

# **Applications of the weak gravitational lens effect**

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

vorgelegt von Thomas Erben  
aus München

München, den 25. Juli 2000

1. Gutachter: Prof. Dr. Peter Schneider

2. Gutachter: Prof. Dr. Ralf Bender

Tag der mündlichen Prüfung: 14.12.2000

# Contents

<b>1</b>	<b>Einleitung und Zusammenfassung</b>	<b>5</b>
<b>2</b>	<b>Introduction</b>	<b>7</b>
<b>3</b>	<b>Theoretical groundwork</b>	<b>9</b>
3.1	Standard Cosmology . . . . .	9
3.1.1	Redshift and Distances . . . . .	11
3.1.2	The Cosmic Microwave Background Radiation . . . . .	12
3.2	Structure formation . . . . .	13
3.3	Gravitational lensing . . . . .	15
3.3.1	Light propagation in inhomogeneous universes . . . . .	16
3.3.2	The lensing effect of isolated mass concentrations . . . . .	17
3.3.3	The lensing effect of large-scale structures . . . . .	23
<b>4</b>	<b>Weak gravitational lensing</b>	<b>25</b>
4.1	Lensing quantities and image ellipticities . . . . .	26
4.2	The basic ideas of weak gravitational lensing studies . . . . .	28
<b>5</b>	<b>Combining magnification and shear information to constrain cluster mass profiles</b>	<b>31</b>
5.1	Maximum Likelihood estimators . . . . .	31
5.2	Motivation and basic ideas for our work . . . . .	32
5.3	The calculation of the likelihood functions . . . . .	35
5.4	The expected average log-likelihood . . . . .	36
5.5	The lens models . . . . .	38
5.6	Results from the average log-likelihood analysis . . . . .	42
5.7	Error analysis . . . . .	42
5.8	The effect of uncertainties in the unlensed number counts . . . . .	47
5.9	Conclusions . . . . .	51
<b>6</b>	<b>Technical aspects of weak lensing observations</b>	<b>54</b>
6.1	Practical problems for weak lensing studies . . . . .	54
6.2	The KSB algorithm . . . . .	58
6.3	Practical problems with the KSB formulae . . . . .	61
6.4	Motivation for our simulations . . . . .	63
6.5	The introduction of a weighting scheme in shear estimates . . . . .	64
6.6	Semi-analytical simulations with Gaussian profiles . . . . .	64
6.6.1	Estimation of $P^g$ . . . . .	68
6.6.2	Weighting scheme . . . . .	68
6.6.3	Accuracy of shear estimates . . . . .	68
6.6.4	Summary . . . . .	70
6.7	The SkyMaker simulations . . . . .	72
6.7.1	The image characteristics of our SkyMaker simulations . . . . .	72
6.7.2	Results of the SkyMaker simulations . . . . .	73
6.8	Conclusions . . . . .	75
<b>7</b>	<b>Detection of a dark matter concentration near the cluster Abell 1942</b>	<b>82</b>
7.1	Finding matter concentrations with weak gravitational lensing . . . . .	82
7.2	The data . . . . .	84
7.3	The detection of a dark cluster candidate . . . . .	86
7.4	Properties of our dark clump candidate . . . . .	87
7.5	Conclusions . . . . .	93

<b>8</b>	<b>First detection of cosmic shear</b>	<b>97</b>
8.1	Theoretical predictions for cosmic shear . . . . .	97
8.2	The data . . . . .	101
8.3	Analysis and measured signal . . . . .	101
8.4	Analysis of possible systematics . . . . .	102
8.5	Cosmological constraints . . . . .	104
<b>9</b>	<b>Outlook</b>	<b>113</b>
<b>A</b>	<b>The Stuff program</b>	<b>119</b>
<b>B</b>	<b>Object detection and selection</b>	<b>119</b>

# 1 Einleitung und Zusammenfassung

Der schwache Gravitationslinseneffekt ist im letzten Jahrzehnt zu einem der wichtigsten Werkzeuge zur Vermessung der Materieverteilung unseres Universums geworden. Albert Einstein sagte bereits Anfang dieses Jahrhunderts im Rahmen seiner Allgemeinen Relativitätstheorie Lichtablenkung durch Masse voraus. Hierdurch sehen wir das Licht weit entfernter Objekte nicht in seiner intrinsischen Erscheinungsform, sondern durch Vordergrundmassen verzerrt. Seitdem die Form von lichtschwachen Galaxien auf Aufnahmen von hochauflösenden CCD-Kameras mit hoher Genauigkeit meßbar ist, können wir die kohärente Verzerrung von Hintergrundquellen zum Studium von Vordergrundmassen nutzen. Hierbei ist es nicht nur möglich, die Gesamtmasse zu bestimmen, durch die das Licht abgelenkt wird, sondern es kann sogar deren zwei-dimensionale Projektion aufgelöst werden. Da dieser Linseneffekt lediglich auf die Masse eines Objekts sensitiv ist, haben wir hier das ideale Werkzeug, um insbesondere die so-genannte *dunkle Materie* von Objekten, die kein Licht im elektromagnetischen Spektrum ausstrahlt, aufzuspüren.

Die Analyse des Linsensignals bekannter Galaxienhaufen hat bereits wichtige neue Erkenntnisse erbracht, z.B. die Existenz massiver Haufen bei hoher Rotverschiebung. Seit Kaiser & Squires 1993 gezeigt haben wie die Masse eines Galaxienhaufens mit Hilfe des Linseneffekts parameterfrei rekonstruiert werden kann, haben zahlreiche Gruppen sowohl die Theorie dieser Rekonstruktionen weiterentwickelt, als auch Techniken zur Datenanalyse optimiert. Die Analyse des Effekts in Galaxienhaufen ist inzwischen beinahe zur Routine geworden.

In den letzten Jahren ist das Interesse an der Verwendung des Linseneffekts zur Untersuchung von Massen in unserem Universum stark angestiegen. Ein sehr junger Teilbereich ist die Untersuchung des Linseneffekts einzelner Galaxien (das so-genannte *Galaxy-Galaxy Lensing*). Der Linseneffekt einzelner Galaxien auf die Form dunkler Hintergrundgalaxien ist so klein, daß für ein signifikantes Signal die Verzerrung vieler Galaxien statistisch überlagert werden muß. Desweiteren benötigen wir für eine verlässliche Formbestimmung von Galaxien Datenmaterial von hoher Qualität. Während für die Analyse von Galaxienhaufen Daten mit einem Blickfeld von wenigen Bogenminuten ausreichend sind, benötigen wir für quantitative Galaxy-Galaxy-Lensing Analysen Beobachtungen über einen möglichst großen Winkelbereich. Solche Untersuchungen sind nur mit so-genannten *Weitwinkel-Kameras*, d.h. mit Instrumenten, die ein halbes Quadratgrad des Himmels oder sogar mehr mit einer einzigen Aufnahme abdecken können, in einer vernünftiger Menge an Beobachtungszeit durchführbar.

Seit ca. eineinhalb Jahren hat unsere Linsengruppe Zugang zu mehreren Teleskopen, die mit einer Weitwinkel-Kamera ausgestattet wurden. Der Autor dieser Arbeit hat an folgenden Teilbereichen des schwachen Linseneffekts gearbeitet, die sich die phantastischen neuen Möglichkeiten des *wide-field imaging* zunutze machen:

- Das erste wissenschaftliche Projekt ist rein theoretischer Natur. Eine Vordergrundmasse verzerrt nicht nur die Bilder ihrer Hintergrundobjekte, sondern sie verändert auch deren beobachtete Anzahldichte. Wir haben untersucht, wie diese beiden Linseneffekte theoretisch optimal kombiniert werden können, um parametrisierte Modelle für Haufenprofile einzuschränken. Durch Vielteilchen-Simulationen haben wir gelernt, daß sich Galaxienhaufen über einen großen Massenbereich sehr ähnlich verhalten, sobald wir ihr Massenprofil außerhalb des Zentralbereichs betrachten. Das Profil sollte dort durch ein *universelles Massenprofil* beschrieben werden können. Der Linseneffekt, der direkt sensitiv auf die Masse ist, ist das ideale Werkzeug für Untersuchungen in den äußeren, oft lichtschwachen Bereichen bekannter Galaxienhaufen. Wir konnten in unserer Arbeit zeigen, daß durch die gleichzeitige Nutzung von Objektverzerrung und Änderung der Anzahldichte Massenprofile deutlich besser eingeschränkt werden als durch die Verwendung von nur einem der beiden Effekte. Der praktische Nutzen der Methode für quantitative Analysen muß sich erst noch erweisen. Die Messung des Anzahldichte Effekts benötigt eine hinreichend akkurate Kenntnis über die Dichte ungelinster Objekte. Es ist noch nicht klar, wie genau diese bestimmbar ist. Nichtsdestoweniger wurde unser Algorithmus bereits von einer britischen Gruppe auf Daten angewendet.
- Die Möglichkeit, die dunkle Materie mit dem Linseneffekt direkt zu vermessen, beschränkt

sich nicht auf den Anteil dunkler Materie in bereits bekannten Galaxienhaufen. Insbesondere können wir neuartige, nahezu komplett aus dunkler Materie bestehende, Objekte finden, falls diese existieren. Weitwinkelaufnahmen ermöglichen uns in Zukunft die systematische Suche nach massiven Objekten mit dem Gravitationslinseneffekt, ohne auf eine Korrelation zwischen Masse und Licht angewiesen zu sein. Bei der Analyse der ersten Weitwinkelaufnahmen, die uns 1997 zur Verfügung standen, haben wir einen ersten Kandidaten für einen Haufen, der hauptsächlich aus dunkler Materie besteht, gefunden. Das starke Linsensignal, um dessen Position keine Überdichte von Galaxien im Optischen gefunden wurde, konnte auf einer zweiten Weitwinkelaufnahme, die ein Jahr nach unserer ersten Analyse aufgenommen wurde, bestätigt werden. Die Natur unserer Entdeckung ist bis jetzt ungeklärt, und es wurde ein umfangreiches Beobachtungsprogramm zur Klärung dieser Frage begonnen. Eine zweite Gruppe hat 2000 über die Entdeckung eines weiteren dunklen Haufens mit dem Linseneffekt berichtet.

- Der Linseneffekt ist in der Lage, die Materiedichte von großräumigen Strukturen (den *large scale structures*) in unserem Universum direkt zu vermessen. Theoretische Arbeiten zur möglichen Entdeckung des Linseneffekts dieser Strukturen sagen lediglich einen sehr kleinen Effekt voraus. Desweiteren ist eine große Datenmenge zur statistisch signifikanten Messung des Linsensignals nötig. Andererseits ist die Möglichkeit, die statistischen Eigenschaften der Materieverteilung im Universum direkt zu bestimmen, ausgesprochen reizvoll. In diesem Gebiet brachten die neuen Weitwinkel Kameras den entscheidenden Durchbruch. Über einen Zeitraum von vier Jahren konnte die nötige Menge an hochqualitativen Daten gesammelt werden, und wir konnten in einer internationalen Arbeitsgruppe die winzigen Verzerrungen von Galaxien durch großräumige Strukturen, die so genannte *kosmische Scherung*, signifikant nachweisen.
- Speziell für den Nachweis der kosmischen Scherung war neben dem Bedarf einer großen Datenmenge eine sehr genaue Formbestimmung und somit eine verlässliche Messung des Verzerrungseffekts von Objekten das Hauptproblem. Neben der Verzerrung durch Gravitationslinsen verändern Lichtstreuungs-Effekte in der Erdatmosphäre und durch die Teleskopoptik die Form von Objekten maßgeblich. Diese Störeinflüsse übersteigen oft sogar das erwartete Linsensignal. Um Vertrauen in unsere wissenschaftlichen Resultate zu gewinnen, wurden umfangreiche Simulationen mit dem derzeit favorisierten Algorithmus zur Verzerrungsmessung unter Berücksichtigung bekannter Störungen durchgeführt. Wir konnten zeigen, daß wir die Linsenverzerrung mit hinreichender Genauigkeit messen können. Desweiteren konnten wir einen Algorithmus entwerfen, die riesige Datenmenge von Weitwinkel Teleskopen nahezu automatisch für Linsenanalysen zu verarbeiten.

## 2 Introduction

In the past decade, weak gravitational lensing has become one of the most promising tools to study the matter content in our universe. The effect utilises the fact that a foreground mass distorts the light from galaxies in its background by its gravitation. It was realised that this distortion is coherent and measurable. It allows the direct reconstruction of the projected, two-dimensional surface mass density by which it is caused. The study of well-known galaxy clusters has provided us with invaluable information about the matter content and the matter distribution in these systems. It also revealed new, important facts, for instance the existence of massive galaxy clusters at high redshifts. In this field of cluster mass reconstructions a lot of theoretical and observational work has been done to improve the techniques and to optimally use the available data. The effect is routinely analysed now.

With time, interest has progressively shifted to masses having only a tiny distortion effect on the shape of background galaxies. To these studies belongs the so-called galaxy-galaxy lensing effect which provides constraints on dark matter haloes of individual galaxies. Background galaxies do not have a round shape intrinsically, and to detect the lens effect of individual galaxies, the signal around many of these objects has to be statistically stacked. Moreover, the reliable measurement of galaxy shapes needs observational data of high quality. While for lens studies of galaxy clusters, a relatively small field-of-view (several arcminutes) of data around the cluster centre is sufficient, qualitative studies of the galaxy-galaxy lensing require cameras covering a large field-of-view (30 arcminutes or more) with a single exposure, if these projects are to be executed in a reasonable observing time. Several telescopes have been equipped with such a *wide field* camera in the past years.

The major part of this thesis covers applications of the weak lensing effect utilising the fantastic new possibilities wide-field imaging offers. The author has worked on the following aspects:

- The first part is not directly related to wide-field imaging. Besides the distortion of background galaxies, gravitational lensing also changes the number counts of galaxies behind the mass (the so-called magnification effect). It was investigated in a theoretical work how the two lensing signatures, the distortion of shape and the change in number counts can be combined to constrain parametric models of cluster mass profiles. As simulations of galaxy-cluster formation suggest that massive objects follow a universal profile this work is of particular interest if it is applied to a well defined statistical sample of clusters. As the lens effect is only sensitive to the mass, we do not depend on the light distribution to trace the shape of these objects.
- The last comment in the previous paragraph immediately provokes turning around the chain of argument. Instead of using the lens effect to study already well-known objects, we can use its sensitivity to mass to detect new, up to now undiscovered mass distributions. This especially opens up the possibility to discover massive objects completely consisting of dark matter with very little light emission in the electromagnetic spectrum. Applying a weak lensing cluster finder to wide-field imaging data immediately allows us to derive the cosmological mass function within the analysed area. Of course it is not clear whether objects consisting only of dark matter exist, but the lens effect is the tool of choice for this research. We will report on the first possible candidate of a dark clump within this thesis.
- Besides massive objects, lensing also traces directly the matter distribution of large-scale structures. Theoretical work on the lensing effect of large-scale structure suggests only a tiny correlated distortion of light emitting galaxies and even in 1999 it was unclear whether the very challenging measurement can be done with currently available data-sets and data processing techniques. Also in this area, wide-field imaging brought the decisive breakthrough. In an international collaboration, with the analysis of 1.7 square degrees of high-quality data that had to be accumulated over four years, we have been able to obtain a highly significant measurement of the tiny coherent ellipticity distortions caused by large-scale structures - the so-called cosmic shear. The presentation of this work is part of this thesis.

- Besides the high data quality that weak lensing studies require, the key to a successful experiment is the reliable conversion of measured image shapes to the so-called *shear*, characterising the distortion caused by a gravitational lens. To be able to process large amounts of wide-field imaging data in an automatic way and to have confidence in the resulting object catalogs and hence in the scientific results, we performed numerous simulations with state-of-the-art weak lensing tools. We will show how accurately we can currently measure weak lensing distortions and that weak lensing data can indeed be processed nearly completely automatically.

The thesis is organised as follows: The first two sections give an overview of the cosmological background, an introduction to lensing and the weak gravitational lens effect. The third section presents the work on the combination of distortion and magnification effects to constrain parametrised cluster mass models. The next section deals with technical problems to measure galaxy shapes in observational praxis. In the following section we report on the first possible discovery of a massive dark clump detected purely with weak lensing techniques. Finally, our measurement of the cosmic shear effect, the lensing caused by large-scale structure, is described. We finish with an outlook of the work presented.



### 3 Theoretical groundwork

**Summary:**

*In this chapter we summarise the theory of cosmology and gravitational lensing necessary to follow the rest of the thesis. First, the so-called “standard cosmology” that was introduced by Albert Einstein is described. It is a solution of his general relativistic field equations assuming that the space we live in is isotropic and homogenous. The evolution of localised matter concentrations like single galaxies or galaxy clusters, that make the universe appear very clumpy on small scales, cannot be described within its framework. Hence, the model of an isotropic and homogenous universe has to be complemented by a theory for structure formation. After a discussion of some important, unsolved problems concerning the nature and the amount of matter in the universe, gravitational lensing will be discussed. A thorough introduction to cosmology can be found in [52] for instance.*

#### 3.1 Standard Cosmology

Albert Einstein’s general relativity describes the universe as a four-dimensional space-time, characterised by a metric tensor  $g_{\alpha\beta}$ . The dynamics of this tensor, containing the geometrical properties of space-time, are directly determined by the present matter and radiation density. The quantities are connected via Einstein’s field equations:

$$G_{\alpha\beta} = \frac{8\pi G}{c^2} T_{\alpha\beta} + \Lambda g_{\alpha\beta}. \quad (1)$$

$c$  is the vacuum speed of light,  $G_{\alpha\beta}$  is a function of first and second derivatives of the metric tensor,  $\Lambda$  is the so-called cosmological constant and  $T_{\alpha\beta}$  is the stress energy tensor for the matter and radiation present.  $T_{\alpha\beta}$  has the form of a homogenous fluid, characterised by its density  $\rho(\vec{x}, t)$  and pressure  $p(\vec{x}, t)$ .

To arrive at a metric tensor describing the universe one puts strong constraints on its possible form by the *cosmological principle*. It consists of the following two postulates:

- No point in the universe can be distinguished from any other point (The universe everywhere looks the same, no matter where we set our origin).
- The distribution of matter around our position is homogenous and isotropic in space.

Robertson and Walker (1935, 1936) have derived the most general line element satisfying these constraints:

$$ds^2 = g_{\alpha\beta} g^{\alpha\beta} = c^2 dt^2 - R(t)^2 du^2; \quad du^2 = dw^2 + f_K^2(w)(d\phi^2 + \sin^2(\theta)d\theta^2),^* \quad (2)$$

where  $t$  is the time,  $R(t)$  the cosmic scale factor and  $du$  the line element of a homogenous and isotropic three-dimensional space. This line element consists of angular coordinates  $\phi$  and  $\theta$  and a radial coordinate  $w$ . An isotropic three-dimensional space has a constant curvature  $K$ , and  $f_K(w)$  takes the form

$$f_K(w) = \begin{cases} K^{-1/2} \sin(K^{1/2}w) & (K > 0) \text{ (closed space)} \\ w & (K = 0) \text{ (flat space)} \\ (-K)^{-1/2} \sinh[(-K)^{1/2}w] & (K < 0) \text{ (open space)} \end{cases}. \quad (3)$$

The three-dimensional space, described by  $du$ , directly scales with  $R(t)$ . From the homogeneity postulate follows that this factor can depend only on  $t$ . Otherwise the cosmic expansion would be different at different places. Similarly, the matter density  $\rho(t)$  and the pressure  $p(t)$  in the universe

---

\*The metric is only unique up to an arbitrary coordinate transformation. So the reader is likely to find it in a different form when consulting the literature. We used a form most convenient in later sections.

can be functions of time only. With these constraints, the field equations [eq. (1)] reduce to the so-called ‘‘Friedmann equations’’. They determine the time evolution of  $R(t)$  and relate the space curvature  $K$  to matter and radiation,

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{R^2} + \frac{\Lambda}{3}, \quad (4)$$

and

$$\left(\frac{\ddot{R}}{R}\right) = -\frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}, \quad (5)$$

where dots denote derivatives with respect to  $t$ . Once we have specified an equation of state  $p = p(\rho)$  and an initial condition for  $R$  the Friedmann equations determine the temporal evolution of our model universe. Ordinary matter, which is often called dust, has  $p \ll \rho c^2$  (so we use  $p = 0$  for the equation-of-state of ordinary matter), while radiation and other forms of relativistic matter have  $p = \rho c^2/3$ . As an initial condition we use  $R(t_0) = 1$  for the present time  $t_0$ .

