a moderately contaminated 3-D cluster with $S = 4.4$. See the caption of Fig. 4.6 for a description of the panels.
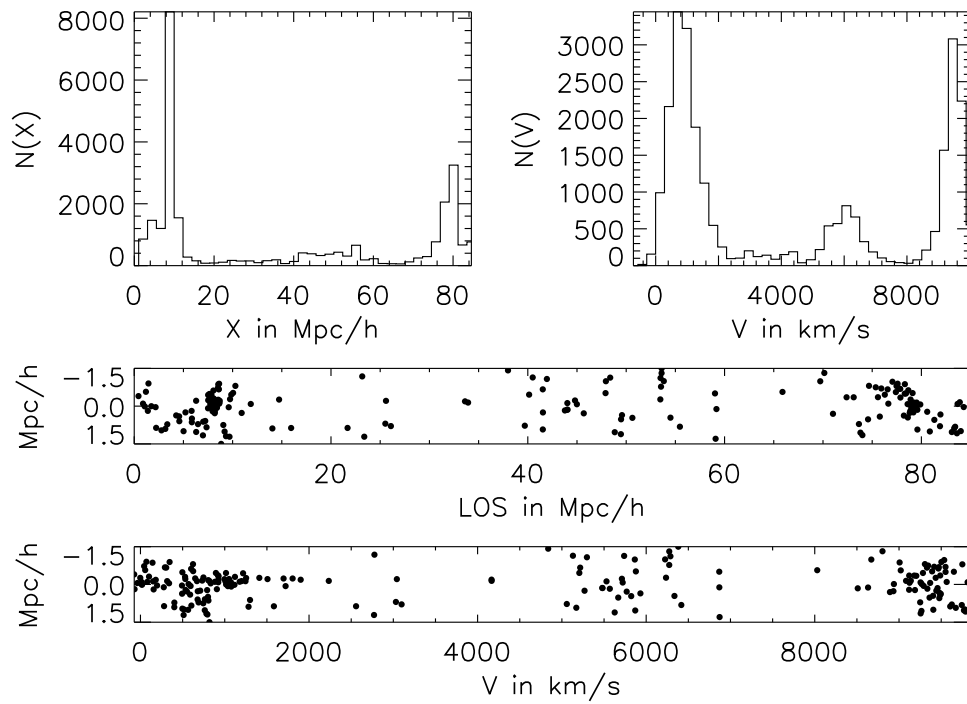
Figure A.5: Structure of moderately large 3-D group with S between 3 and 4. The 3-D object is contaminated by projection along the line-of-sight, leading to an increased S of 3.4. See the caption of Fig. 4.6 for a description of the panels.
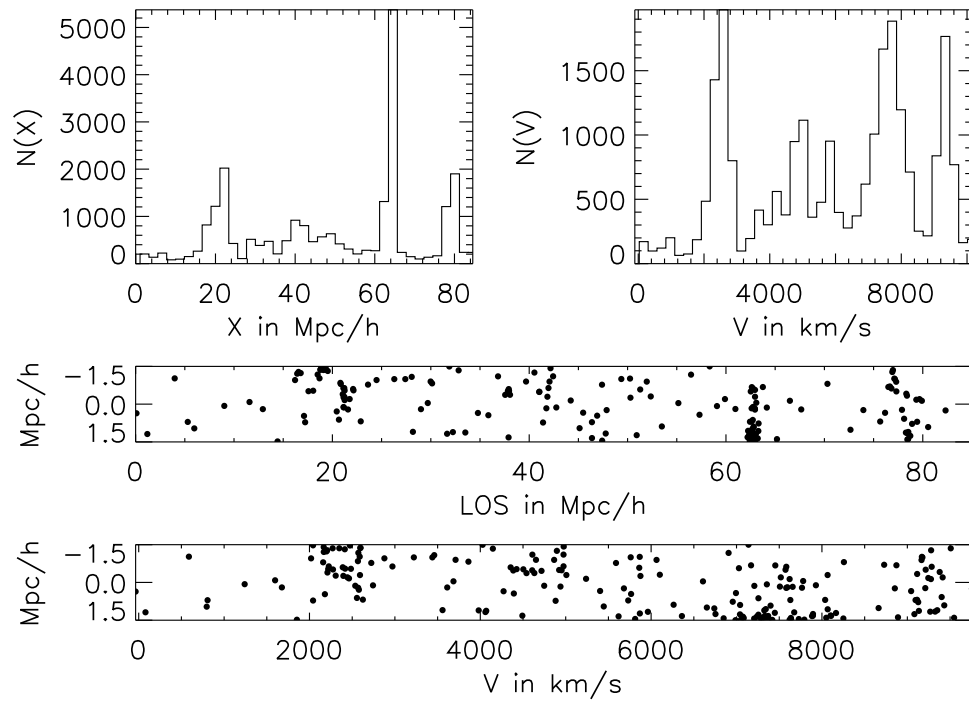
Figure A.6: Structure of a spuriously detected object which does not correspond to a 3-D cluster. The high contamination along the line-of-sight leads to $S = 3.8$. See the caption of Fig. 4.6 for a description of the panels.