The cosmological constant  $\Lambda$  was originally introduced by Einstein to allow a static universe ( $\dot{R} = \ddot{R} = 0$  for all times). Today we know that we live in a dynamic, always expanding universe. The velocity of this expansion is characterised by the so-called *Hubble constant*  $H_0 = H(t_0)$

$$H_0 = \frac{\dot{R}_0}{R_0}. \quad (6)$$

Measurements of the Hubble constant today are within the limits  $50 \text{ km}/(\text{s Mpc}) \leq H_0 \leq 100 \text{ km}/(\text{s Mpc})$ . So, one usually writes  $H_0 = 100h \text{ km}/(\text{s Mpc})$  with  $h \in [0.5; 1]$ . If not stated otherwise, we use  $h = 0.5$  throughout this thesis. With the further definitions

$$\rho_{\text{cr}} := \frac{3H_0^2}{8\pi G} \quad \Omega_0 := \frac{\rho_0}{\rho_{\text{cr}}} \quad \Omega_\Lambda = \frac{\Lambda}{3H_0^2} \quad (7)$$

we can rewrite the first Friedmann equation to read in the case of a universe dominated by ordinary matter ( $p=0$ ):

$$H^2(t) = H_0^2 \left( R^{-3}(t)\Omega_0 - R^{-2}(t)\frac{Kc^2}{H_0^2} + \Omega_\Lambda \right). \quad (8)$$

For the present time  $t_0$  this becomes:

$$K = \left(\frac{H_0}{c}\right)^2 (\Omega_0 + \Omega_\Lambda - 1). \quad (9)$$

Since we live in a matter dominated universe today, the sign of the space curvature is determined by the matter density  $\rho$  in units of a critical density  $\rho_{\text{cr}}^\dagger$  and the contribution of  $\Omega_\Lambda$ .

$$\begin{aligned} \Omega = \Omega_0 + \Omega_\Lambda > 1 &\implies \text{closed universe} \\ \Omega = \Omega_0 + \Omega_\Lambda = 1 &\implies \text{flat universe} \\ \Omega = \Omega_0 + \Omega_\Lambda < 1 &\implies \text{open universe.} \end{aligned} \quad (10)$$

Because of the formal equality of  $\Omega_\Lambda$  to the density parameter  $\Omega_0$ , the cosmological constant is interpreted today as vacuum energy density. Although the determination of  $\Omega_0$  and  $\Omega_\Lambda$  is the most fundamental problem, since the above equations have been written down nearly one century ago the issue is still one of the most active fields of current research in theoretical and observational cosmology. Before describing how we try to approach the problem by studying extragalactic objects, we consider some important consequences of the standard cosmology.

---

<sup>†</sup>In the absence of  $\Omega_\Lambda$ ,  $\rho_0 = \rho_{\text{cr}}$  is just the condition for a flat universe.

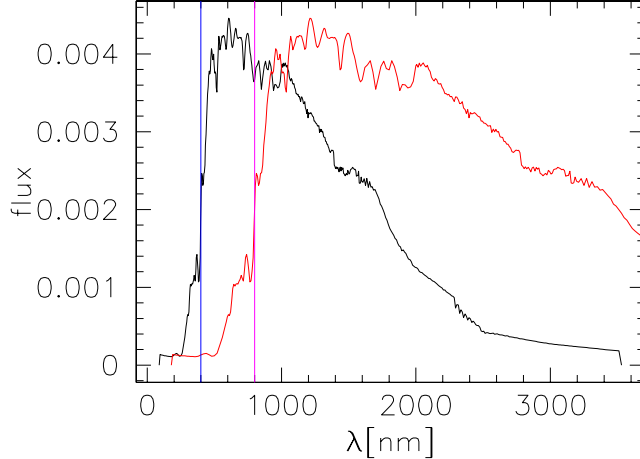


Figure 1: Redshift: The black curve shows a model for the spectral energy distribution (SED) of an elliptical galaxy at  $z = 0$ . The red curve shows the observed flux for a galaxy at redshift  $z = 1$ . The flux is normalised so that the total flux is equal to 1. The blue and magenta lines show the limits of the visible part of the electromagnetic spectrum. We notice that the optical flux is highly dimmed for high-redshift elliptical galaxies and we have to go to the infrared to catch their light. We note also the following: Measuring the flux in a wavelength range  $\lambda_1.. \lambda_2$ , we do not get the flux that the object emits in this range but our measurement depends on the object’s redshift. To calculate the flux in the range  $\lambda_1.. \lambda_2$  out of the observed flux, one needs to know the SED and the redshift. This correction of observed to intrinsic flux is called *K-correction*.

### 3.1.1 Redshift and Distances

When a distant astronomical object emits light of the wavelength  $\lambda_e$  we measure a shifted wavelength  $\lambda_s$  because of the Hubble expansion. The so-called *redshift*  $z$  is defined as

$$z := \frac{\lambda_s - \lambda_e}{\lambda_e}. \quad (11)$$

We can show that the scale factor  $R$  is connected to the redshift by the simple relation

$$\frac{R_0}{R} = 1 + z. \quad (12)$$

The redshift of an object can directly be observed via the *spectral energy distribution* which is the emitted energy per time and per frequency interval (see Fig. 1). As the redshift is an observable we usually transform dependencies of the cosmic time  $t$  to dependencies of  $z$  via eq. (12).

As light from highly redshifted objects travels longer to us than that from low redshift objects, the term redshift is often understood as a synonym for distance to an object. To be more precise, we have to define the term “distance” with the metric of the space time. We shortly introduce the four most common distance measures in cosmology: the proper distance, the comoving distance, the angular diameter distance and the luminosity distance:

The *proper distance* or physical distance  $D_{\text{prop}}(z_1, z_2)$  between two events  $z_1$  and  $z_2$ , where  $z_2 > z_1$ , 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  $dD_{\text{prop}} = -c dt = -c dR \dot{R}^{-1} = -c dR (RH)^{-1}$ . This leads to

$$D_{\text{prop}}(z_1, z_2) = \frac{c}{H_0} \int_{R(z_2)}^{R(z_1)} [R^{-1}\Omega_0 + (1 - \Omega_0 - \Omega_\Lambda) + R^2\Omega_\Lambda]^{-1/2} dR. \quad (13)$$

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  $z_1$ ,  $dD_{\text{com}} = dw$ . Since light rays propagate with  $ds = 0$ , we derive  $cdt = -Rdw$  from the metric, and therefore  $dD_{\text{com}} = -R^{-1}cdt = -cdR(R\dot{R})^{-1} = cdR(R^2H)^{-1}$ . Thus

$$D_{\text{com}}(z_1, z_2) = \frac{c}{H_0} \int_{R(z_2)}^{R(z_1)} [R\Omega_0 + R^2(1 - \Omega_0 - \Omega_\Lambda) + R^4\Omega_\Lambda]^{-1/2} dR. \quad (14)$$

The *angular-diameter distance*  $D_{\text{ang}}(z_1, z_2)$  is the most important distance measure for the rest of this thesis. It is defined as in Euclidian space. It relates the physical cross section  $\delta A$  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

$$\frac{\delta A}{4\pi R^2(z_2) f_K^2(w_2 - w_1)} = \frac{\delta\omega}{4\pi}, \quad (15)$$

where  $R(z_2)$  is the scale factor at emission time and  $f_K(w_2 - w_1)$  is the radial coordinate distance between the observer and the source, it follows that

$$D_{\text{ang}}(z_1, z_2) = \left( \frac{\delta A}{\delta\omega} \right)^{1/2} = R(z_2) f_K(w_2 - w_1). \quad (16)$$

The *luminosity distance*  $D_{\text{lum}}(z_1, z_2)$  is defined by the relation between the luminosity  $L$  of an object at  $z_2$  and the flux  $S$  received by the observer at  $z_1$ . We can show the following relation of the luminosity distance to the angular diameter distance:

$$D_{\text{lum}}(z_1, z_2) = \left( \frac{R(z_1)}{R(z_2)} \right)^2 D_{\text{ang}}(z_1, z_2). \quad (17)$$

### 3.1.2 The Cosmic Microwave Background Radiation

As another consequence of the standard cosmology we consider the Cosmic Microwave Background radiation. This will not be needed in the rest of the thesis but its discovery is the most impressive confirmation of the cosmological principle and thus for the standard cosmology.

Today we believe that the universe was born in a singular event, the so-called *Big Bang Theory*, that is  $R(t) \xrightarrow{t \rightarrow 0} 0$ . We assume that at some point  $t_e$  in the early development there was a complete equilibrium between matter and radiation that can be described as a Planck distribution with a well defined temperature  $T$ . The photon density at this time was:

$$n(\omega', t_e) d\omega' = \frac{1}{\pi^2 c^3} \frac{\omega'^2 d\omega'}{\exp(\hbar\omega'/kT_e) - 1}. \quad (18)$$

With time, a photon changes its frequency  $\omega$  because of the redshift by

$$\omega = \omega' \frac{R(t_e)}{R(t)}. \quad (19)$$

On the other hand, as all physical distances scale with  $R(t)$ , a density scales as the volume with  $R(t)^{-3}$ . The photon number is conserved and hence

$$n(\omega, t) d\omega = \left[ \frac{R(t_e)^3}{R(t)} \right] n(\omega', t_e) d\omega'. \quad (20)$$

Inserting eq. (18) and eq. (19) leads to

$$n(\omega, t) d\omega = \frac{1}{\pi^2 c^3} \frac{\omega^2 d\omega}{\exp(\hbar\omega/kT) - 1} \quad T(t) = T(t_e) \frac{R(t_e)}{R(t)}. \quad (21)$$

Consequently, the expansion of the universe transforms a Planck distribution again to a Planck distribution with the temperature scaling as  $R(t)^{-1}$ . If our assumptions that the universe originated from a small, hot and very dense region are correct, the radiation from the time when matter and radiation were in equilibrium should still be measurable with a lower temperature. This predicted background radiation was first found by Penzias and Wilson (1965) and confirmed in many subsequent experiments. The measured temperature today is:

$$T(t_0) = 2.728 \pm 0.006K. \quad (22)$$

Measuring the temperature in all possible directions of the sky, we find fluctuations of the order

$$\frac{\Delta T}{T} \leq 10^{-5}. \quad (23)$$

At least at very early times the universe was extremely homogenous. The deviations from complete homogeneity are the reason for the generation of structures like galaxies and galaxy clusters that we see on small scales. The description of the formation of these structures is summarised in the next subsection.

### 3.2 Structure formation

The state-of-the-art theory for structure formation states that the structures we see today originated from small density fluctuations at very early times that evolve according to gravitational instability. The origin of these fluctuations is still under debate. In the currently favoured scenario (*The Inflation theory*) these fluctuations originated from primordial quantum fluctuations in the very early universe that were blown up during an inflationary phase. This inflation theory leads to fluctuations that can be described as a Gaussian random field in the very early stages. Besides the origin of the matter fluctuations there are also problems in understanding their composition. It is now evident that there is a large amount of matter in the universe that does not emit electromagnetic radiation at any wavelength and that shows up only by its gravitational forces, the *Dark Matter*. Some observational facts that can only be explained by dark matter are:

- The rotation curves for spiral galaxies, i.e. the rotation speed of stars and gas as a function of distance from the galaxy centre, are also flat at very large radii. If we assume that all the matter is connected with the luminous core of these objects we would expect a decline of the rotation curve outside the visible matter. This discrepancy is reconciled by the assumption of a dark, quasi-spherical and massive *halo* surrounding the visible part of the galaxy.
- The relative amount of dark matter is even larger for galaxy clusters<sup>‡</sup>. As we will describe in the next section, the gravitational lens effect can reconstruct their mass distributions. These estimates agree fairly well with methods based on assumptions about virial equilibrium for the conversion of the X-ray flux to a mass estimate, for instance. The measured masses suggest that the baryon fraction is only about 10% – 20% of the total mass.

The family of the so-called Cold Dark Matter scenarios (CDM henceforth) is nowadays the preferred model for the nature of dark matter. In these scenarios the dark matter consists of non-baryonic, not relativistic (hence cold) particles that have a small velocity dispersion when matter is decoupling from radiation. So the pressure of the particles can be neglected and the evolution of the density perturbation can be described completely by gravitational forces. In contrast, an alternative scenario, the so-called Hot Dark Matter models, consist of relativistic particles with a high velocity dispersion at the time of decoupling. As a consequence, free streaming of the relativistic particles wipe out all structure on small scales. Structures in a HDM model form in a top-bottom way, so that there are first only very large bound objects which later fragment into smaller objects. In the CDM case structures first form on very small scales and later merge to

---

<sup>‡</sup>The usual way to express this is to quote the ratio of the mass to the emitted light, the so-called *mass to light ratio*. Typical values are  $M/L = 30M_{\odot}/L_{\odot}$  for galaxies and  $M/L = 300M_{\odot}/L_{\odot}$  for galaxy clusters, respectively.

larger objects. The CDM scenarios have been extensively tested in numerical  $N$ -body simulations and are currently favoured over the HDM scenarios. Perhaps the CDM models are not the complete truth and reality lies somewhere in between, in the *Mixed Dark Matter* models.

To describe the evolution of density fluctuations in CDM models we introduce the *density contrast*  $\delta(\vec{x})$ :

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \quad (24)$$

which is defined as the deviation of the density  $\rho(\vec{x}, t)$  from the average cosmic density  $\bar{\rho}(t)$ . The temporal and spatial evolution of the matter density  $\rho(\vec{x}, t)$ , and thus of the density contrast, is described by the standard Newton equations for the evolution of density and velocity of a pressureless fluid under the influence of a gravitational field. Although these equations are non-relativistic, they give the correct picture for the evolution of  $\rho$ . With  $p = 0$  the equations are:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) &= 0 && \text{(continuity equation),} \\ \frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} &= -\nabla \phi && \text{(Euler equation),} \\ \Delta \phi &= 4\pi G \rho && \text{(Poisson equation).} \end{aligned} \quad (25)$$

The solution of this system is not trivial in general. For very early times with small density fluctuations  $\delta \ll 1$  we can linearise the above equations after rewriting them in terms of the density contrast, comoving distances  $\vec{x} = \vec{r}/R$  and peculiar velocities  $\vec{u} = \vec{v} - \dot{R}\vec{x}$  to obtain

$$\frac{\partial^2 \delta}{\partial t^2} + \frac{2\dot{R}}{R} \frac{\partial \delta}{\partial t} - \frac{3H_0^2 \Omega_0}{2R^3} \delta = 0. \quad (26)$$

The general solution of eq. (26) is given by

$$\delta(\vec{x}, t) = D_+(t) \Delta_+(\vec{x}) + D_-(t) \Delta_-(\vec{x}), \quad (27)$$

where  $D_+(t)$  increases while  $D_-(t)$  decreases with time. We assume that at present times only the *growing mode*  $D_+(t) \Delta_+(\vec{x})$  is present. In the case of the *linear evolution* matter inhomogeneities of different scales from the initial field simply scale with time according to  $D_+(t)$ , and perturbations on different scales do not mix with each other. As we describe the initial spatial fluctuations as a Gaussian random field, the perturbations are completely characterised by the mean of the density contrast ( $\langle \delta \rangle = 0$  by definition) and the two-point correlation function  $\langle \delta(\vec{x}) \delta^*(\vec{y}) \rangle$ . This is valid as long as the evolution can be described by the linearised form eq. (26). To describe the correlation function  $\langle \delta(\vec{x}) \delta^*(\vec{y}) \rangle$ , we usually consider the Fourier transform of the density contrast:

$$\hat{\delta}(\vec{k}, t) = \int d^3x \delta(\vec{x}, t) e^{i\vec{k}\vec{x}}. \quad (28)$$

Considering the Fourier transform of  $\langle \delta(\vec{x}) \delta^*(\vec{y}) \rangle$  leads to

$$\langle \delta(\vec{x}) \delta^*(\vec{y}) \rangle \rightarrow \langle \hat{\delta}(\vec{k}) \hat{\delta}^*(\vec{k}') \rangle = (2\pi)^3 \delta_{\text{D}}(\vec{k} - \vec{k}') P_{\delta}(|\vec{k}|) \quad (29)$$

which defines the *power spectrum*  $P_{\delta}(k)$ .  $\delta_{\text{D}}$  denotes the Dirac delta function. For the last equation we used that  $\delta$  is a homogenous, isotropic field. This power spectrum is one of the fundamental quantities describing structure formation. For the form of the linear evolved power spectrum, an often used fitting formula is from Bond & Efstathiou 1984 (see [9])

$$P(|\vec{k}|) = \frac{A|\vec{k}|}{[1 + [aq + (bq)^{3/2} + (cq)^2]^{\nu}]^{2/\nu}}, \quad (30)$$

where  $q = \Gamma^{-1} k$  with the shape parameter  $\Gamma$ ,  $a = 6.4 h^{-1} \text{ Mpc}$ ,  $b = 3 h^{-1} \text{ Mpc}$ ,  $c = 1.7 h^{-1} \text{ Mpc}$ , and  $\nu = 1.13$ . The normalisation  $A$  of this spectrum cannot be calculated theoretically but it has to be determined from observations. This can be done on different scales:

$\Omega_0$	$\Omega_\Lambda$	$\Gamma$	$\sigma_8$	model name
1.0	0	0.25	0.6	EdS(0.6,0.25)
1.0	0	0.25	1.0	EdS(1,0.25)
1.0	0	0.5	0.6	EdS(0.6,0.5)
0.3	0	0.25	1.0	OCDM(1,0.25)
0.3	0.7	0.25	1.0	$\Lambda$ CDM(1,0.25)

Table 1: Cosmological models often used for theoretical predictions

- We can use the temperature fluctuations of the cosmic microwave background. The COBE satellite measured these fluctuations on scales of  $\geq 7^\circ$ , and so these measurements fix the amplitude on these large scales.
- If we assume that the observed light is tracing mass we can use the local variance of galaxy counts to determine the normalisation. One often uses the variance  $\sigma_8$  of galaxies in a radius of 8 Mpc. The original measurements for  $\sigma_8$  were close to unity.
- With the same technique, the variance of the galaxy cluster counts can also be used to normalise the power spectrum.

On scales of galaxies and galaxy clusters, the density contrast has grown, and the evolution there can no longer be described with linear equations. Perturbations of different scales interact, and the density field becomes non-Gaussian. On these small scales, the higher-order moments of the density contrast are non-zero. Although the field can no longer be described completely by the power spectrum, it is still a well defined quantity. A fitting formula for the form of the power spectrum in this case is given in Peacock & Dodds 1996 (see [51]). Different variants of the CDM scenario are theoretically characterised by different combinations of the parameters  $\Omega_0$ ,  $\Omega_\Lambda$ , the shape parameter  $\Gamma$  and the normalisation  $\sigma_8$ . Table 1 gives an overview of cosmological models under consideration nowadays.

Current research focuses on the question in which of these model universes we live in. As already indicated above, these investigations are done on two different scales. On the one hand, the cosmic microwave background measures temperature fluctuations, and hence density fluctuations, on large scales and hence in the linear regime. Most of the studies, however, focus on the small, non-linear scales, as single galaxies or galaxy clusters can easily be observed in a large part of the universe.

A very sensitive test to the nature of the universe is the determination of the so-called *mass function*. This function gives the number of objects with mass  $M$  at redshift  $z$ . In Fig. 2 the number of expected massive objects for different redshifts and the five cosmological models of Table 1 is plotted. From this plot we see that the number of massive objects at high redshift discriminates strongly between high- and low-density universes. Most of the efforts underway to determine this mass function rely on light at some wavelength as a tracer of mass (like the temperature-mass relation in  $X$ -ray studies of clusters or a color sequence of early-type galaxies for optical surveys). The shortcoming of this approach is that we may miss an unknown class of objects, completely dominated by dark matter. It is obvious that the existence of such objects may substantially change our current understanding about the universe. To find these objects, or rule out their existence, we need a cosmological tool that is sensitive directly to the mass of compact objects and does not need further assumptions about how light traces mass. With gravitational lensing we introduce such a tool in the next subsection.

### 3.3 Gravitational lensing

Gravitational lensing describes the effect that any matter deflects electromagnetic radiation. If light, coming from a distant object, is lensed on its way to us by an intervening mass concentration,

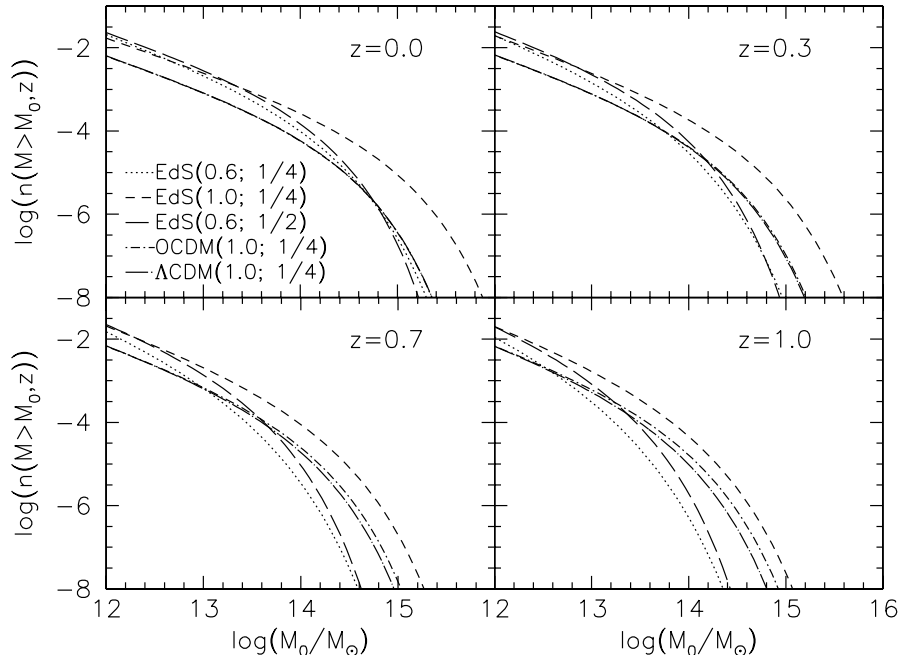


Figure 2: The number density of massive objects (number of objects with mass  $M \geq M_0$  per  $\text{Mpc}^{-3}$  comoving volume). We see that this number provides a powerful discriminator for cosmological models if we can determine the mass function at high redshift ( $z \geq 0.7$ ). For low redshift most models are degenerate in the mass function. The calculations were done with the semi-analytical Press-Schechter formalism.

the apparent position and also the observed shape of the object is different from what it would be without the lens. We will describe in the next sections how this lens effect can be utilised to investigate the mass distribution lying between us and distant objects emitting light. We will see that the lens effect can be used to study massive, non-linear objects on the one hand but also the large-scale density contrast  $\delta$ .

### 3.3.1 Light propagation in inhomogeneous universes

To describe light propagation in our inhomogeneous universe correctly we have to take into account that light travels on geodesics of the homogenous background cosmology as well as its deflections due to mass inhomogeneities. To derive an equation of motion for a light ray moving through the universe we consider the situation in Fig. 3. There we consider two light rays, starting at our position. Let one of the light rays be influenced by the background cosmology only, the second one by small-scale mass inhomogeneities in addition. Considering this constellation we can derive an equation of motion for the comoving transverse separation  $\vec{x}(w)$  as a function of the comoving distance from the observer  $w$ , the space curvature and the Newtonian potential  $\Phi(\vec{x}(w), w)$  associated with the mass concentrations. If not stated otherwise, quantities with an arrow, like  $\vec{x}$ , denote a two-dimensional vector from now on. We can derive for  $\vec{x}(w)$

$$\frac{d^2 \vec{x}}{dw^2} + K \vec{x} = -\frac{2}{c^2} \left[ \nabla_{\perp} \Phi \left( \vec{x}(\vec{\theta}, w), w \right) \right], \quad (31)$$

where  $\nabla_{\perp} = (\partial x_1, \partial x_2)$  is the transverse gradient operator in comoving coordinates. Although the final form of this equation is that of a simple harmonic oscillator with an inhomogeneity term, its rigorous derivation is very cumbersome. It can be found in Bartelmann & Schneider (2000)



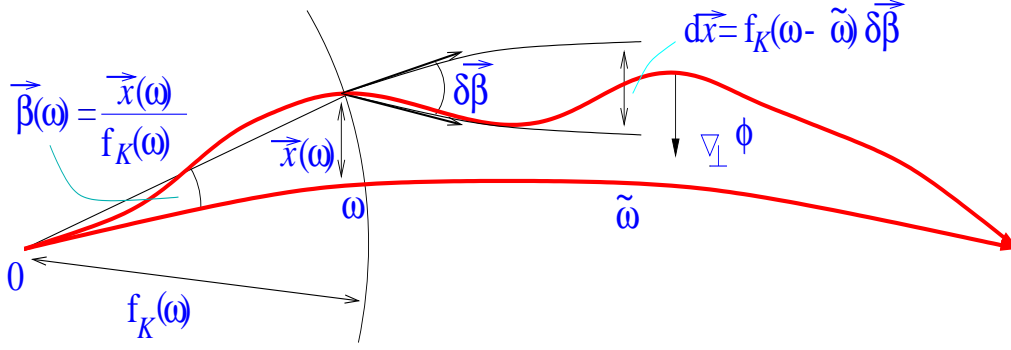


Figure 3: The light path in an inhomogeneous universe (Figure 1.6 from [13]): Two light rays travelling through our universe are considered. The lower one moves on geodesics of the isotropic and homogenous background cosmology. The upper one is deflected by mass inhomogeneities in addition. We can derive a relation for the transverse separation  $\vec{x}(w)$  as a function of the comoving distance  $w$  from the observer. The mass inhomogeneities are described by a Newtonian potential  $\Phi$  whose force is acting transversely (expressed by the  $\nabla_{\perp}$  operator) to the lower light ray. For more details see the text.

(see [4]) for instance. The homogeneous part of this equation

$$\frac{d^2 \vec{x}}{dw^2} + K \vec{x} = 0 \quad (32)$$

describes light propagation in the homogenous and isotropic Robertson-Walker universe. The solution of eq. (31) is formally given by

$$\vec{x}(\vec{\theta}, w) = f_K(w) \vec{\theta} - \frac{2}{c^2} \int_0^w dw' f_K(w - w') \left[ \nabla_{\perp} \Phi \left( \vec{x}(\vec{\theta}, w'), w' \right) \right]. \quad (33)$$

To fix the solution of the homogenous equation we have used  $\vec{x}(0) = 0$  and  $\frac{d\vec{x}}{dw}(0) = \vec{\theta}$  (the light rays start at our position with the angle  $\vec{\theta}$  between them). As the transverse separation  $\vec{x}(\vec{\theta}, w)$  also appears in the Newton potential on the right-hand-side, this is still an integral equation. We first use this equation to study isolated, massive objects like galaxy clusters and later to investigate the cosmological density contrast  $\delta$ .

### 3.3.2 The lensing effect of isolated mass concentrations

To study the lensing effect of isolated massive objects we make the following assumptions:

- All of the lensing effect is caused by the massive object that is localised at a comoving distance  $w_d$ . We neglect the lensing effect of large-scale structures. Thus, the Newtonian potential in this case only contains a contribution from the mass concentration.
- To arrive at an explicit equation for  $\vec{x}(\vec{\theta}, w)$  we use that the angles by which the light rays are deflected and hence the Newtonian potential are very small<sup>§</sup>. We will show this fact below. So we can expand eq. (33) in powers of the Newtonian potential. This is done formally by inserting eq. (33) into itself

$$\vec{x}(\vec{\theta}, w) = f_K(w) \vec{\theta} - \frac{2}{c^2} \int_0^w dw' f_K(w - w') \nabla_{\perp} \Phi \left( f_K(w) \vec{\theta} - \frac{2}{c^2} \int_0^w \dots, w' \right). \quad (34)$$

Keeping only the first term yields:

$$\vec{x}(\vec{\theta}, w) = f_K(w) \vec{\theta} - \frac{2}{c^2} \int_0^w dw' f_K(w - w') \left[ \nabla_{\perp} \Phi \left( f_K(w') \vec{\theta}, w' \right) \right]. \quad (35)$$

<sup>§</sup>We mean  $|\Phi| \ll c^2$  if we say that the Newtonian potential is small.

- We assume that the (comoving) distances between observer-lens and lens-source are much larger than the spatial extent of the mass concentration. This is known as *thin screen approximation* and the Newtonian potential is assumed to be different from zero only in a very narrow region around  $w_d$ . As a consequence the integrand in eq. (35) only contributes in an interval  $w_d - \epsilon, w_d + \epsilon$ , and we can simplify it to

$$\vec{x}(\vec{\theta}, w) = f_K(w)\vec{\theta} - \frac{2}{c^2}f_K(w - w_d) \int_{w_d - \epsilon}^{w_d + \epsilon} dw' \left[ \nabla_{\perp} \Phi \left( f_K(w')\vec{\theta}, w' \right) \right]. \quad (36)$$

We now convert the derivatives of the Newtonian potential and the integrand itself from comoving  $w$  to physical, proper coordinates  $r := D_{\text{prop}}$  by:

$$\nabla_{\perp} = R\nabla_{\perp}^r \quad dr = R dw. \quad (37)$$

We further use that the integrand itself is confined to a narrow band in the  $r$  direction and we can extend the integration limits to infinity. This leads to

$$\int_{w_d - \epsilon}^{w_d + \epsilon} dw' \left[ \nabla_{\perp} \Phi \left( f_K(w')\vec{\theta}, w' \right) \right] \approx \int_{-\infty}^{\infty} dr \left[ \nabla_{\perp}^r \Phi \left( f_K(w')\vec{\theta}, r(w) \right) \right]. \quad (38)$$

Finally, we rewrite eq. (36) in terms of the angle  $\vec{\beta} := \vec{x}/f_K(w)$ , that is the angle under which the observer at  $w = 0$  sees the light ray influenced by the background cosmology and the mass concentration, at the end of its journey, relative to the angle  $\vec{\theta}$  of the light ray influenced by the Robertson-Walker cosmology only. We obtain

$$\vec{\beta} = \vec{\theta} - \frac{f_K(w - w_d)}{f_K(w)} \int_{-\infty}^{\infty} dr \left[ \nabla_{\perp}^r \Phi \left( f_K(w')\vec{\theta}, r(w) \right) \right]. \quad (39)$$

We now can interpret the integral in the last term as the deflection angle  $\hat{\alpha}$  caused by the mass. The ratio  $f_K(w - w_d)/f_K(w)$  is just the ratio of the angular diameter distances between mass concentration-source and observer-source, respectively. Writing the angular diameter distance between  $w_d$  and  $w$  as  $D_{\text{ds}}$  and that between 0 and  $w_s$  as  $D_s$ , we get the following form of the so-called *lens equation* eq. (39)

$$\vec{\beta} = \vec{\theta} - \frac{D_{\text{ds}}}{D_s} \hat{\alpha}(\vec{\theta}). \quad (40)$$

We now give a second, heuristic derivation of it as it is done most often in the literature. This is done mainly to introduce terms that we need later on.

Figure 4 visualises the so-called *lens geometry*, a picture often used to derive the lens equation. There we assume a two-dimensional mass distribution at a distance  $D_d$  from an observer<sup>¶</sup>. Let us call this plane the *lens plane*. Parallel to this plane at a larger distance  $D_s$  we define the *source plane*. Furthermore, an *optical axis* going perpendicular through these planes and crossing the observer is introduced. On each plane we define a coordinate system having its origin at the intersection with the optical axis. A light emitting object be at coordinates  $\vec{\eta}$  in the source plane (the *source*). This light crosses the lens plane at  $\xi$ , is bent by the mass concentration and reaches the observer. If we use the thin-screen approximation so that we can approximate the light path by two straight lines from source plane to lens plane and from lens plane to observer, we can read from Fig. 4 the following relations when interpreting the distances  $D_d$ ,  $D_s$  and  $D_{\text{ds}}$  as angular diameter distances:

$$\frac{\vec{\eta}'}{\vec{\xi}'} = \frac{D_s}{D_d}, \quad \vec{\eta}' + \hat{\alpha}(\vec{\xi}')D_{\text{ds}} = \vec{\eta}. \quad (41)$$

<sup>¶</sup>The assumption of a two-dimensional mass distribution corresponds to the thin-screen approximation above. It implies that we can deal with a three-dimensional mass distribution  $\rho(x, y, z)$  through its two-dimensional projection

$\Sigma(x, y) = \int_{-\infty}^{\infty} \rho(x, y, z) dz$  in the lens plane if  $x$  and  $y$  are the physical coordinates in this plane.

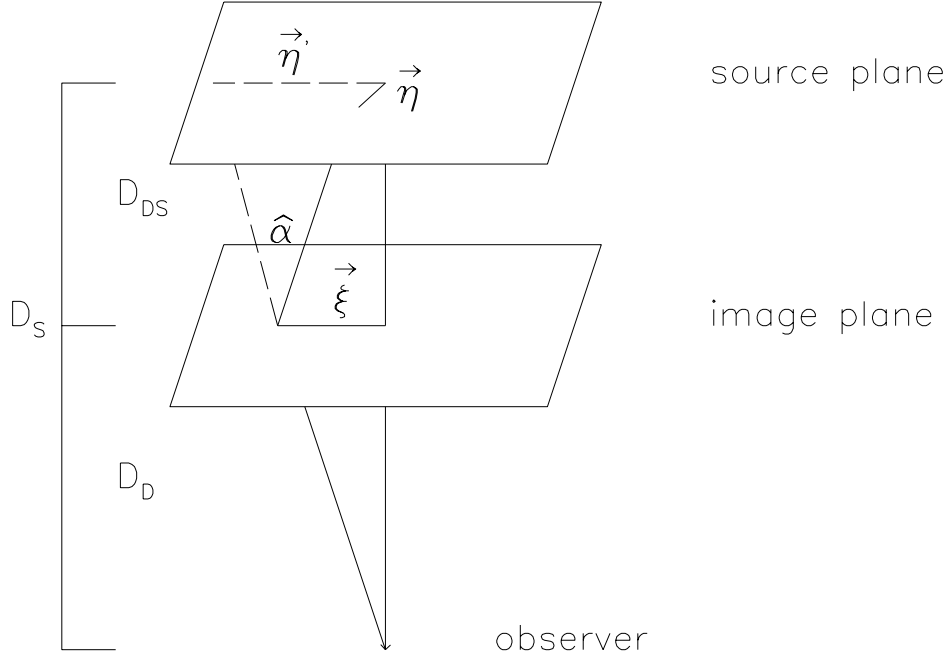


Figure 4: The lens geometry

A combination of these equations leads to:

$$\vec{\eta} = \frac{D_s}{D_d} \vec{\xi} - D_{ds} \hat{\alpha}(\vec{\xi}). \quad (42)$$

This equation connects the source plane coordinate  $\vec{\eta}$  on which we would see the object in the absence of a lens with the coordinate  $\vec{\xi}$  under which we actually see the image (for this the lens plane is also called the *image plane*). Changing the distances  $\vec{\eta}$  and  $\vec{\xi}$  to angles between light rays and optical axis:

$$\vec{\theta} := \frac{\vec{\xi}}{D_d} \quad \vec{\beta} := \frac{\vec{\eta}}{D_s} \quad (43)$$

we again arrive at the lens equation eq. (40) connecting the angles  $\vec{\beta}$  and  $\vec{\theta}$ .

We now consider light deflection by a point mass  $M$  as an example. The mass is put at the origin of the lens plane and its Newtonian potential is given by:

$$\Phi(\vec{\xi}, d) = -\frac{GM}{\sqrt{|\vec{\xi}|^2 + d^2}} \quad (44)$$

where we understand  $d$  as the proper distance along the optical axis. The integration of eq. (38) leads then to:

$$\hat{\alpha} = \frac{4GM\vec{\xi}}{c^2|\vec{\xi}|^2}. \quad (45)$$

The quantity  $R_s = \frac{2GM}{c^2}$  is called the Schwarzschild radius of an object with mass  $M$ , and for all cosmological objects that we consider in this thesis it is much smaller than the *impact parameter*  $|\vec{\xi}|$ . For instance for the Sun, the Schwarzschild radius is about 3km, while light that is deflected by it has an impact parameter at least as large as its radius ( $7 \cdot 10^5$ km). Thus, the deflection angles caused by cosmological objects are indeed very small. This fact justifies the third assumption that we made in the derivation of eq. (36). If we do not have a single point mass but a surface mass density  $\Sigma(\vec{\xi})$  in the lens plane, the deflection angle becomes:

$$\hat{\alpha}(\vec{\theta}) = \frac{4G}{c^2} \int_{\mathbb{R}^2} d^2\vartheta \Sigma(\vec{\vartheta}) \frac{\vec{\theta} - \vec{\vartheta}}{|\vec{\theta} - \vec{\vartheta}|^2}. \quad (46)$$

For the following it is useful to rescale the surface mass density  $\Sigma(\vec{\xi})$  and the deflection angle  $\hat{\alpha}$  to dimension-less quantities by:

$$\kappa := \frac{\Sigma}{\Sigma_{\text{cr}}}; \quad \vec{\alpha}(\vec{\theta}) := \frac{D_{\text{ds}}}{D_s} \hat{\alpha}(\vec{\theta}), \quad (47)$$

where  $\Sigma_{\text{cr}} := \frac{c^2 D_s}{4\pi G D_a D_{\text{ds}}}$  is the so-called *critical density*. With these definitions the lens equation obtains the final, simple form,

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}), \quad (48)$$

where  $\vec{\alpha}(\vec{\theta})$  is now

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\vartheta \kappa(\vec{\vartheta}) \frac{\vec{\theta} - \vec{\vartheta}}{|\vec{\theta} - \vec{\vartheta}|^2}. \quad (49)$$

We also introduce the *lens potential*

$$\psi(\vec{\theta}) := \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\vartheta \kappa(\vec{\vartheta}) \ln |\vec{\theta} - \vec{\vartheta}| \quad (50)$$

that has the following properties:

$$\nabla\psi(\vec{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\vartheta \kappa(\vec{\vartheta}) \frac{\vec{\theta} - \vec{\vartheta}}{|\vec{\theta} - \vec{\vartheta}|^2} = \vec{\alpha}(\vec{\theta}), \quad (51)$$

$$\Delta\psi(\vec{\theta}) = 2 \int_{\mathbb{R}^2} d^2\vartheta \kappa(\vec{\vartheta}) \delta_D(\vec{\theta} - \vec{\vartheta}) = 2\kappa(\vec{\theta}). \quad (52)$$

**Properties of the lens mapping** We mentioned in the preamble of section 3.3 that gravitational lensing not only changes the apparent positions of sources but also their shapes. This statement will be quantified in this section. Let us consider the Jacobi-Matrix  $A_{ij}$  of the lens mapping

$$A_{ij} := \left( \frac{\partial\beta_i}{\partial\theta_j} \right) = \delta_{ij} - \underbrace{\alpha_{i,j}}_{\psi_{,ij}}. \quad (53)$$

As  $\psi_{,ij} = \psi_{,ji}$ ,  $A_{ij}$  is a symmetric matrix that can be written as

$$A = \begin{pmatrix} 1 - \kappa + \gamma_1 & +\gamma_2 \\ +\gamma_2 & 1 - \kappa - \gamma_1 \end{pmatrix}, \quad (54)$$

with the abbreviations

$$\gamma_1(\vec{\theta}) := -\frac{1}{2}(\psi_{,11} - \psi_{,22}) = -\frac{1}{2} \left( \frac{\partial\beta_1}{\partial\theta_1} - \frac{\partial\beta_2}{\partial\theta_2} \right) \quad \gamma_2(\vec{\theta}) := -\psi_{,12} = -\frac{\partial\beta_1}{\partial\theta_2}. \quad (55)$$

A symmetric matrix can always be diagonalised, and  $A$  reads in its eigenvalue system

$$A = \begin{pmatrix} 1 - \kappa & 0 \\ 0 & 1 - \kappa \end{pmatrix} + \begin{pmatrix} |\gamma| & 0 \\ 0 & -|\gamma| \end{pmatrix}, \quad (56)$$

where  $|\gamma| := \sqrt{\gamma_1^2 + \gamma_2^2}$ . The lens mapping describes a mapping from the lens plane to the source plane. As we can observe only the images, we are interested in the inverse mapping. Assuming that an object subtends a sufficiently small area so that the lens mapping does not change on these scales, the lens effect on the source can be approximated by the linear equation

$$\vec{\theta} = A^{-1}\vec{\beta} \sim \begin{pmatrix} \frac{1}{1-\kappa+|\gamma|} & 0 \\ 0 & \frac{1}{1-\kappa-|\gamma|} \end{pmatrix} \vec{\beta}. \quad (57)$$

We now see the meaning of  $\kappa$  and  $\gamma$ . In the absence of the shear  $\gamma$  the object is stretched isotropically by a factor  $1/(1 - \kappa)$ . If  $\kappa = 0$  and  $|\gamma| \neq 0$  the image is stretched along the one and compressed along the other eigenvector. In the general case, a circular source with radius  $a_0$  is transformed to an ellipse with semi-minor/major axis along the eigenvectors and moduli

$$a_{\pm} = \frac{a_0}{|1 - \kappa \pm |\gamma||}. \quad (58)$$

Besides this change in object shape, gravitational lensing also changes the observed flux from a source. As it conserves the number of photons (light is only bent but neither absorbed nor emitted) the *surface brightness*  $I(\vec{\theta})$  (defined as the number of photons that reach us per time interval and per solid angle) is conserved as well. A rigorous derivation of this can be found for instance in Schneider et al. 1992 (see [60]). As gravitational lensing changes the solid angle in which we see a source, the total observed flux, that is the surface density integrated over the extent of the object, is changed as well. The influence of lensing on the solid angle is given by the Jacobi determinant. We obtain

$$\frac{\Omega_l}{\Omega_s} = \frac{F_l}{F_s} = \frac{1}{\det A} = \frac{1}{|(1 - \kappa)^2 + |\gamma|^2|} \quad (59)$$

for the ratio of observed flux  $F_l$  and intrinsic flux  $F_s$ . The curves on which  $\det A = 0$  are called critical curves<sup>||</sup>. If they are mapped to the source plane we call them caustics. On these curves images are formally infinitely magnified. This infinite magnification can never be observed as we would have to take into account wave effects of the light, neglected here. Instead, *giant luminous arcs* and multiple images are observed in this case. As the *strong lensing* effect that deals with critical lenses is not the subject of this thesis we do not go further into detail. Figure 5 shows some of its consequences nevertheless. We now consider axially symmetric lenses as an example of the formalism.