Bibliography

Abell G.: *ApJS* **3** (1958) 211

Bahcall N.: *ARAA* **26** (1988) 631

Banday A. et al.: *ApJ* **475** (1997) 393

Bardeen J., Bond J., Kaiser N. & Szalay A.: *ApJ* **304** (1986) 15

Bartelmann M.: *A&A* **303** (1995) 643

Bartelmann M. et al.: *A&A* **330** (1998) 1

Bartelmann M. & Schneider P.: *A&A* **345** (1999a) 17

Bartelmann M. & Schneider P.: *submitted to Phys. Rep.* (1999b)

Bartelmann M. & Steinmetz M.: *MNRAS* **283** (1996) 431

Bernardeau F.: *A&A* **338** (1998) 375

Bernardeau F., van Waerbeke L. & Mellier Y.: *A&A* **322** (1997) 1

Binney J. & Tremaine S.: *Galactic Dynamics*, Princeton University Press, Princeton, New Jersey, (1987)

Binney J.J., Davies R.L. & Illingworth G.D.: *ApJ* **361** (1990) 78

Birkinshaw M.: *Phys. Rep.* **310** (1999) 97

Blandford R. & Narayan R.: *Ann. Rev. Astr. Ap.* **30** (1992) 311

Blandford R., Saust A., Brainerd T. & Villumsen J.: *MNRAS* **251** (1991) 600

Bond J. & Efstathiou G.: *ApJ* **285** (1984) L45

Carlstrom J. et al.: Imaging the Sunyaev–Zel'dovich Effect, in: *Particle Physics and the Universe* (Edited by Bergstrom L., Carlson P. & Fransson C.), Nobel Symposium, 1999

Cen R.: *ApJ* **485** (1997) 39

Cole S. & Lacey C.: *MNRAS* **281** (1996) 716

Couchman H., Thomas P. & Pearce F.: *ApJ* **452** (1995) 797

Couchman H., Thomas P. & Pearce F.: *astro-ph/9603116* (1996)

Davis M., Efstathiou G., Frenk C. & White S.: *ApJ* (1985)

Diaferio A., Kauffmann G., Colberg J. & White S.: *MNRAS* (1999)

Dolag K., Bartelmann M. & Lesch H.: *A &A* **348** (1999) 351

Eke V., Cole S. & Frenk C.: *MNRAS* **282** (1996) 263

Erben T. et al.: *submitted to A&A* (1999) astro–ph/9907134

Etherington I.: *Phil. Mag.* **15** (1933) 761

Fahlman G., Kaiser N., Squires G. & Woods D.: *ApJ* **437** (1994) 56

Fort B. & Mellier Y.: *AARA* **5** (1994) 239

Frenk C., White S., Efstathiou G. & Davis M.: *ApJ* **351** (1990) 10

Gaztanaga E. & Bernardeau F.: *A&A* **331** (1998) 829

Gregory P. & Condon J.: *ApJS* **75** (1991) 1011

Gunn J. & Gott J.: *ApJ* **176** (1972) 1

Gunn J.E.: *ApJ* **150** (1967a) 737

Gunn J.E.: *ApJ* **147** (1967b) 61

van Haarlem M., Frenk C. & White S.: *MNRAS* **287** (1997) 817

Hu W. & Tegmark M.: *ApJ* **514** (1999) L65

Hui L.: *ApJ* **5149** (1999) L9

Jain B. & Seljak U.: *ApJ* **484** (1997) 560

Jain B. & van Waerbeke L.: *accepted by ApJL* (1999)