**Axially symmetric lenses** We now want to specialise the formulae given in the previous section for spherical mass distributions  $\rho(x, y, z) = \rho(\sqrt{x^2 + y^2 + z^2})$  and so cylindrically symmetric surface mass densities  $\kappa(\vec{\theta}) = \kappa(|\vec{\theta}|)$ . Mass distributions of extragalactic objects and dark matter haloes are most often modelled by symmetric models so that it is of interest to study their properties. Introducing polar coordinates  $(\theta, \phi)$  in the lens plane and limiting our considerations to points on  $(\theta, 0)$ <sup>\*\*</sup>, we get for the components of the deflection angle:

$$\alpha_1(\theta) = \frac{1}{\pi} \int_0^\infty \vartheta d\vartheta \kappa(\vartheta) \int_0^{2\pi} d\varphi \frac{\theta - \vartheta \cos(\varphi)}{(\theta)^2 + (\vartheta)^2 - 2\theta\vartheta \cos(\varphi)}, \quad (60)$$

$$\alpha_2(\theta) = \frac{1}{\pi} \int_0^\infty \vartheta d\vartheta \kappa(\vartheta) \int_0^{2\pi} d\varphi \frac{-\vartheta \sin(\varphi)}{(\theta)^2 + (\vartheta)^2 - 2\theta\vartheta \cos(\varphi)}. \quad (61)$$

<sup>||</sup>This is why  $\Sigma_{cr}$  is called critical density. If a lens has  $\kappa = 1$  somewhere it has a critical curve.

<sup>\*\*</sup>for all other points we get the quantities by a simple rotation of the coordinate system

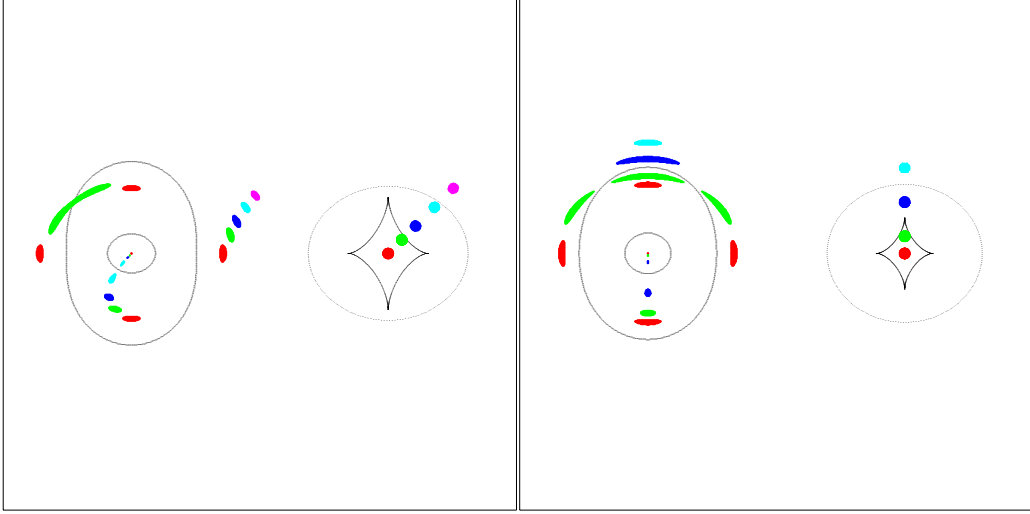


Figure 5: Critical lenses (Figure 19 from [47]): We show several constellations of sources and images of a lens with an elliptical mass profile. The left part of the panels shows the configuration in the lens, the right part that in the source plane. The closed curves show critical lines and caustics, respectively. A source lying outside the caustics produces exactly one image. When crossing a caustic, two new images appear. If the source lies exactly on a caustic we observe so-called *giant arcs*. A transparent lens always produces an odd number of images (see [60]). This does not hold when we consider singular mass profiles.

The second component is identically zero. For the first one we obtain

$$\alpha_1(\theta) = \frac{2}{\theta} \int_0^\theta \vartheta d\vartheta \kappa(\vartheta) \equiv \frac{m(\theta)}{\theta}, \quad (62)$$

where we have defined the dimension-less mass  $m(\theta)$  within  $\theta$ . Because of eq. (48), optical axis, source and image are all lying in one plane for axisymmetric lenses. The distortion matrix reads in this case

$$A = \begin{pmatrix} 1 + \frac{m(\theta)}{(\theta)^2} - 2\kappa & 0 \\ 0 & 1 - \frac{m(\theta)}{(\theta)^2} \end{pmatrix} \quad (63)$$

$$= \begin{pmatrix} 1 - \kappa & 0 \\ 0 & 1 - \kappa \end{pmatrix} + \begin{pmatrix} \frac{m(\theta)}{(\theta)^2} - \kappa & 0 \\ 0 & -\frac{m(\theta)}{(\theta)^2} + \kappa \end{pmatrix}. \quad (64)$$

We see first that  $A$  is in its eigenvalue system. Thus, a source is always distorted along the  $\theta$  direction or perpendicular to it. Quantifying the distortion with the semi-major and semi-minor axis of an ellipse introduced in eq. (58) we say, the image is tangentially aligned if the semi-major axis is perpendicular to  $\theta$ . The components of the shear read

$$\gamma_1 = \frac{m(\theta)}{(\theta)^2} - \kappa; \quad \gamma_2 = 0. \quad (65)$$

The simplest and most often used example of circular models is the family of *power law* models

$$\kappa(\theta) = a \left( \frac{\theta}{\theta_E} \right)^{-q}. \quad (66)$$

They are here characterised by a strength  $a$ , the slope  $q$  and the *Einstein radius*  $\theta_E^{\dagger\dagger}$ . The Einstein radius is a characteristic scale in lens theory and defined such that the mean surface mass density within  $\theta_E$ ,  $\bar{\kappa}(\theta_E) := \frac{m(\theta_E)}{\theta_E^2}$ , is unity. It approximately limits the region within which strong lensing features (giant arcs, multiple images) occur. The models are unphysical as the density reaches infinity at the origin. Moreover, the total mass can become infinite. Hence these models are used primarily when studying the mass distribution outside the Einstein radius. It will be used for this purpose in section 5. The shear modulus of the model is

$$|\gamma(\theta)| = \left(1 - \frac{2a}{2-q}\right) \left(\frac{\theta}{\theta_E}\right)^{-2} + a \left(\frac{\theta}{\theta_E}\right)^{-q} \left(\frac{q}{2-q}\right). \quad (67)$$

For the special case  $a = 0.5$ ,  $q = 1$  we get the *singular isothermal sphere*<sup>‡‡</sup>:

$$\kappa(\theta)_{SIS} = |\gamma(\theta)|_{SIS} = 0.5 \left(\frac{\theta_E}{\theta}\right). \quad (68)$$

### 3.3.3 The lensing effect of large-scale structures

The discussion above showed that the lensing effect can not probe the full three-dimensional structure of massive objects but only their projections along our line-of-sight. So, if we want to study the large-scale density contrast  $\delta$ , we expect that only a two-dimensional projection of this quantity can be probed. In contrast to isolated massive objects, large-scale structures are present all along the line-of-sight and not only on a single lens plane. Nevertheless we are able to define a surface mass density  $\kappa$  for the density contrast between us and a distance  $w$  so that this case can be treated in a similar way as the single lens plane case. We return to eq. (33) and proceed in the following way:

1. Also here we make the assumption that the Newtonian potential produced by the large-scale structures is small and that we are allowed to make a series expansion as a function of it. Keeping only the first term leads us to eq. (35) as above.
2. As in the single lens plane, a distortion matrix can be defined

$$A(\vec{\theta}, w) = \frac{\partial \vec{\beta}}{\partial \vec{\theta}}, \quad (69)$$

where we used that  $\vec{\beta} = \vec{x}/f_K(w)$  would be the separation angle of the two light rays in the absence of any mass. Differentiating eq. (35) yields

$$A_{ij}(\vec{\theta}, w) = \delta_{ij} - \frac{2}{c^2} \int_0^w dw' \frac{f_K(w-w') f_K(w')}{f_K(w)} \Phi_{,ik}(\vec{x}(\vec{\theta}, w'), w') A_{kj}(\vec{\theta}, w'). \quad (70)$$

3. The last equation is an integral equation similar in form to eq. (33) which was simplified by making a series expansion in the Newtonian potential. We use the same strategy here and make a series expansion of the distortion matrix  $A$  that is terminated after the first term, leading to

$$A_{ij}(\vec{\theta}, w) = \delta_{ij} - \frac{2}{c^2} \int_0^w dw' \frac{f_K(w-w') f_K(w')}{f_K(w)} \Phi_{,ij}(f_K(w')\vec{\theta}, w'). \quad (71)$$

When we imagine large-scale structures as a pile of individual lens planes along the line-of-sight, the total amplification matrix of such a system would be given as a sum over products

<sup>††</sup>The dimensionless strength  $a$  and the Einstein radius  $\theta_E$  depend on the velocity dispersion of the object and on the distance ratio  $D_{ds}/D_s$ .

<sup>‡‡</sup>This model reads in physical parameters  $\kappa(\theta) = 2\pi \cdot (\sigma/c)^2 \cdot D_{ds}/(D_s\theta)$ , where  $\sigma$  is the velocity dispersion of the galaxy. Hence  $\theta_E = 4\pi \cdot (\sigma/c)^2 \cdot D_{ds}/D_s$ .

of amplification matrices. This total amplification matrix would contain *coupling terms* between the different lenses, i.e. products of the shears and surface mass densities. The approximation that we made here means that we neglect all the quadratic and higher-order coupling terms.

4. We note that the resulting matrix in eq. (71) is also symmetric like that from eq. (53) for the single lens plane. So, in analogy to this case, a lensing potential  $\psi$  can be defined

$$\psi(\vec{\theta}, w) = \frac{2}{c^2} \int_0^w dw' \frac{f_K(w-w')}{f_K(w')} \frac{f_K(w')}{f_K(w)} \Phi(f_K(w')\vec{\theta}, w'). \quad (72)$$

5. With this lensing potential we are now able to define formally a surface mass density  $\kappa$  and a shear  $\gamma$  for the large-scale structures case

$$\kappa(\vec{\theta}, w) = \frac{1}{2}(\psi_{,11} + \psi_{,22}), \quad (73)$$

$$\gamma_1(\vec{\theta}, w) = -\frac{1}{2}(\psi_{,11} - \psi_{,22}); \quad \gamma_2(\vec{\theta}, w) = -\psi_{,12} \quad (74)$$

As a final step we now relate  $\kappa$  to the density contrast  $\delta$ . One can show (see [52]) that the density contrast obeys a kind of Poisson equation,

$$\nabla^2 \Phi = \frac{3H_0^2 \Omega_d}{2R} \delta. \quad (75)$$

Furthermore, we use the fact that a term  $\Phi_{,33}$  cancels upon integration along  $w$ . So, we finally obtain from eq. (72), eq. (73) and eq. (75) the following expression for  $\kappa$

$$\kappa(\vec{\theta}, w) = \frac{3}{2} \left( \frac{H_0}{c} \right)^2 \Omega_d \int_0^w dw' \frac{f_K(w-w') f_K(w')}{f_K(w)} \frac{\delta(f_K(w')\vec{\theta}, w')}{R(w')}. \quad (76)$$

In this way we can treat the projected density contrast  $\delta$  completely in the same way as a single, projected mass concentration. A crucial assumption here was that we can neglect higher-order terms in the amplification matrix  $A_{ij}$ . This is equivalent to the assumption  $\kappa \ll 1$  and  $|\gamma| \ll 1$ . We will use it frequently in the next section where the *weak lensing* effect is discussed.



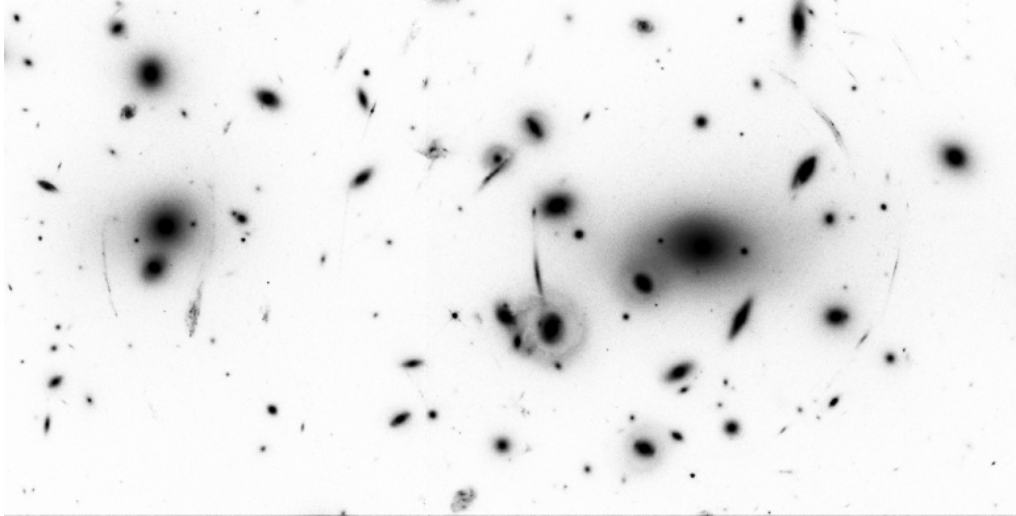


Figure 6: The lensing cluster Abell 2218 (NASA HST Archive)

## 4 Weak gravitational lensing

### Summary:

*In this section we introduce the concepts of weak gravitational lensing. The global goal is to investigate mass that acts as a gravitational lens on its background galaxies. First the quantities, characterising a gravitational lens (the surface mass density  $\kappa$  and the shear  $\gamma$ ), are connected to observables. Afterwards we introduce the basic ideas how one of the primary goals of weak lensing, reconstructing the lensing surface mass density, can be achieved.*

The term *weak gravitational lensing* is a very young one. It came up in about 1993 when Kaiser & Squires (see [33]) first showed how we can use the distortion (the shear  $\gamma$ ) that gravitational lensing causes on background images, to reconstruct the surface mass density  $\kappa$ . This idea is evident if we use the *strong lensing* regime of a massive galaxy cluster. Figure 6 shows an optical HST\* image of the central part of the rich cluster Abell 2218. We note giant arcs that immediately give strong constraints on the positions and shapes of critical curves of the lens. Also several multiple image configurations are produced by this cluster. A possible strategy here is to use a parametrised model for the cluster lens and adapt the parameters so that the critical curves and multiple image systems are reproduced. This kind of *lens modelling* has been very successful in the past years. For instance, an early model of the cluster Abell 370 predicted multiple image systems that were later confirmed when better observations became available (see [39] and [38]).

Lensing investigations in the outer parts of galaxy clusters, where the distortion effect on background galaxies is only small, are completely different. In this so-called *weak lensing regime* (it is formally defined by  $\kappa \ll 1$  and  $|\gamma| \ll 1$ ) the very weak shape distortions, caused by the lens, cannot be measured from individual galaxies. We will see below that we can instead measure statistically coherent distortion patterns when considering a large number of galaxies. In contrast to strong lensing, all these weak lensing investigations are therefore of statistical nature. Figure 7 visualises the strong and weak lensing regimes of a gravitational lens again. Before going into detail, we study how the quantities  $\kappa$  and  $\gamma$  are connected to observables.

---

\*Hubble Space Telescope

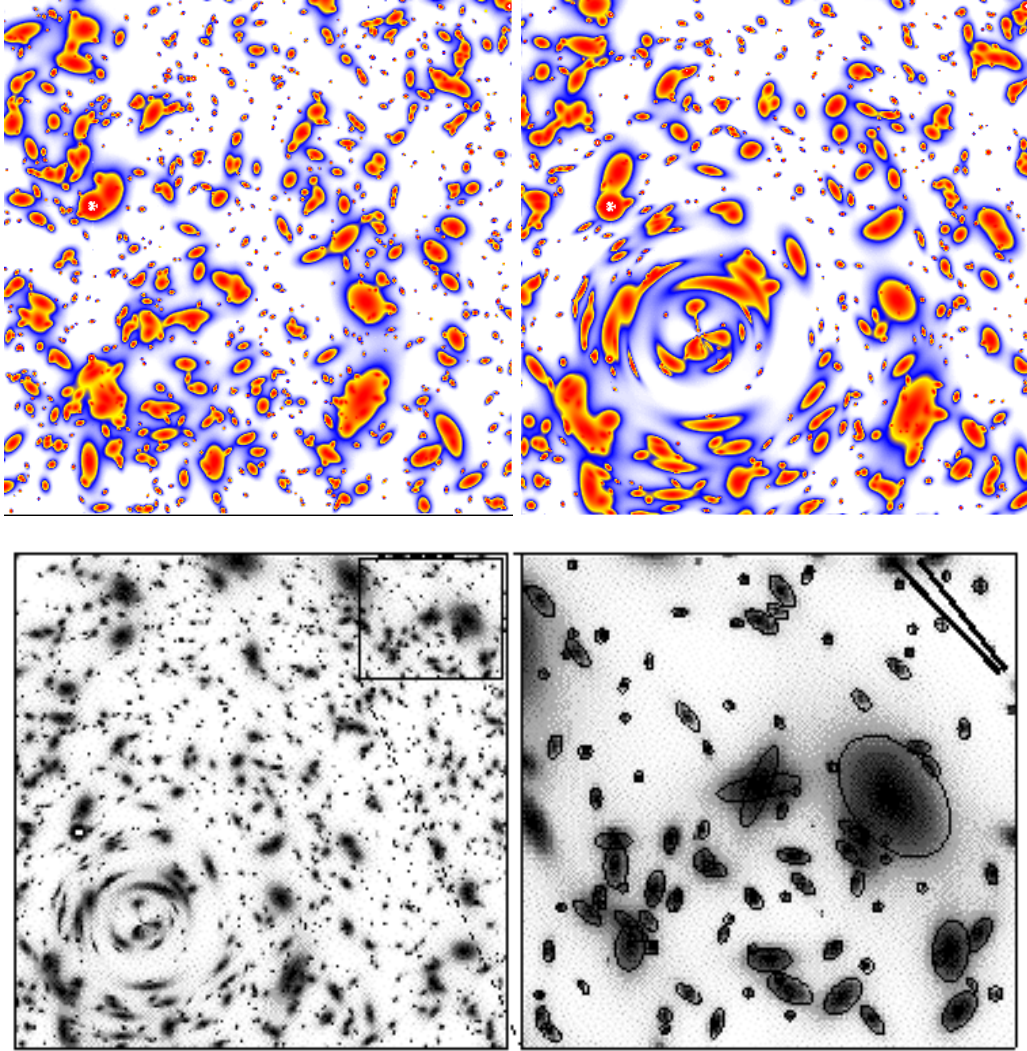


Figure 7: Visualisation of strong and weak lensing (Figure 2 from [46]): The upper left panel shows a *blank field* with a random distribution of galaxies. In the upper right panel we have inserted a gravitational lens in front of the galaxies. We see that near to the lens centre, the *background sources* are strongly distorted and aligned coherently around the lens centre like in the image of Abell 2218 (Fig. 6). In the lower panels a larger area around the lens is shown. Zooming into the part that has the largest distance from the lens centre we do not see an obvious distortion of individual galaxy images. But performing a statistical analysis of the object shapes we discover that they are preferentially aligned tangential to the mass centre (this is indicated by the sticks; the lower stick shows the direction of the shear caused by the cluster, the upper stick the measured mean shear from 92 galaxies in the zoomed area.)

#### 4.1 Lensing quantities and image ellipticities

Photometry gives us information about the light distribution of sources, and so it is used to parametrise their shape. We will use the surface brightness  $I(\vec{\theta})$  for this purpose. For the following we assume that we can uniquely identify the light belonging to a galaxy. Equivalent is the assumption that all objects are well separated from each other. First, we identify the position of

an object  $\vec{\theta}$  by its *centre of light*

$$\bar{\vec{\theta}} = \frac{\int d^2\theta I(\vec{\theta})\vec{\theta}}{\int d^2\theta I(\vec{\theta})}. \quad (77)$$

In analogy to this *first moment* of light we now consider the second moments

$$Q_{ij} := \frac{\int d^2\theta I(\vec{\theta})(\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int d^2\theta I(\vec{\theta})}; \quad 1 \leq i, j \leq 2. \quad (78)$$

With these second moments the complex *image ellipticities*

$$\chi := \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22}}. \quad (79)$$

are defined. For an object with elliptical isophotes and an axis ratio  $b/a = r \leq 1$ ,  $\chi$  reads

$$\chi = \frac{1 - r^2}{1 + r^2} e^{2i\vartheta}, \quad (80)$$

where  $\vartheta$  is the angle between the abscissa and the semi-major axis. In principle the numerator of eq. (79) already yields a measure for the deviation from a round object. It is divided by a rotational invariant of the second moment tensor so that it does not depend on the angular size of the object. But we are free to choose any combinations of invariants (trace and the determinant of the  $Q_{ij}$  tensor) for this normalisation. The following ellipticity  $\epsilon$

$$\epsilon := \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}, \quad (81)$$

that is

$$|\epsilon| = \frac{1 - r}{1 + r} \leq 1 \quad (82)$$

for an elliptical object will turn out more practical for the theoretical considerations later in this chapter. When we turn to ellipticity measurements in observational data in section 6 we will come back to the quantity  $\chi$ . We can express the observed image ellipticity  $\epsilon$  as a function of the intrinsic one  $\epsilon^s$  that we would observe in the absence of a lens. Because of the invariance of the surface brightness  $I(\vec{\theta})$  under gravitational lensing, we have the following relation between the intrinsic  $I^s(\vec{\beta})$  and the lensed  $I(\vec{\theta})$

$$I(\vec{\theta}) = I^s(\vec{\beta}(\vec{\theta})). \quad (83)$$

The extent of small, faint background sources is typically much smaller than the scales on which a lens changes its strength, so that we use the linearised lens equation  $\vec{\beta} = \vec{\beta}_0 + A(\vec{\theta}_0)(\vec{\theta} - \vec{\theta}_0)$  leading to the following relation between second moments in the lens and source plane

$$Q^s = AQA^T = AQA, \quad (84)$$

where  $A^T$  is the transponent of the distortion matrix that is equal to  $A$  itself. As a consequence, the transformations for the ellipticities of source and lens plane read

$$\chi^s = \frac{\chi + 2g + |g|^2\chi^*}{1 + |g|^2 + 2\text{Re}(g\chi^*)} \quad (85)$$

and

$$\epsilon^s = \begin{cases} \frac{\epsilon + g}{1 + g^*\epsilon} & : |g| \leq 1 \\ \frac{1 + g\epsilon^*}{\epsilon^* + g^*} & : |g| > 1 \end{cases}. \quad (86)$$

Here the  $*$  stands for complex conjugation,  $\text{Re}$  for the real part, and we call  $g = \gamma/(1 - \kappa)$  the *reduced shear*.  $\gamma := \gamma_1 + i\gamma_2$  is the shear in a complex notation. These transformations have the following properties:

- Only the combination  $\mathbf{g} = \gamma/(1-\kappa)$  of  $\kappa$  and  $\gamma$  appears. Hence the transformation  $(1-\kappa) \rightarrow \lambda(1-\kappa)$  and  $\gamma \rightarrow \lambda\gamma$  with an arbitrary number  $\lambda$  leads to the same observed ellipticities. Consequently, a lens having  $\kappa$  and  $\gamma$  cannot be distinguished from one with the scaled quantities by observing ellipticities. This degeneracy is called the *mass sheet degeneracy*
- The quantity  $\epsilon$  looks more complex at first sight because of the case distinction. But for the weak lensing case,  $\kappa \ll 1$  and  $|\gamma| \ll 1$  and hence always  $|\mathbf{g}| \leq 1$ . A linear approximation leads to the simple relation

$$\epsilon^s \approx \epsilon + \mathbf{g} \approx \epsilon + \gamma \quad (87)$$

what will turn out to be very useful in the next section.

## 4.2 The basic ideas of weak gravitational lensing studies

Looking at eq. (86) we note that we could directly observe the reduced shear  $\mathbf{g}$  of a lens if we knew the intrinsic ellipticity of a source. Galaxies are intrinsically not round, but have a very broad ellipticity distribution instead. To use the galaxy ellipticities nevertheless for shear estimation we rely on the following assumption:

The intrinsic orientation of galaxies is random. In other words: the probability distribution for galaxy ellipticities only depends on the modulus of the ellipticity, but not on its phase.

This assumption is fundamental for all weak lensing applications and, so far, it has not been seriously questioned<sup>†</sup>. Measurements of the ellipticity distribution  $p(|\epsilon^s|)$  can very well be represented by a Gaussian

$$p(|\epsilon^s|) = \frac{\exp(-|\epsilon^s|^2/\sigma_\epsilon^2)}{\pi\sigma_\epsilon^2(1 - \exp(-1/\sigma_\epsilon^2))}. \quad (88)$$

Here,  $p(|\epsilon^s|)d\epsilon_1^s d\epsilon_2^s$  gives the probability to find an ellipticity with components between  $[\epsilon_1^s; \epsilon_1^s + d\epsilon_1^s]$  and  $[\epsilon_2^s; \epsilon_2^s + d\epsilon_2^s]$ . The width  $\sigma_\epsilon$  is typically 0.2. We note that we can only observe the lensed ellipticity distribution. But to first order it is equivalent to the intrinsic distribution of the sources in the weak lens case. Because of the above assumption we get for the mean ellipticity  $\langle \epsilon^s \rangle$  over an ensemble of unlensed ellipticities

$$\langle \epsilon^s \rangle = \iint p(\epsilon_1^s, \epsilon_2^s) \epsilon^s d\epsilon_1^s d\epsilon_2^s = 0. \quad (89)$$

With this consideration, eq. (86) becomes in the weak lensing regime for a local ensemble of galaxies around a position  $\vec{\theta}$

$$\langle \epsilon \rangle = \langle \epsilon^s \rangle - \mathbf{g}(\vec{\theta}) = -\mathbf{g}(\vec{\theta}) \approx -\gamma(\vec{\theta}), \quad (90)$$

and we can statistically determine the shear of a lens.

To make use of the observable  $\gamma$  we return to the lensing potential eq. (50). The surface mass density  $\kappa$  and the gravitational shear  $\gamma$  are both combinations of second derivatives of this potential and the quantities are connected linearly. One can show that

$$\kappa(\vec{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\vartheta \operatorname{Re}[K(\vec{\theta} - \vec{\vartheta})\gamma^*(\vec{\vartheta})] + \text{const}, \quad (91)$$

with the complex kernel  $K(\vec{\theta}) = \frac{\theta_1^2 - \theta_2^2 + 2i\theta_1\theta_2}{|\vec{\theta}|^4}$ . The publication of this formula in Kaiser & Squires 1993 (see [33]) which allows the reconstruction of the lens surface mass density  $\kappa$  from the observable shear  $\gamma$  marked the beginning of weak lensing research. We make some remarks to eq. (91):

---

<sup>†</sup>Very recently there have been claims that galaxy ellipticities have a small intrinsic correlation on scales  $\theta \leq 1'$  (see [25] and [12]). This could affect measurements of cosmic shear (see section 8) on small scales.

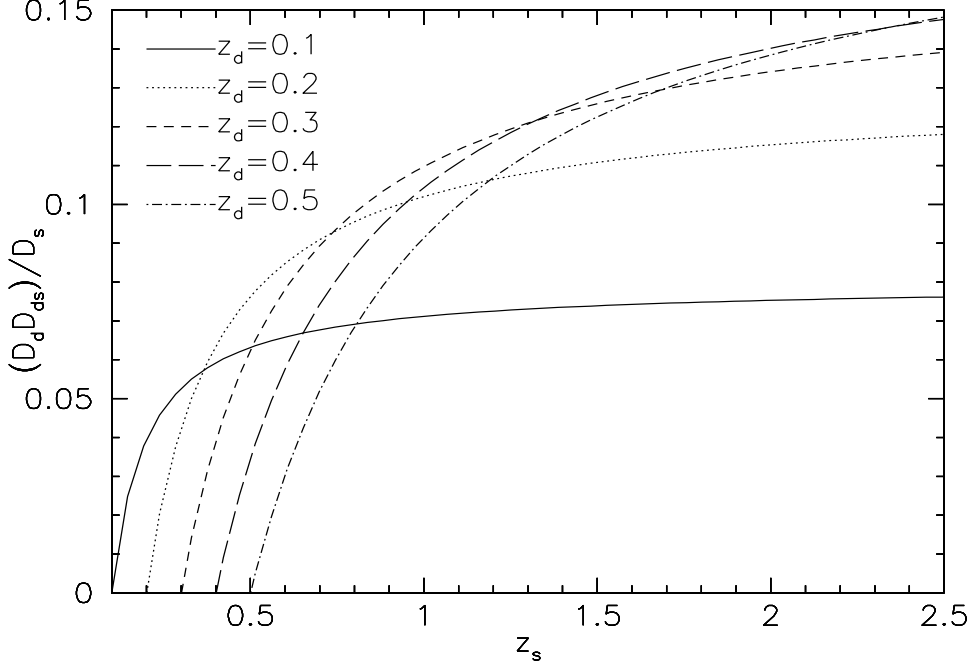


Figure 8: The lens strength  $D_d D_{ds} / D_s$ : The plotted curves show the redshift dependence of  $\gamma$  and  $\kappa$  as a function of the source redshift. We note that there is only a weak redshift dependence for low-redshift lenses but the relative position of source and lens becomes increasingly more important for higher-redshift lenses.

- The inversion formula (91) can be easily transformed to a sum over the discrete source images that we observe

$$\kappa(\vec{\theta}) \simeq -\frac{1}{\pi n} \sum_{j=1}^N \text{Re}[K(\vec{\theta} - \vec{\vartheta}_j) \langle \epsilon^* \rangle(\vec{\vartheta}_j)], \quad (92)$$

where we can obtain an unbiased estimate of the shear  $\gamma$  by local averages over  $\epsilon$  [see eq. (90)]. The sum extends over  $N$  gridpoints where the shear has been estimated. The number-density of these grid-points is  $n$ . Hence, it is straightforward to apply the formalism to data.

- With the reconstruction of  $\kappa$  we can not yet obtain a physical mass. For this the critical surface mass density  $\Sigma_{\text{cr}}$ , that depends on the redshifts of the lens and the sources, has to be specified. In fact, we have up to now completely neglected that the shearstrength  $|\gamma|$  depends on the redshift of lens and source. Using eq. (90) in the form given assumes that all galaxies included in the averaging are at the same redshift. This is a very good approximation if the lens is at low redshift. Figure 8 shows this. For studies of low-redshift clusters, typically a mean redshift  $\langle z_s \rangle$  of the sources is assumed, and  $\Sigma_{\text{cr}}$  and, finally, a physical mass for the cluster is calculated with this mean (the redshift of the cluster lens  $z_d$  is assumed to be known). On the other hand, the redshift distribution of the sources has a strong influence on the mass estimate of high-redshift clusters. Seitz & Schneider 1996 (see [64]) have demonstrated how to generalise eq. (90) and the subsequent formulae in cases where the redshift distribution of the sources is known.
- The integral of eq. (91) is an integral over the whole  $\mathbb{R}^2$ . In practice, we have data only on a limited field-of-view, and a straight application of the inversion leads to border effects. This problem has been solved by working out inversion formulae performing an exact mass reconstruction on finite data fields. (so-called *finite field inversion formulae*).

- If we want to apply the above formula also for regimes where  $\kappa \ll 1$  no longer holds we have to take into account that only the reduced shear  $\mathbf{g} = \boldsymbol{\gamma}/(1 - \kappa)$ , but not the shear itself can be measured. In eq. (91) we can substitute  $\boldsymbol{\gamma}^*$  by  $\mathbf{g}^*(1 - \kappa)$  which leads to an integral equation for  $\kappa$  that can be solved iteratively.
- The integral in eq. (91) is only determined uniquely up to an additive constant  $\kappa_0$ . This constant is often chosen such that the reconstructed  $\kappa$  equals zero at the border of the field which in turn leads to a lower limit for the mass. We note that this additive constant is not only an artefact of the integral. Looking at eq. (65) we see that  $\kappa = \text{const}$  implies  $\boldsymbol{\gamma} = 0$ . Thus, a constant sheet of mass does not produce any shear and we cannot infer it from lensing based on shape measurements.

Taking another look at the two upper panels of Fig. 7 we notice that the number density of objects around the lens changes in the presence of a mass. The lens changes the area on which we apparently see sources according to the Jacobi matrix of the mapping, and it reads

$$\frac{A_s}{A_l} = |\det A| = \frac{1}{\mu} = |(1 - \kappa^2) - |\boldsymbol{\gamma}|^2| \quad (93)$$

where  $A_l$  and  $A_s$  are the areas with and without the lens respectively. On the other hand, the magnification enhances the apparent flux of objects as discussed in section 3.3.2. If our observed galaxy sample is limited by the observed flux, gravitational lensing on the one hand lowers the observed number density because of the area effect, on the other hand sources that may be unobservable come into our sample via the flux magnification. This is known as *magnification bias*. The observed number density of sources then reads:

$$n(\vec{\theta}, S) = \frac{1}{\mu}(\vec{\theta})n_0\left(\frac{S}{\mu(\vec{\theta})}\right), \quad (94)$$

where  $n_0(S)$  is the number density of sources with flux greater than  $S$  that we would observe in the absence of lensing. If we assume a power law of the form  $n_0 \sim S^{-\beta}$  for the unlensed number counts, we get

$$n(\vec{\theta}) = n_0\mu(\vec{\theta})^{\beta-1}. \quad (95)$$

For observations in blue optical bands we obtain a slope  $\beta$  of about unity, and so there is no net effect. But in the red part we observe  $\beta \approx 0.5$ , and we expect a depletion of the unlensed number counts<sup>†</sup> We note that for the weak lensing regime one obtains

$$\mu(\vec{\theta}) \approx 1 + 2\kappa(\vec{\theta}), \quad (96)$$

so that measuring the magnification is another method for studying the mass distribution of a lens. If we want to utilise the magnification bias for studying the lens properties we need accurate knowledge about the unlensed number counts. This substitutes the assumption of a random mean intrinsic ellipticity when utilising the shear effect. We will discuss the magnification effect in more detail in the next section.

---

<sup>†</sup>if we consider the number counts as function of the apparent magnitude  $m$ ,  $n_0(\vec{\theta}) \sim 10^{\beta' m}$ , we get for the lensed number counts  $n(\vec{\theta}) = n_0\mu^{2.5\beta'-1}$ . For the blue optical bands the slope  $\beta' \approx 0.4$  but for the red optical bands below 0.2.

## 5 Combining magnification and shear information to constrain cluster mass profiles

### Summary:

Together with Peter Schneider and Lindsay King, a theoretical work was done on how shear and magnification information can be combined to optimally constrain parameters of cluster mass profiles. For this we have used a Maximum Likelihood formalism. In the first part of this chapter the ideas behind Maximum Likelihood estimators are introduced and then the work itself is presented. It was published in [61].

### 5.1 Maximum Likelihood estimators

Let us assume we have made several measurements of a quantity  $y_j = y(x_j, a_i)$  ( $j = 1 \dots N$ ) that depends on some known quantities  $x_j$  and some unknown quantities  $a_i$  that we want to determine. We denote the complete set of the  $N$  different, independent measurements of  $y$  by  $\hat{y}$ . Furthermore, we have given a function  $P(\hat{y}|a_i)$  that returns the probability of getting the measurement set  $\hat{y}$  if the parameters  $a_i$  are known. This knowledge seems irrelevant for the solution of our problem at first sight. But we can turn the question around and ask: how probable is the parameter set  $a_i$  with the given measurement of  $\hat{y}$ ? This probability  $P(a_i|\hat{y})$  is called *likelihood*. We would then intuitively choose that parameter set which maximises the likelihood. This way of parameter estimation is called *Maximum Likelihood estimation*. A connection between the probabilities  $P(\hat{y}|a_i)$  and  $P(a_i|\hat{y})$  is given by *Bayes theorem* stating

$$P(a_i \cap \hat{y}) = P(\hat{y}|a_i)P(a_i) = P(a_i|\hat{y})P(\hat{y}), \quad (97)$$

where  $P(a_i \cap \hat{y})$  is the probability that  $a_i$  is correct and the value  $\hat{y}$  is measured.  $P(a_i)$  is the probability for the parameter set  $a_i$  and  $P(\hat{y})$  for that of  $\hat{y}$ . To use the above formula for calculating the likelihood  $P(a_i|\hat{y})$  we need an assumption about  $P(a_i)$  ( $P(\hat{y})$  is the probability to get a specific measurement value and so a constant). The required assumption is called *Bayes postulate* and it says:

In the absence of any knowledge about the parameter set  $a_i$ , all possible sets should be treated equally and  $P(a_i)$  should be set to a constant.

If we accept this postulate, the known probability  $P(\hat{y}|a_i)$  is sufficient for a maximum likelihood estimation. We get for the likelihood  $\mathcal{L}$

$$\mathcal{L} = P(a_i|\hat{y}) = P(\hat{y}|a_i) \text{const.} \quad (98)$$

For our parameters we then accept the parameters  $a_i$  maximising the likelihood  $\mathcal{L}$  (or equivalently minimising the *log-likelihood*  $l := -\ln \mathcal{L}$ ).

- One of the strengths of the maximum likelihood analysis is that additional information easily can be included in the formalism. Suppose we have a second, independent measurement  $\hat{y}'(x'_j, a_i)$  and a function  $P'(\hat{y}'|a_i)$  for the probability of the result  $\hat{y}'$  if our model has the true parameters  $a_i$ . For this we can, in analogy to above, perform a maximum likelihood analysis by maximising  $\mathcal{L}' := P'(\hat{y}'|a_i)$ . But in addition our information from the measurements  $\hat{y}$  and  $\hat{y}'$  can be composed to a combined likelihood function

$$\mathcal{L}_{\text{tot}} := \mathcal{L}\mathcal{L}'. \quad (99)$$

By maximising this function we then get the best parameter values  $a_i$  using all the available information. This formalism can naturally be extended to  $N$  independent measurements.

- The other side of the coin is that estimating errors for the obtained values of  $a_i$  is quite difficult in general. We consider the application of the maximum likelihood technique to measurements whose errors are distributed according to a Gaussian. In this case, the probability to find a set of independent measurements  $y_j$  is given by

$$P \sim \prod_{j=1}^N \exp \left[ -\frac{1}{2} \left( \frac{y_j - y(x_j, a_i)}{\sigma_j} \right)^2 \right], \quad (100)$$

and minimising the log-likelihood  $-\ln P$  leads to the usual  $\chi^2$  analysis to estimate the parameters  $a_i$

$$\chi^2 = \sum_{j=1}^N \left( \frac{y_j - y(x_j, a_i)}{\sigma_j} \right)^2. \quad (101)$$

In this case it is known that the dispersion of obtained best fit parameters  $a_i$  follows a  $\chi^2$  distribution of  $\nu$  degrees of freedom around the true values, where  $\nu$  is the number of parameters  $a_i$ . If we limit ourselves to two free parameters  $a_1$  and  $a_2$  (this will be the case below), we introduce so-called *confidence regions* that are closed curves in the plane of  $a_1$  and  $a_2$  and contain 68.3%, 90% and 95.4% etc. of the estimates for  $a_1$  and  $a_2$ . In the case of the  $\chi^2$  statistics the confidence intervals, that we use later on, are:

$$\begin{aligned} \chi^2 - \chi^2_{\min} = 1, 15 &\Rightarrow 68.3\% && \text{of all estimates} \\ \chi^2 - \chi^2_{\min} = 2, 3 &\Rightarrow 90\% && \text{of all estimates} \\ \chi^2 - \chi^2_{\min} = 3, 1 &\Rightarrow 95.4\% && \text{of all estimates} \\ \chi^2 - \chi^2_{\min} = 4.6 &\Rightarrow 99\% && \text{of all estimates} \\ \chi^2 - \chi^2_{\min} = 5.9 &\Rightarrow 99.7\% && \text{of all estimates} \\ \chi^2 - \chi^2_{\min} = 9.2 &\Rightarrow 99.99\% && \text{of all estimates} \end{aligned} \quad (102)$$

In the general case, if the errors are not necessarily Gaussian distributed, the probability function for the log-likelihood can usually not be calculated analytically. In this case the confidence regions have to be determined with computer intensive *Monte Carlo simulations*. Below we will use the confidence regions of the  $\chi^2$  distribution as an approximation for error estimates in our analysis first, and verify their validity with simulations later.