Jain B., Seljak U. & White S.: *accepted by ApJ* (1999) astro–ph/9901191, JSW

Jenkins A. et al.: *ApJ* **499** (1998) 20

Kaiser N.: *ApJ* **388** (1992) 272

Kaiser N.: *ApJ* **439** (1995) L1

Kaiser N.: *ApJ* **498** (1998) 26

Kaiser N. et al.: *submitted to ApJ* (1998) astro–ph/9809268

Kaiser N. & Squires G.: *ApJ* **404** (1993) 441

Kaiser N., Squires G. & Broadhurst T.: *ApJ* **449** (1995) 460

Kauffmann G., Colberg J., Diaferio A. & White S.: *MNRAS* **303** (1999a) 188

Kauffmann G., Colberg J., Diaferio A. & White S.: *MNRAS* **307** (1999b) 529

Kolb E. & Turner M.: *The Early Universe*, Addison Wesley, Reading, Massachusetts, (1972)

Kormendy J. & Knapp G. (Eds.): *Dark matter in the universe*, Proc. IAU Symp. 117, Boston, (1987)

Kruse G. & Schneider P.: *MNRAS* **302** (1999a) 821, (KS1)

Kruse G. & Schneider P.: *accepted for publication in MNRAS* (1999b) astro–ph/9904192, (KS2)

Longair M.: *Our evolving universe*, CUP, Cambridge, (1996)

Lucy L.B.: *AJ* **79** (1974) 745

Lucy L.B.: *A&A* **289** (1994) 983

Marzke R., Huchra J. & Geller M.: *ApJ* **428** (1994) 43

Mellier Y., van Waerbeke L., Bernardeau F. & Le Fevre O.: Weak Lensing with MEGA-CAM and the VLT, in: *Looking Deep in the Southern Sky*, 1999 p. 59

Miralda-Escude J.: *ApJ* **380** (1991) 1

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

NAG, *User's Manual*, Numerical Algorithms Group Ltd., mark 17 edition, (1995)

Narayan R. & Bartelmann M.: Lectures on Gravitational Lensing, in: *1995 Jerusalem Winter School* (Edited by Dekel A. & Ostriker J.), 1997

Navarro J., Frenk C. & White S.: *ApJ* **462** (1996) 563, NFW

Navarro J., Frenk C. & White S.: *ApJ* **490** (1997) 493, NFW

Oort J.: *Ann. Rev. Astron. Astrophys.* **21** (1983) 373

Padmanabhan T.: *Structure formation in the universe*, CUP, Cambridge, (1995)

Peacock J.: *Cosmological Physics*, CUP, Cambridge, (1999)

Peacock J. & Dodds S.: *MNRAS* **280** (1996) L19

Pearce F. & Couchman H.: *NewA* **2** (1997) 411

Peebles P.: *The Large Scale Structure of the Universe*, Princeton University Press, Princeton, New Jersey, (1980)

Peebles P.: *Principles of Physical Cosmology*, Princeton University Press, Princeton, New Jersey, (1993)

Pen U., Seljak U. & Turok N.: *Phys. Rev. Lett.* **79** (1997) 1611

Press W. & Schechter P.: _ApJ_ **187** (1974) 425

Reblinsky K. & Bartelmann M.: _A&A_ **345** (1999a) 1

Reblinsky K. & Bartelmann M.: in: _Gravitational Lensing: Recent Progress and Future Goals_, Proc. Conf. Boston, 1999

Reblinsky K. & Bartelmann M.: _in preparation_ (2000)

Reblinsky K., Kruse G., Jain B. & Schneider P.: _A&A_ **351** (1999) 815

Richardson W.: _J. Opt. Soc. Am._ **62** (1972) 55

Rood H.: _Rep. Prog. Phys._ **44** (1981) 1077

Rybicki G. & Lightman A.: _Radiative Processes in Astrophysics_, John Wiley & Sons, New York, (1979)

Sarazin C.: _Rev. Mod. Phys._ **58** (1986) 1

Sarazin C.: _X–ray emissions from clusters of galaxies_, CUP, New York, (1988)

Schechter P.: _ApJ_ **203** (1976) 297

Schneider P.: _A&A_ **302** (1995) 639

Schneider P.: _MNRAS_ **283** (1996) 837, S96

Schneider P. & Kneib J.: in: _The Next Generation Space Telescope: Science Drivers and Technological Challenges_, 34th Liège Astrophysics Colloquium, 1998 p. 89