## 5.2 Motivation and basic ideas for our work

In section 4.2 we saw that we can get information about the surface mass density  $\kappa$  of a lens by observing shear causing coherent distortions of background galaxies, and by measuring the magnification causing a depletion in red galaxy number counts behind the lens. While the techniques utilising the shear method have experienced large improvements and extensions in the past years, and quantitative analysis has been done with it for more than a dozen galaxy clusters up to now (see e.g. [65], [26], [36]), the use of the magnification is still in its infant stage. Up to now it has been detected in at least three clusters (see [67], [20], [23]). This is counter-intuitive at first glance as the measurement of accurate object shapes is in practice much more difficult than pure object detection and flux measurement (see also section 6). The reasons why the magnification is being utilised only now are the following:

- We will show below a simple consideration about the signal-to-noise ratio of the detection for the two lensing signatures. It turns out that a significant detection for a shear signal is easier to obtain than for the magnification effect.
- The assumption of the shear method, that the mean intrinsic ellipticity of unlensed sources vanishes, is well founded and allows a simple calculation of errors in shear estimates. Calculating the reduced shear  $\mathbf{g}$  with  $N$  galaxies, the uncertainty is  $\sigma_\epsilon/\sqrt{N}$  for the Gaussian ellipticity distribution of eq. (88). It can be obtained from the background galaxies themselves.



In contrast, for the magnification an external calibration for the unlensed number counts is necessary. One strategy is to use *empty fields* near the lensing cluster as a reference. This requires additional data besides the images for the lensing cluster and can be obtained only for free with instruments covering *wide fields* with a single exposure. Although these data are becoming available now, it is not yet sure how accurate we can determine the unlensed number counts. As possible systematic errors in these immediately propagate to our measurements, the overall usefulness of the magnification method for quantitative analysis is still under debate.

Let us now consider with a simple approximation how easily a highly significant shear and magnification signal can be detected. If we limit our measurement to a region of approximately constant reduced shear, the signal in this case is the modulus of the reduced shear  $\mathbf{g}$  and the noise is  $\sigma_\epsilon/\sqrt{N}$  (see above). The signal-to-noise in this case is:

$$\left(\frac{S}{N}\right)_\gamma = \frac{|\mathbf{g}|\sqrt{N_\gamma}}{\sigma_\epsilon}. \quad (103)$$

For the magnification, the signal is the difference between the observed number of galaxies  $\mu^{\beta-1}N_\mu$  and the expected number of galaxies  $N_\mu$  in the absence of lensing, that is  $\Delta N = |\mu^{\beta-1} - 1|N_\mu$ . If we neglect the effect of galaxy clustering, the noise is the Poisson noise  $\sqrt{N_\mu}$ , and the signal-to-noise ratio for a magnification detection reads

$$\left(\frac{S}{N}\right)_\mu = |\mu^{\beta-1} - 1|\sqrt{N_\mu}. \quad (104)$$

Specialising these expressions to the weak lensing limit  $\mathbf{g} \approx \gamma$ ,  $\mu \approx 1 + 2\kappa$ ,  $\Delta N \approx 2\kappa|1 - \beta|N_\mu$  and considering the ratio of the signal-to-noise estimates we get:

$$\frac{(S/N)_\gamma}{(S/N)_\mu} = \frac{|\gamma|}{\kappa} \frac{1}{2\sigma_\epsilon|1 - \beta|} \sqrt{\frac{N_\gamma}{N_\mu}}. \quad (105)$$

As discussed in section 4.2 observations in a near infrared passband have  $\beta \approx 0.5$ . Furthermore we consider the case of a singular isothermal sphere [eq. (68)] where  $|\gamma| = \kappa$  and take  $\sigma_\epsilon = 0.2$ . Then the above ratio becomes

$$\frac{(S/N)_\gamma}{(S/N)_\mu} \approx 5\sqrt{\frac{N_\gamma}{N_\mu}}. \quad (106)$$

To make this ratio of order unity and thus to get the same significance for a magnification detection as for a shear detection, the number of galaxies for which we can measure only position and flux would have to exceed 25 times those for which accurate object shapes can be obtained. Typical galaxy number densities on high-quality data are  $n_\gamma \approx 20 - 30 \text{ arcmin}^{-2}$  and  $n_\mu \approx 100 - 120 \text{ arcmin}^{-2}$ . Hence, the ratio  $N_\mu/N_\gamma$  is typically  $\approx 4$ . From this simple estimate we see that the shear is easier to detect.

One of the interesting questions concerning cluster mass profiles is whether they can be described by a universal profile. Navarro, Frenk & White 1996 (see [48]) have shown with simulations that dark matter haloes can be fitted by a two parameter model (the so-called *NFW profile*) over a large range of mass. An important ingredient to test this hypothesis is to investigate the slope of cluster profiles at large radii. The projected mass density  $\kappa$  of the NFW profile has there a  $1/\theta^2$  behaviour, whereas the often used SIS drops as  $1/\theta$ .

Weak lensing, tracing directly the mass, is the ideal tool for this work. Tom Broadhurst claimed in 1995 that not the shear, but the magnification is the lensing effect of choice when determining the slopes of profiles. The investigation of the substance of this claim was the motivation for this work. Furthermore, so far no technique has been known that combines shear and magnification information to use all available information for constraining parametrised cluster profiles. We

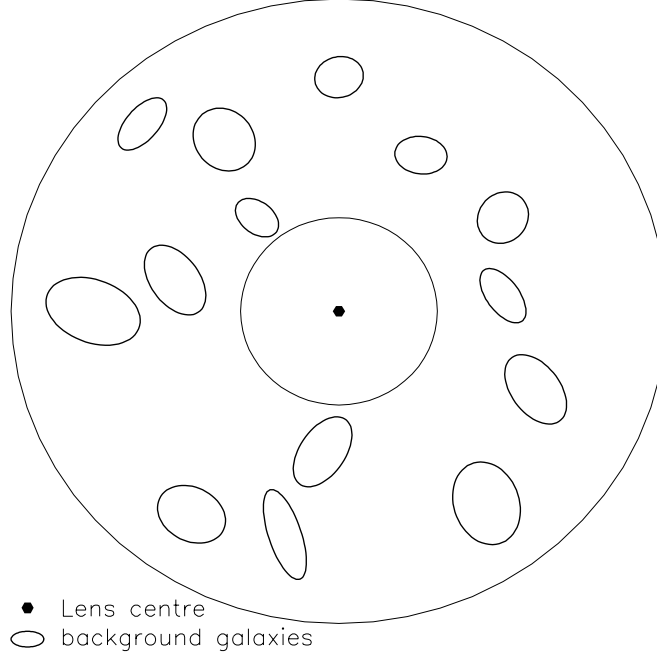


Figure 9: Likelihood construction: We consider background galaxies in an annulus around a galaxy cluster. The basic idea of our likelihood approach to constrain a parametrised model of the cluster consists of the questions: What is the probability to find  $N_\mu$  galaxies at the positions where we observe them and what is the probability that  $N_\gamma$  galaxies have the observed ellipticity distribution?

started with a parametrised model for the cluster profile and constructed an algorithm for this combined analysis using maximum likelihood techniques.

The idea for the construction of the likelihood function is visualised in Fig. 9. We consider all background galaxies in an annulus with  $\theta_{\text{in}} < \theta < \theta_{\text{out}}$  around a galaxy cluster. If we have a parametrised model for the cluster we can ask the following questions:

- What is the probability for a certain parameter set  $\pi_i^\mu$  to find  $N_\mu$  galaxies in the annulus and what is the probability to find the galaxies at the positions where we observe them. This probability will be the likelihood  $\mathcal{L}_\mu$  based on magnification information.
- What is the probability for a certain parameter set  $\pi_i^\gamma$  that the  $N_\gamma$  background galaxies (this sample will in general be different from the  $N_\mu$  galaxies in the first case), for which we can measure reliable shapes, have the observed ellipticity distribution. This probability will form the likelihood  $\mathcal{L}_\gamma$  based on gravitational shear.

Dealing with both questions at the same time will result in a combined likelihood  $\mathcal{L}_{\text{tot}}$ .

As we see in Fig. 6, the centre of galaxy clusters is dominated by some bright and large galaxies that make the detection and analysis of background galaxies behind and close to them impossible. For this reason this area cannot be used for our analysis and an annulus excluding this region is chosen.  $\theta_{\text{min}}$  is chosen larger than the Einstein radius so that we avoid the strong lensing regime and we can use the weak lensing limit in the following. Also, the area of the annulus itself may be contaminated by bright stars or bright foreground galaxies, and we then have to mask out these regions as well. This masking (influencing the available area and thus the number density

of objects) has a tremendous effect on the magnification method as demonstrated in Taylor et al. 1998 (see [67]).

### 5.3 The calculation of the likelihood functions

For the construction of the likelihood functions in this section the following three assumptions are used:

- The unlensed number counts  $n_\mu$  are known exactly. We will investigate the consequences of giving up this assumption later on.
- The cluster lens is at a fixed low redshift and hence all sources are assumed to lie at the same redshift. Relaxing this assumption was not done in the course of this work.
- We know the size of the Einstein radius of our galaxy cluster. This radius is an observable by the location of giant arcs. Giving up this constraint was not investigated in this work as well. Throughout this work an Einstein radius of  $\theta_E = 0.5'$  is used.

We start with the likelihood function using magnification information. The expected number of galaxies in our annulus with  $\theta_{\text{in}} < \theta < \theta_{\text{out}}$  is

$$\langle N_\mu \rangle = n_\mu \int d^2\theta [\mu(\vec{\theta})]^{\beta-1}, \quad (107)$$

where the integral extends over our annulus. Neglecting the effect of galaxy clustering (we will come back to this below),  $N_\mu$  follows a Poisson distribution. Moreover the probability that the  $i$ -th observed galaxy is located at  $\vec{\theta}_i$  is proportional to  $n(\vec{\theta}_i) \propto [\mu(\vec{\theta}_i)]^{\beta-1}$ . Then the likelihood  $\mathcal{L}_\mu$  becomes

$$\mathcal{L}_\mu = P(N_\mu; \langle N_\mu \rangle) \prod_{i=1}^{N_\mu} \frac{[\mu(\vec{\theta}_i)]^{\beta-1}}{\int d^2\theta [\mu(\vec{\theta})]^{\beta-1}}, \quad (108)$$

with

$$P(N; \langle N \rangle) = \frac{\langle N \rangle^N}{N!} \exp(-\langle N \rangle). \quad (109)$$

Using that the denominator of the product in eq. (108) equals  $\langle N_\mu \rangle / n_\mu$  [see eq. (107)] the factor  $\langle N_\mu \rangle^{N_\mu}$  of the Poisson distribution cancels. Furthermore, the factors  $N_\mu!$  and  $n_\mu^{N_\mu}$  do not depend on the lens model and are unimportant when we determine the maximum of the likelihood. So, the relevant terms of the log-likelihood read

$$l_\mu := -\ln \mathcal{L}_\mu = n_\mu \int d^2\theta [\mu(\vec{\theta})]^{\beta-1} + (1-\beta) \sum_{i=1}^{N_\mu} \ln \mu(\vec{\theta}_i). \quad (110)$$

We now turn to the likelihood function that uses the ellipticity information. The probability distribution for the observed ellipticities can be calculated from that of the intrinsic ellipticities by

$$p_\epsilon(\epsilon|\mathbf{g}) = p_\epsilon^s(\epsilon^s(\epsilon, \mathbf{g})) \left| \frac{\partial^2 \epsilon^s}{\partial^2 \epsilon} \right| = p_\epsilon^s(\epsilon^s(\epsilon, \mathbf{g})) \frac{(|\mathbf{g}|^2 - 1)^2}{|\epsilon \mathbf{g}^* + 1|^4} \quad (111)$$

if we use the  $|\mathbf{g}| < 1$  case of eq. (86). The corresponding probability density  $p_\epsilon(\epsilon_i)$  for each lensed galaxy in the catalog is determined by using  $\mathbf{g}(\vec{\theta}, \pi^\gamma)$  in eq. (111), and the likelihood and log-likelihood functions are given by

$$\mathcal{L}_\gamma = \prod_{i=1}^{N_\gamma} p_\epsilon(\epsilon_i | \mathbf{g}(\vec{\theta}_i, \pi^\gamma)); \quad l_\gamma = - \sum_{i=1}^{N_\gamma} \ln p_\epsilon(\epsilon_i | \mathbf{g}(\vec{\theta}_i)). \quad (112)$$

Instead of using the exact expressions from eq. (111) we make the following simplifications: For the intrinsic ellipticity distribution we assume the Gaussian profile of eq. (88). In the following we approximate the distribution of observed ellipticities again with a Gaussian that reproduces the first and second moments of the true distribution from eq. (111). This assumption is reasonable as long as the modulus of the reduced shear is not close to unity. Because of eq. (87) the first moment of this Gaussian is simply the reduced shear  $-\mathbf{g}$ . We can show that the second moment, the dispersion  $\sigma$  can be well approximated by

$$\sigma[\mathbf{g}(\vec{\theta})] \approx (1 - |\mathbf{g}(\vec{\theta})|^2) \sigma_\epsilon . \quad (113)$$

The shear likelihood function then becomes

$$\mathcal{L}_\gamma = \prod_{i=1}^{N_\gamma} \frac{1}{\pi \sigma^2[\mathbf{g}(\vec{\theta}_i)]} \exp\left(-\frac{|\epsilon_i + \mathbf{g}(\vec{\theta}_i)|^2}{\sigma^2[\mathbf{g}(\vec{\theta}_i)]}\right) , \quad (114)$$

and the corresponding log-likelihood function becomes, dropping irrelevant additive constants,

$$l_\gamma = \sum_{i=1}^{N_\gamma} \left[ \frac{|\epsilon_i + \mathbf{g}(\vec{\theta}_i)|^2}{\sigma^2[\mathbf{g}(\vec{\theta}_i)]} + 2 \ln \sigma[\mathbf{g}(\vec{\theta}_i)] \right] . \quad (115)$$

The likelihood/log-likelihood function for the combination of shear and magnification information is then obtained by multiplying/adding the corresponding terms.

$$\mathcal{L}_{\text{tot}} = \mathcal{L}_\mu \mathcal{L}_\gamma \quad l_{\text{tot}} = l_\mu + l_\gamma . \quad (116)$$

In the following only the log-likelihood function will be considered. We will now calculate analytic expressions for the “expected” likelihood functions.

## 5.4 The expected average log-likelihood

For a given set of observables for a galaxy cluster, the log-likelihood functions defined in the previous section can be calculated; the minimum yields the best-fitting model parameters, and the width of the log-likelihood function yields the confidence region for the parameter estimate from this data set as discussed in section 5.1. Let us assume that we could draw other, independent sets of observables from the same cluster (we assume that we have different realisations for the total number and positions of galaxies and the intrinsic ellipticities, drawn from the underlying probability distributions). For all these realisations we could perform a log-likelihood analysis. In this way we could determine a *average loglikelihood* adding up the different likelihoods. This would yield the “expected” errors for this model. One way to estimate this average log-likelihood function is by Monte Carlo simulations. In the present case we even can perform all necessary *ensemble-averaging* processes analytically.

Given a lens model – hereafter called the true model, and characterised by subscripts ‘t’ – the necessary ensemble-average of a quantity  $X$  is provided by

$$\langle X \rangle := \sum_{N=0}^{\infty} P(N; \langle N \rangle_t) \times \left[ \prod_{i=1}^N \int d^2\theta_i p_t(\vec{\theta}_i) \int d^2\epsilon_i p_t(\epsilon_i) \right] X , \quad (117)$$

where the first term averages over the probability to find  $N$  galaxies when the expectation value from the true model is  $\langle N \rangle_t$ , the second term is the integration over the probability that galaxy number  $i$  lies at  $\vec{\theta}_i$ , and the final term integrates over the ellipticity distribution. Note that this prescription has to be extended if we include new statistical quantities in the calculation of our log-likelihoods. For instance the inclusion of a redshift distribution for the background galaxies would result in an additional averaging.

We want to apply this averaging on our log-likelihood function  $l_\mu$  and  $l_\gamma$  now. Starting with  $l_\mu$  we see from eq. (108) that it does not depend on the image ellipticities, and the final integral in eq. (117) just contributes a factor of 1, giving

$$\langle l_\mu \rangle = \sum_{N_\mu=0}^{\infty} P(N_\mu; \langle N_\mu \rangle_t) \left[ \prod_{i=1}^{N_\mu} \int d^2\theta_i p_t(\vec{\theta}_i) \right] l_\mu, \quad (118)$$

with

$$\langle N_\mu \rangle_t = n_\mu \int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1}. \quad (119)$$

The first term in the log-likelihood function eq. (108) depends neither on  $\vec{\theta}_i$  nor on the actual number  $N_\mu$  of galaxies, and thus the averaging operator in eq. (118) leaves this term unchanged. The second term of  $l_\mu$  consists of a sum over individual galaxy images; therefore, for the  $j$ -th galaxy image, only the term  $i = j$  in the spatial averaging in eq. (118) contributes, the other integrations just yield a factor one. Since there are  $N_\mu$  equal terms, one obtains

$$\begin{aligned} \langle l_\mu \rangle &= n_\mu \int d^2\theta [\mu(\vec{\theta})]^{\beta-1} \\ &+ (1 - \beta) \sum_{N_\mu=0}^{\infty} P(N_\mu; \langle N_\mu \rangle_t) \frac{N_\mu}{\int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1}} \\ &\times \int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1} \ln \mu(\vec{\theta}). \end{aligned} \quad (120)$$

With  $\sum_{i=0}^{\infty} \frac{x^i}{i!} = \exp(x)$ , and using eq. (119) we get

$$\begin{aligned} \langle l_\mu \rangle &= n_\mu \int d^2\theta [\mu(\vec{\theta})]^{\beta-1} \\ &+ n_\mu(1 - \beta) \int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1} \ln \mu(\vec{\theta}). \end{aligned} \quad (121)$$

Turning now to the average log-likelihood for the shear we insert eq. (115) in eq. (117). We note that  $l_\gamma$  is a sum over galaxy images. So the integration works term by term resulting in  $N_\gamma$  identical terms. This then immediately allows one to perform the sum over  $N_\gamma$ , as before, so that

$$\begin{aligned} \langle l_\gamma \rangle &= \langle N_\gamma \rangle_t \int d^2\theta \frac{[\mu_t(\vec{\theta})]^{\beta-1}}{\int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1}} \\ &\times \int d^2\epsilon \frac{1}{\pi\sigma_t^2(\vec{\theta})} \exp\left(-\frac{|\epsilon + \mathbf{g}_t(\vec{\theta})|^2}{\sigma_t^2(\vec{\theta})}\right) \\ &\times \left( \frac{|\epsilon + \mathbf{g}(\vec{\theta})|^2}{\sigma^2(\vec{\theta})} + 2 \ln \sigma(\vec{\theta}) \right), \end{aligned} \quad (122)$$

and  $\langle N_\gamma \rangle_t$  is defined in analogy to eq. (119), with  $n_\mu$  replaced by  $n_\gamma$ . We obtain

$$\begin{aligned} \langle l_\gamma \rangle &= n_\gamma \int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1} \\ &\times \left( \frac{|g(\vec{\theta}) - g_t(\vec{\theta})|^2 + \sigma_t^2(\vec{\theta})}{\sigma^2(\vec{\theta})} + 2 \ln \sigma(\vec{\theta}) \right). \end{aligned} \quad (123)$$

Before presenting results with this average log-likelihood functions we introduce the lens models that we considered.

## 5.5 The lens models

Two families of two-parameter lens models have been considered. As the most important aspect of our work was to see how well we can recover the slope of cluster mass profiles in the weak lensing regime, the first of our models (hereafter Family A) is the power-law model from eq. (66). As the central part of the mass profiles is not included in our analysis we do not care about the unphysical properties of the model there. Besides the slope we also try to infer the amplitude  $a$  with our likelihood analysis; therefore the free parameters of our model are  $a$  and  $q$ . As a second mass profile (hereafter Family B) we chose the lens model described in Sect. 8.1.5 of SEF. It reads:

$$\kappa(\theta) = \kappa_0 \frac{1 + p(\theta/\theta_c)^2}{[1 + (\theta/\theta_c)^2]^{2-p}}, \quad (124)$$

where  $\theta_c$  is a *core radius*, and  $\kappa_0$  the central surface mass density. For large angular radii  $\theta \gg \theta_c$ , it behaves like a power-law model determined by the exponent  $p$ , whereas it is nearly constant for  $\theta \ll \theta_c$ . To be able to make comparisons between the models, we demand that they should coincide for large radii where they both behave like a power law. For this comparison we now convert the quantities  $\theta_c$ ,  $\kappa_0$  and  $p$  into the corresponding quantities  $\theta_E$ ,  $a$  and  $q$ , respectively, from Family A.

The Einstein radius  $\theta_E$  is defined so that  $\bar{\kappa} = 1$  within it. This leads to

$$\theta_E = \theta_c \sqrt{\kappa_0^{1/(1-p)} - 1} =: \theta_c \sqrt{W}, \quad (125)$$

and only exists if  $\kappa_0 > 1$ . This eliminates  $\theta_c$ . For large angular radii the model then behaves like

$$\kappa \rightarrow \kappa_0 p W^{p-1} (\theta/\theta_E)^{2p-2}. \quad (126)$$

The comparison with Family A then transforms  $\kappa_0$  and  $p$  into  $a$  and  $q$

$$a = \kappa_0 p W^{p-1} \quad q = 2 - 2p \quad \Rightarrow \quad 1/W = [2a/(2-q)]^{2/q} - 1 \quad \kappa_0 = (1+W)^{q/2}. \quad (127)$$

The form of Family B that is used henceforth finally reads

$$\kappa(\theta) = \left( \frac{2a}{2-q} \right) \frac{\left( \frac{2a}{2-q} \right)^{2/q} - 1 + \left( \frac{2-q}{2} \right) \left( \frac{\theta}{\theta_E} \right)^2}{\left[ \left( \frac{2a}{2-q} \right)^{2/q} - 1 + \left( \frac{\theta}{\theta_E} \right)^2 \right]^{1+q/2}}. \quad (128)$$

The constraint  $\kappa_0 > 1$  for the existence of an Einstein translates into

$$a > \frac{2-q}{2}. \quad (129)$$

With the formula for the shear given in eq. (65) we can calculate the reduced shear and the magnification for the two families. In Fig. 10 we plot, for four different combinations of  $a$  and  $q$ , the radial mass profile  $\kappa(\theta)$ , the magnification signature  $\mu^{\beta-1}$  for  $\beta = 0.5$ , and the modulus of the reduced shear  $|\mathbf{g}|$ , for both Families of models. The upper panels show the radial dependence of the surface mass densities  $\kappa(\theta)$  which are all fairly different over the range of angles plotted. The bottom panels show the radial dependence of the reduced shear; here we see, in particular for Family A, that models with the same amplitude but different slope have nearly the same curves  $|\mathbf{g}(\theta)|$ ; hence, it will be quite challenging to determine the radial slope  $q$  for these models from the shear method. The reason for this similarity comes from an accidental cancellation of the slope dependencies of  $1 - \kappa$  and  $\gamma$  in  $\mathbf{g}$  for these models. In contrast to the reduced shear, the magnification signal, plotted as  $\mu^{\beta-1} = \mu^{-0.5}$  in the middle panels, differs clearly between models with different slopes. Hence, despite the conclusion reached in section 5.2 on the relative sensitivity of the shear and magnification methods, from the curves in Fig. 10 one might already expect that the magnification effect could more easily distinguish between different radial slopes. This will now be quantified in the next section.

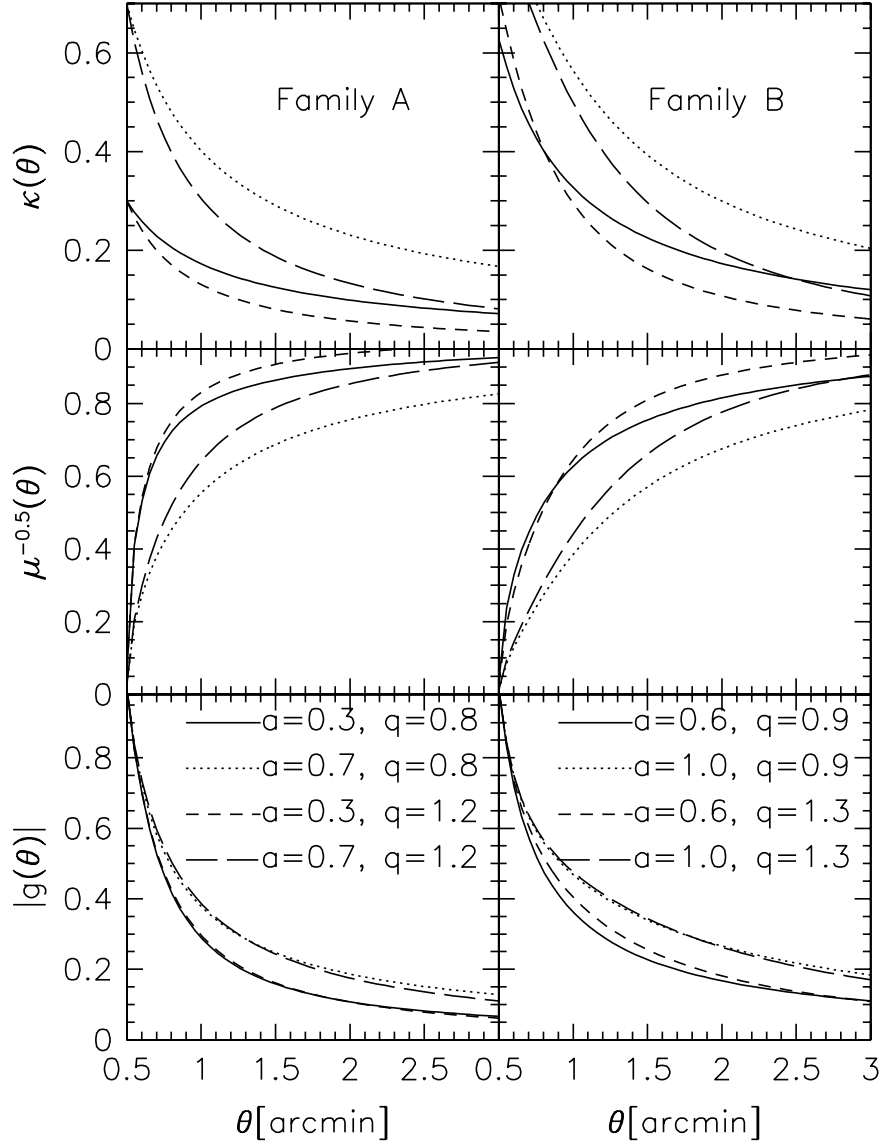


Figure 10: The radial dependence of  $\kappa$  (upper panels),  $|g|$  (lower panels) and  $\mu$  (middle panels) for Families A (left panels) and B (right panels) (Figure 1 from [61])

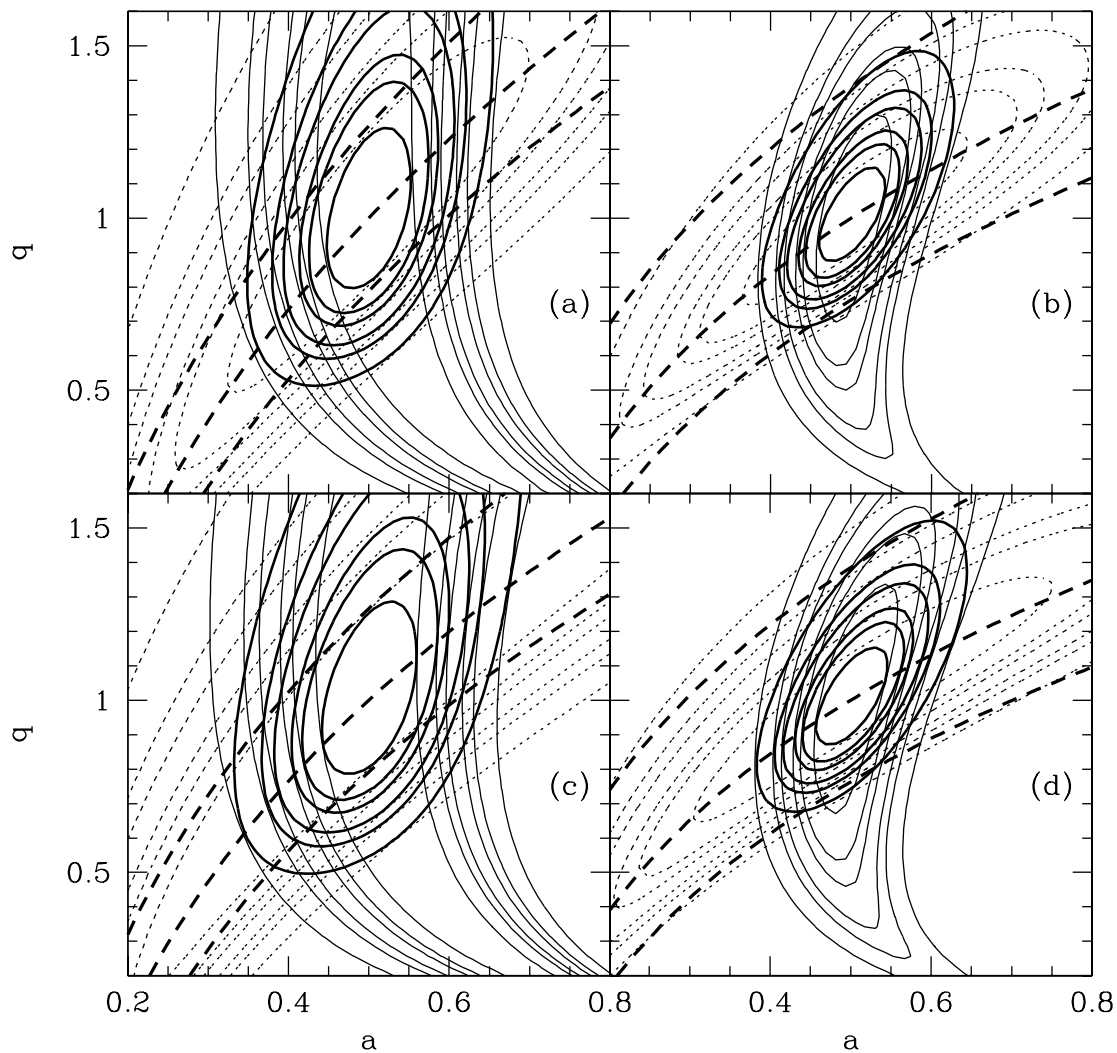


Figure 11: Contours of constant average likelihood for models of Family A (Figure 2 from [61]): The thin solid lines are contours of constant  $\langle l_\gamma \rangle$ , the dashed contours correspond to  $\langle l_\mu \rangle$ , and the heavy solid contours to  $\langle l_{\text{tot}} \rangle$ . Contours are drawn for the error levels listed in eq. (102). The input model is described by  $a = 0.5$ ,  $q = 1.0$ , and corresponds to the minimum of  $\langle l \rangle$ . Parameters are as described in the text, i.e.,  $n_\mu = 120 \text{ arcmin}^{-2}$ ,  $n_\gamma = 30 \text{ arcmin}^{-2}$ ,  $\beta = 0.5$ ,  $\sigma_\epsilon = 0.2$ , and the inner and outer radii of the annulus are (a)  $\theta_{\text{in}} = 0.6'$ ,  $\theta_{\text{out}} = 2.0'$ ; (b)  $\theta_{\text{in}} = 0.6'$ ,  $\theta_{\text{out}} = 4.0'$ ; (c)  $\theta_{\text{in}} = 0.9'$ ,  $\theta_{\text{out}} = 2.0'$ ; (d)  $\theta_{\text{in}} = 0.9'$ ,  $\theta_{\text{out}} = 4.0'$ . The three heavy solid curves are curves of constant  $\langle N_\mu \rangle$ ; the middle one is for  $\langle N_\mu \rangle$  fixed at the value obtained for the input model, the two others are for changes of  $\langle N_\mu \rangle$  by  $\pm 5\%$



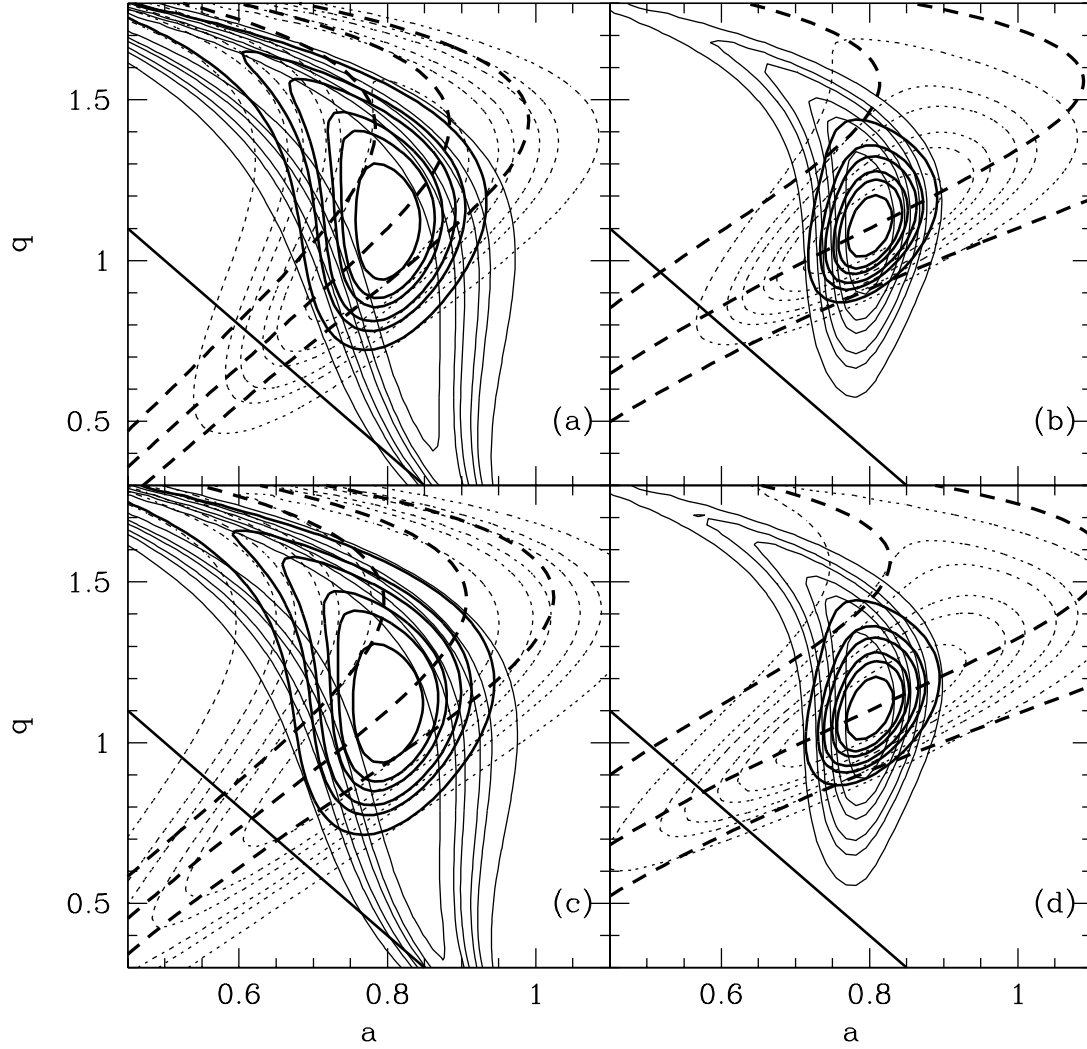


Figure 12: Contours of constant average likelihood for models of Family A (Figure 3 from [61]): The input model is characterised by  $a = 0.8$ ,  $q = 1.1$ . Inner and outer radii are the same as in Fig. 11. The heavy solid line in each panel denotes the limit in eq. (129); models below this line do not correspond to centrally condensed mass distributions with  $\kappa_0 > 1$

## 5.6 Results from the average log-likelihood analysis

We have calculated the average log-likelihood for the Families A and B for four different combinations of inner and outer radii  $\theta_{\text{in}}$  and  $\theta_{\text{out}}$ . In figures 11 and 12 we show contours of constant likelihood for  $\langle l_\gamma \rangle$ ,  $\langle l_\mu \rangle$  and  $\langle l_{\text{tot}} \rangle$  for these combinations. The contours correspond to the confidence regions of the  $\chi^2$  distribution listed in eq. (102). The results for the two Families are qualitatively the same, so we discuss those for Family A in more detail only:

- As we could already expect from the results of Fig. 10, the shear has severe problems in constraining the slope of the profile. The contours for  $\langle l_\gamma \rangle$  are nearly parallel to the  $q$ -axis. On the other hand, the amplitude is constrained very well by the shear.
- In contrast, the contours for  $\langle l_\mu \rangle$  are nicely tilted against those of  $\langle l_\gamma \rangle$ , and they allow a much better constraint on the profile slope. As these contours are degenerate along the diagonal of the  $a$ - $q$  plane, the profile amplitude can also be constrained by the magnification. This mutual complementation also results in a large gain when using shear and magnification information at the same time. The combination is much more powerful in distinguishing lens models with respect to amplitude and slope than each method individually.
- The investigation of different annuli sizes shows that the result is not dominated by the innermost part of the aperture. Although the contours in the lower panels widen slightly when we increase  $\theta_{\text{in}}$ , this is not a large effect. Especially the contours for  $\langle l_{\text{tot}} \rangle$  remain basically unchanged. On the other hand, increasing  $\theta_{\text{out}}$  significantly improves the constraints for the shear and the magnification.
- The lines of *constant*  $\langle N_\mu \rangle$  shown in the plots give a first indication of the high sensitivity of the magnification method to an exact calibration. They were obtained by fixing  $\langle N_\mu \rangle$  and solving eq. (119) numerically for  $a$  as a function of  $q$ . The first thing to note is that the contours for  $\langle l_\mu \rangle$  follow the curve of the true  $\langle N_\mu \rangle$  very closely. This implies that the signal of the magnification is mainly determined by the total number of galaxies in the aperture. The bad news is that the curves for  $\langle N_\mu \rangle \pm 5\%$  are substantially offset. Especially for the larger annuli in the right panels, this small systematic error leads to a highly significant offset considering the magnification contours. We come back to a more thorough analysis of uncertainties in the number count calibration below.

The only remarkable difference for the Family B model is that the advantage of increasing the outer radius of the annulus is even larger than was the case for Family A. These results are very encouraging, but in their derivation we have assumed that we exactly know the unlensed number counts. Before investigating which effect a relaxation of this strong constraint has, we will check in the next section whether the contours of our ensemble-averaged log-likelihoods estimate the real uncertainties in our analysis reliably.

## 5.7 Error analysis

In this section we want to investigate with Monte Carlo simulations whether the assumption of a  $\chi^2$  distribution for our likelihood errors is valid. Furthermore all our analysis was done for only one realisation of  $a$  and  $q$  so far. We will investigate with Family A whether the merits of our approach depend on this choice. 500 Monte Carlo catalogs with a specific parameter set  $\pi_t$  for the Families A and B were produced and the likelihood analysis to recover the best-fitting parameters  $\pi$  was performed afterwards. In more detail, the generation and analysis of the catalogs consisted of the following steps:

- First, inner and outer radii  $\theta_{\text{in}}$  and  $\theta_{\text{out}}$  were chosen. When distributing the galaxies in the annulus we do not need to work in the full two-dimensional space but we can limit everything to a one dimensional interval of length  $\theta_{\text{out}} - \theta_{\text{in}}$  as only axially symmetric matter models are considered (see section 3.3.2). We started distributing galaxies in number and position

following the distribution of eq. (108). For this we first calculated the expected number of galaxies in the absence of lensing  $\langle N_0 \rangle = n_0 \pi (\theta_{\text{out}}^2 - \theta_{\text{in}}^2)$  and a total number  $N_0$  was drawn randomly from a Poisson distribution with mean  $\langle N_0 \rangle$ . Then, these  $N_0$  galaxies were provisionally distributed in angular radius, by drawing a random number  $\xi$  uniformly distributed in  $[0, 1]$ , and assigning a radius of  $\theta = \sqrt{\xi(\theta_{\text{out}}^2 - \theta_{\text{in}}^2) + \theta_{\text{in}}^2}$ . Then, a second uniform deviate  $\eta \in [0, 1]$  was drawn, and a galaxy at  $\theta$  is put into the final ‘catalog’ only if  $[\mu(\theta)]^{\beta-1} \geq \eta$ ; otherwise it was discarded. With this rejection method, the distribution of galaxies follows the distribution  $n(\vec{\theta}_i) \propto [\mu(\vec{\theta}_i)]^{\beta-1}$ . This finishes the catalog generation with respect to the lens magnification effect.

- For the shear we first drew a random ellipticity for every galaxy according to the distribution eq. (88) with  $\sigma_\epsilon = 0.2$ . This was done using the algorithm described in Press et al. 1992 (see [53]). After that the observed ellipticities were calculated using the exact formulae of eq. (86). No errors have been assigned to the observed ellipticities.
- The catalog contained positions and observed ellipticities for a galaxy number density of  $n_\mu = 120 \text{arcmin}^{-2}$ . All these galaxies were used for the magnification analysis. From this catalog we draw randomly galaxies with a number density  $n_\gamma = 30 \text{arcmin}^{-2}$  for the shear analysis. Afterwards  $l_\mu$ ,  $l_\gamma$  and  $l_{\text{tot}}$  were calculated according to eq. (110), eq. (115) and eq. (116). The minimisation of these functions yielded best fitting parameters for each of these cases.