Schneider P. & Weiss A.: _ApJ_ **330** (1988) 1

Schneider P., Ehlers J. & Falco E.: _Gravitational Lenses_, Springer-Verlag, Berlin, (1992)

Schneider P., van Waerbeke L., Jain B. & Kruse G.: _MNRAS_ **296** (1998) 873, SvWJK

Schramm T. & Kayser R.: _A&A_ **299** (1995) 1

Sciama D.: _Modern cosmology and the dark matter problem_, CUP, Cambridge, (1993)

Scoccimarro R. & Frieman J.: _ApJ_ **520** (1999) 35

Seitz C. & Schneider P.: _A&A_ **297** (1995) 287

Seitz C. & Schneider P.: _A&A_ **318** (1997) 687

Seitz S. et al.: _MNRAS_ **298** (1998) 945

Seitz S., Schneider P. & Ehlers J.: _Class. Quant. Grav._ **11** (1994) 2345

Sexl R. & Urbantke H.: _Gravitation und Kosmologie_, Spektrum Akademischer Verlag, (1995)

Shu F.H.: _The Physics of Astrophysics, Vol. 1 and 2_, University Science Books, (1991)

Silk J.: *A short history of the universe*, W.H. Freeman and Company, New York, (1994)

Silk J. & White S.: *ApJ* **226** (1978) L103

Sunyaev R. & Zel'dovich Y.: *Comments Astrophys. Space Phys.* **4** (1972) 173

Sunyaev R. & Zel'dovich Y.: *Comments Astrophys. Space Phys.* **2** (1978) 66

Sunyaev R. & Zel'dovich Y.: *ARAA* **18** (1980) 537

Swarztrauber P.: *Fast Poisson Solvers*, No. 24 in Studies in Numerical Analysis, Academic Press, 1984 p. 319

Treyer M. et al.: *ApJ* **509** (1998) 531

Trimble V.: *ARAA* **25** (1987) 425

Tyson J., Valdes F. & Wenk R.: *ApJ* **349** (1990) L1

van Waerbeke L., Bernardeau F. & Mellier Y.: *A&A* **342** (1999) 15

Weinberg S.: *Gravitation and Cosmology*, John Wiley & Sons, (1972)

White S.: Cosmology and Large-Scale Structure, in: *Les Houches, Session LX* (Edited by Schaeffer R., Silk J., Spiro M. & Zinn-Justin J.), Elsevier Science, Elsevier B.V., 1996

White S., Efstathiou G. & Frenk C.: *MNRAS* **262** (1993) 1023

Yahil A. & Vidal N.: *ApJ* **214** (1977) 347

Zaroubi S., Squires G., Hoffman Y. & Silk J.: *ApJL* **500** (1998) L87

Lebenslauf

Katrin Reblinsky

Persönliche Daten

geboren am	22. August 1970 in Lübeck
Staatsangehörigkeit	deutsch
Familienstand	ledig

Hochschulausbildung

seit Jan. 1997	Promotionsstipendium am Max–Planck–Institut für Astrophysik in Garching bei München
Sept. 1996	Diplomhauptprüfung, Nebenfach: Mathematik
Aug. 1995 – Aug. 1996	Diplomarbeit in theoretischer Physik über *Geladene Schwarze Löcher in nichtabelschen Feldtheorien* betreut von Prof. Dr. J. Kunz–Drolshagen
Mai 1994 – Juli 1994	Studienarbeit an der TRI-University Meson Facility (TRIUMF) in Vancouver, Kanada über *Spinobservable beim protoninduzierten Nukleonknockout an polarisiertem ^{3}He bei 500 MeV*, betreut von Prof. Dr. O. Häusser
Sept. 1993 – April 1994	Austauschstudent (4th year honours program) an der University of British Columbia (UBC) in Vancouver, Kanada
Sept. 1992	Vordiplom
Okt. 1990 – Sept. 1996	Physikstudium an der Carl von Ossietzky Universität in Oldenburg/Oldb.

Schulausbildung

1981 – 1990	Leibniz–Gymnasium Bad Schwartau. Abschluss: Allgemeine Hochschulreife
1977 – 1981	Grundschule Bad Schwartau

Danksagung