The distributions of the best fitting parameters obtained are shown for the two families on the left panels in Fig. 13 and Fig. 14, where they are compared with the ensemble-average log-likelihood. We see that the scatter of the recovered values of  $\pi$  is quite consistent with the expectation from the average log-likelihood for  $l_\mu$ ,  $l_\gamma$  and  $l_{\text{tot}}$ . This pure visual impression is quantified in Fig. 15. There, we generated 10000 Monte Carlo realisations with the prescription described above and calculated for every obtained parameter set the quantity  $2\Delta l = 2(l(\pi_{\text{min}}) - l(\pi_t))$ . Then the probability function  $P(> 2\Delta l)$  was calculated and compared with that from a  $\chi^2$  distribution  $P(> \chi_2^2)$  which we used for determining the significance levels of our log-likelihood contours (see Fig. 15). The deviation of  $P(> 2\Delta l)$  from the  $\chi_2^2$  distribution is very small for the magnification method. It is a little larger for the shear and combined likelihoods when going to the higher significance regime but it is still acceptable: the deviation from  $\chi_2^2$  statistics is such that for the recovered values of  $\pi$  about 94.4% of the realisations lie within the 95.4% confidence interval. The reason for the slightly worse behaviour of the shear and combined method is probably the use of the Gaussian approximation for the observed ellipticity distribution in eq. (114).

We also investigated how well our analysis can distinguish between Families A and B. For this we analysed 500 catalogs generated with Family A with the likelihood functions of Family B and vice versa. The result of this is shown in the right panels of Figs. 13 and 14. In a real analysis the model that gives the lower likelihood would be chosen. When generating the catalogs with family A and recovering the parameters with Family B, in 47% of the realisations  $l_\mu(A) < l_\mu(B)$  indicating that the magnification method cannot discriminate between the two models. The middle panel of Fig. 13 shows that the shear method performs slightly better: in 64% of the realisations  $l_\gamma(A) < l_\gamma(B)$ . The results for recovery using the combined method are shown in the bottom panel;  $l_{\text{tot}}(A) < l_{\text{tot}}(B)$  in 63% of the realisations. For generating catalogs with Family B and analysing with Family A the results are qualitatively the same. Also here the shear performs slightly better in discriminating the models: in 57% of the realisations,  $l_\mu(B) < l_\mu(A)$  and in 73% of the realisations  $l_\gamma(B) < l_\gamma(A)$ . The combined method gives  $l_{\text{tot}}(B) < l_{\text{tot}}(A)$  in 78% of the realisations. Although the discrimination of the two models is quite poor, the slope  $q$  is in good agreement no matter what Family we use for the analysis.

Considering Figs. 11 and 13 we note that in the second case where the true model has a steeper and stronger mass profile the shear performs considerably better in determining the slope of the profile. Hence, the dependence of the results from likelihood analysis on the true values  $a$  and  $q$  is worth investigating. Instead of calculating many average log-likelihoods and looking at the resulting contours, we have used the following approach: We consider a Taylor expansion of the

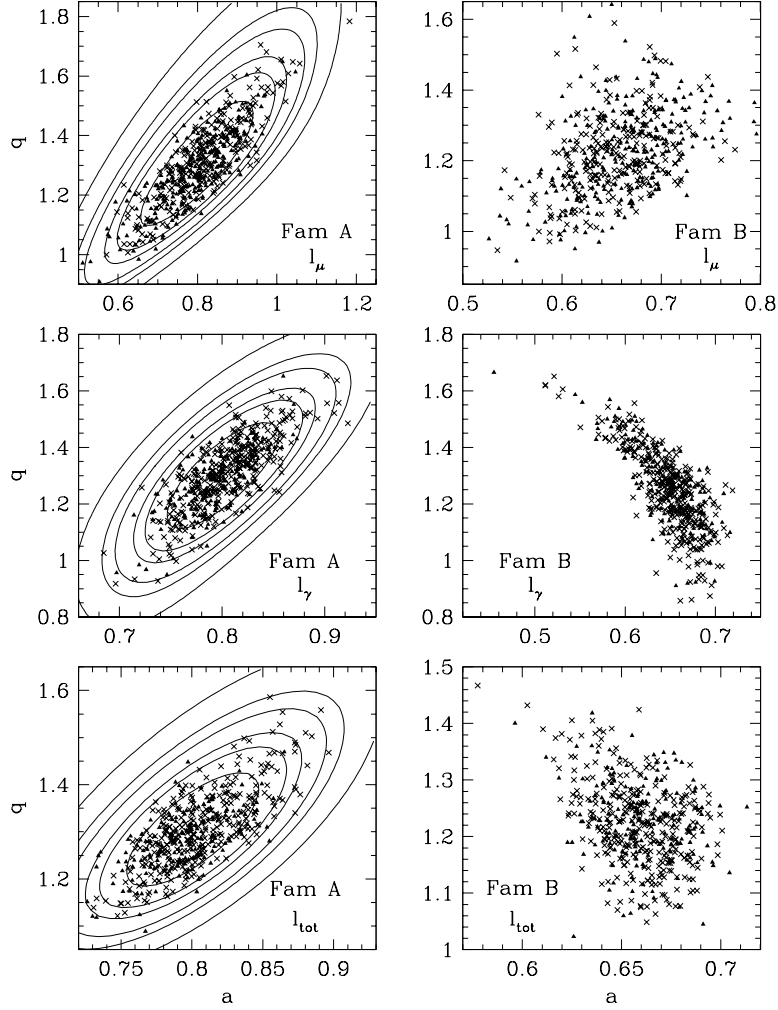


Figure 13: Data generation with Family A, analysis with Families A and B (Figure 3 from [61]): A Family A input model ( $a = 0.8$ ,  $q = 1.3$ ) was used to generate lensed galaxy catalogs. For each of the magnification (top panel), shear (middle panel) and combined methods (bottom panel) the left-hand (right-hand) panel shows parameter recovery under a Family A (Family B) model and the crosses (triangles) indicate where the log-likelihood function is lowest when the minimisation is performed using a Family A (Family B) lens model. On the left-hand panels we mark contours of constant  $\Delta \langle l \rangle$  with levels the same as in Fig. 11.

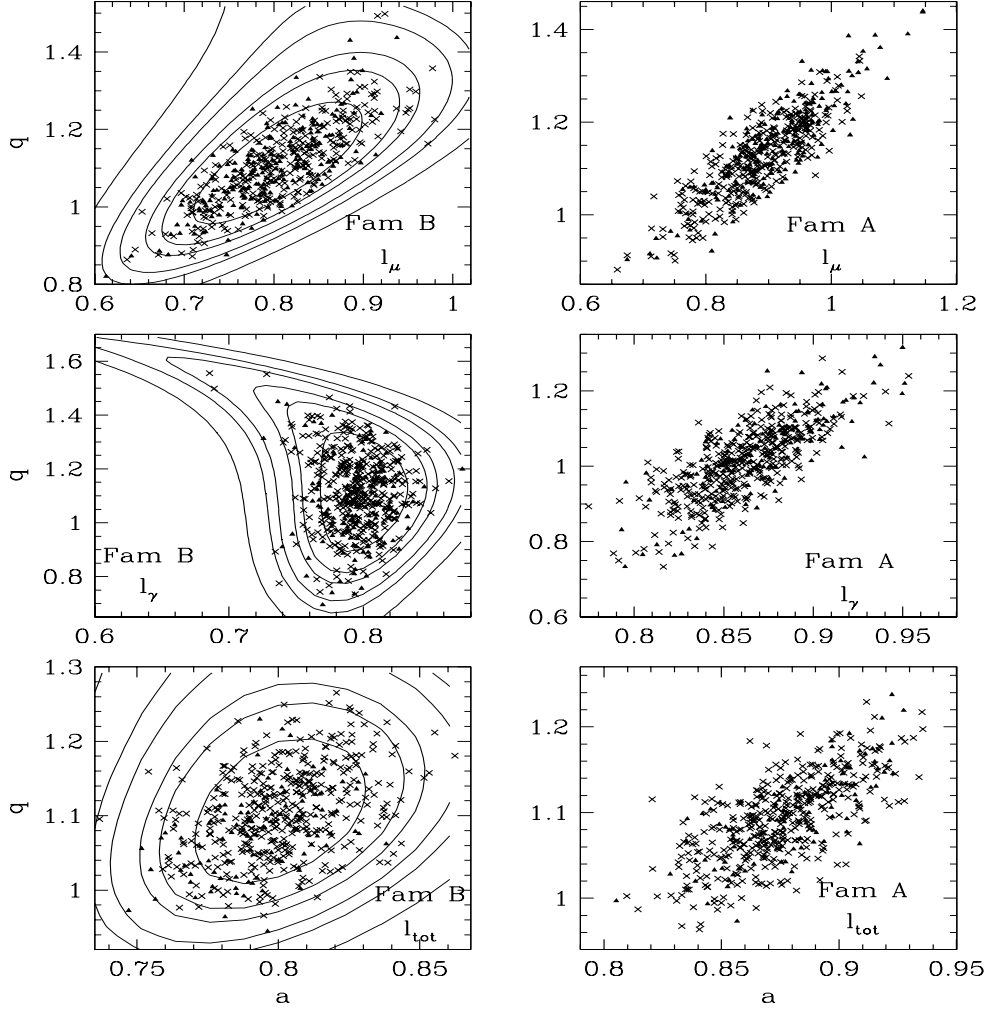


Figure 14: Data generation with Family B, analysis with Families A and B (Figure 4 from [61]): A Family B input model ( $a = 0.8$ ,  $q = 1.1$ ) was used to generate lensed galaxy catalogs. For each of the magnification (top panel), shear (middle panel) and combined methods (bottom panel) the left-hand (right-hand) panel shows parameter recovery under a Family B (Family A) model and the crosses (triangles) indicate where the log-likelihood function is lowest when the minimisation is performed using a Family B (Family A) model. On the left-hand panels we mark contours of constant  $\Delta \langle l \rangle$  with the same levels as in Fig. 11.

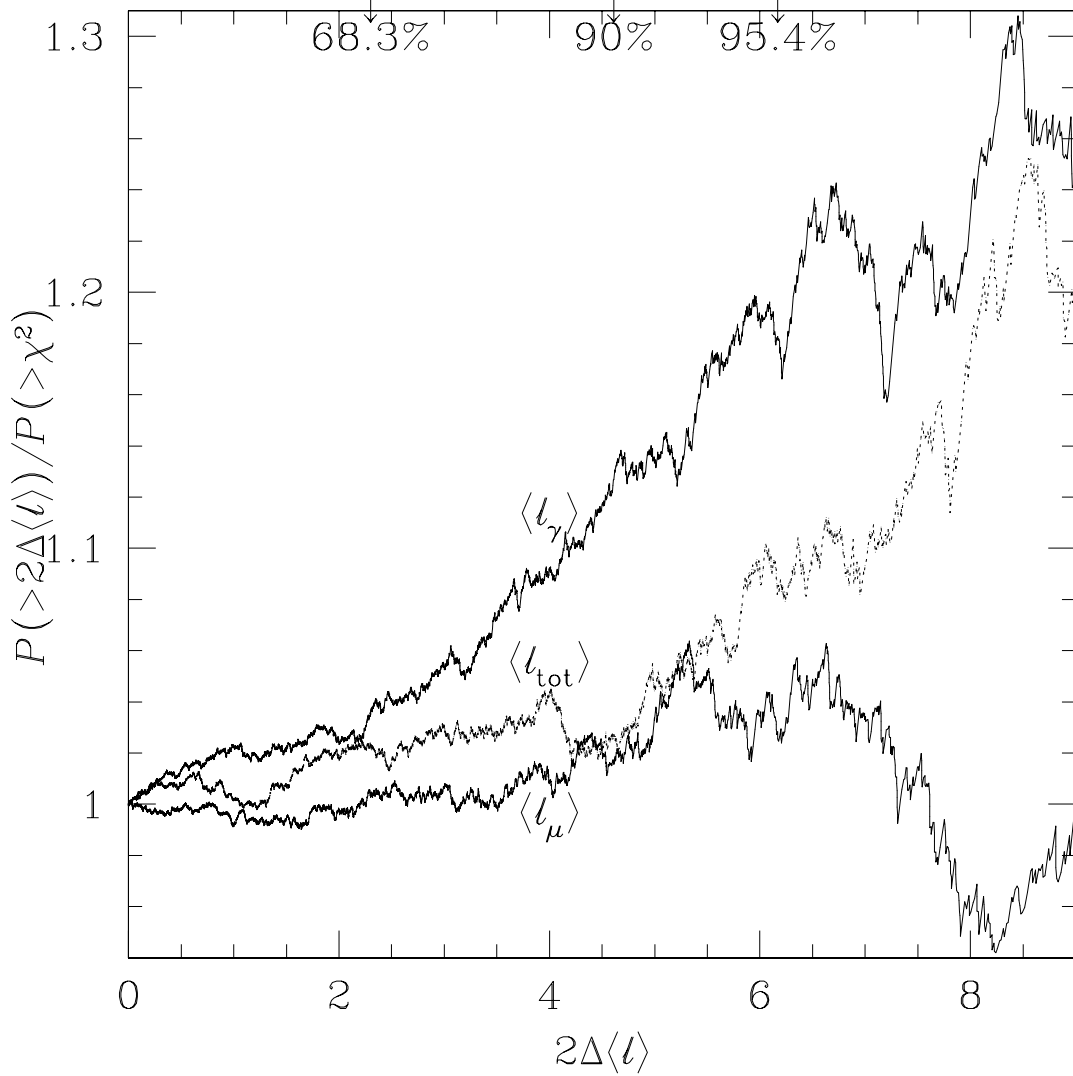


Figure 15: Comparison of the error levels from our average log-likelihood functions with those from the  $\chi^2$  statistics (Figure 5 from [61]): For the magnification (lower curve), shear (upper curve) and combined methods (dashed middle curve), we show the ratio of  $P(> 2\Delta \langle l \rangle)$  to that expected if  $2\Delta \langle l \rangle$  followed a  $\chi^2$  distribution versus  $2\Delta \langle l \rangle$ . On the top edge of the plot the 68.3%, 90%- and 95.4%-confidence intervals are marked.

average log-likelihood

$$\langle l \rangle (\pi) \approx \langle l \rangle (\pi_t) + \frac{1}{2} \frac{\partial^2 \langle l \rangle}{\partial \pi_i \partial \pi_j} (\Delta \pi)_i (\Delta \pi)_j \quad (130)$$

The first derivative of  $\langle l \rangle$  around the minimum at  $\pi_t$  vanishes by definition. The matrix  $V_{ij}$  of second derivatives defines an ellipse around the true values. We take the  $1\sigma$  error ellipse described by  $2\Delta l = 2.30 =: X$  to determine the error bars for  $a$  and  $q$ . The matrix elements  $V_{ij}$  in the magnification case read:

$$V_{ij}^\mu := \frac{\partial^2 \langle l_\mu \rangle}{\partial \pi_i \partial \pi_j} \Big|_{\pi=\pi_t} = (1-\beta)^2 n_\mu \times \int d^2\theta [\mu_t(\vec{\theta})]^{\beta-3} \left( \frac{\partial \mu(\vec{\theta})}{\partial \pi_i} \frac{\partial \mu(\vec{\theta})}{\partial \pi_j} \right) \Big|_{\pi=\pi_t}. \quad (131)$$

In the shear case the analytic calculation turned out to be very cumbersome and the matrix elements were calculated numerically. Next we construct the smallest rectangle surrounding the ellipse and take the half-lengths of its side as the errors. Although we see from Fig. 11 that the description of the contours as ellipses may not be optimal, this simplified prescription should suffice to provide a qualitative handle on the expected uncertainties. Also in the case of an ellipse having its major axis diagonal in the  $a$ - $q$  plane, the errors are likely to be overestimated. With this prescription we obtain for the errors  $\Delta a$  and  $\Delta q$  (after some tedious analysis)

$$\begin{aligned} \frac{\Delta a}{\sqrt{X}} &= \sqrt{\frac{\det V}{V_{11}^2 V_{22}}} + \frac{V_{12}^2}{V_{11}} \sqrt{\frac{1}{V_{22} \det V}}, \\ \frac{\Delta q}{\sqrt{X}} &= \sqrt{\frac{\det V}{V_{11} V_{22}^2}} + \frac{V_{12}^2}{V_{22}} \sqrt{\frac{1}{V_{11} \det V}}. \end{aligned} \quad (132)$$

The errors determined in that way are plotted in Fig. 16, for two choices of the radii of the annulus in which observations are assumed to be available; the small (large) outer radius corresponds to the left (right) panels in Fig. 16. They are plotted as a function of the value of  $a$ , for five different values of  $q$ , and using the shear (solid curves), magnification (dotted curves) and the combination of both methods (dashed curves). The most obvious point to note is that for the smaller outer radius of the annulus, the magnification method yields smaller errors in  $q$  than the shear method, for all values of  $a$  and  $q$  (see lower left panel), whereas for the error in the amplitude  $a$ , the relative merits of both methods depend on  $a$  and  $q$ . For the larger outer aperture radius, the amplitude of the mass distribution is much better determined from the shear method, and the relative accuracy of the slope determination depends mainly on  $q$ : for steep profiles,  $q \geq 1$ , the shear method yields more accurate results, whereas the magnification method is superior for flatter profiles. Except for the shear method, the accuracy increases with decreasing  $q$  and increasing  $a$ , or in other words, with increasing lens strength.

After we have convinced ourselves that our average log-likelihood functions give the correct picture of the errors involved, we use this formalism now to study the consequences of uncertainties in the unlensed number counts in more detail.

## 5.8 The effect of uncertainties in the unlensed number counts

At the end of section 5.6 we have discussed that to first order the magnification information is provided by the total number of galaxies in our annulus. We noticed that even small systematic errors in the calibration of the galaxy number counts can lead to very significant errors when estimating parameters with the magnification. In this section we want to investigate statistical rather than systematic uncertainties in the number counts. We assume that we know the true number density  $\bar{n}_\mu$  with a fractional inaccuracy  $\eta$ . Before investigating the general case we have a closer look at the case  $\eta = \infty$ . This means that we do not have any calibration for the number counts at all but we infer them from the data themselves. It is interesting to see whether in this case, the magnification method still can improve the parameter estimation over using the

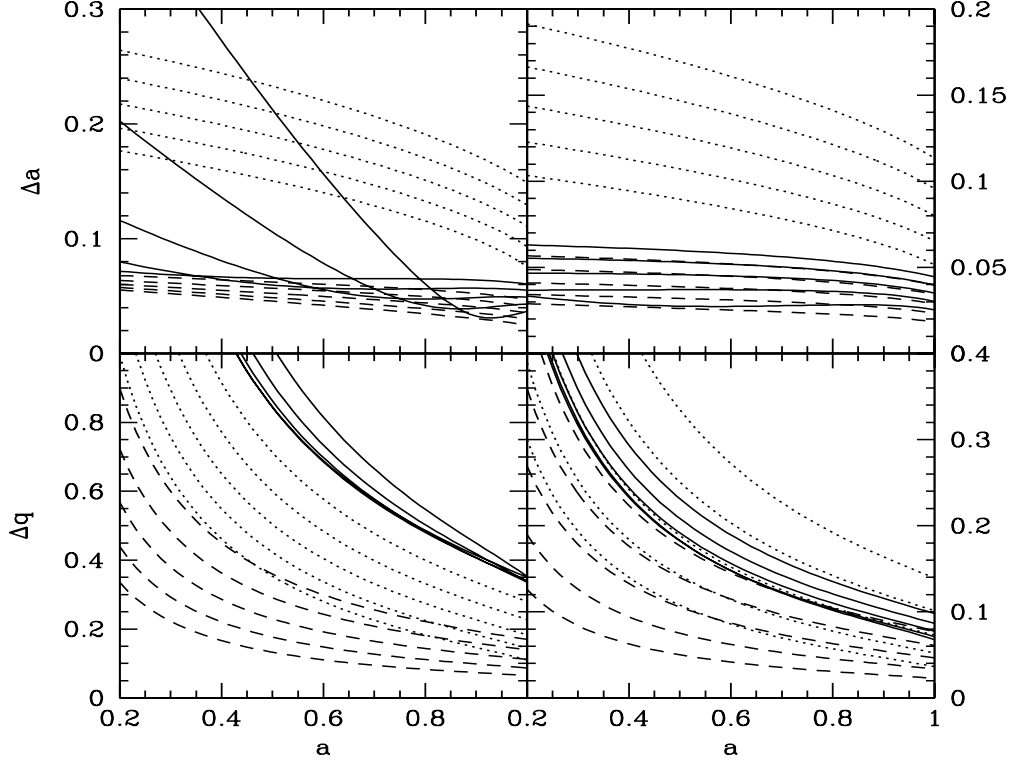


Figure 16: The errors  $\Delta a$  and  $\Delta q$  from the “error ellipse” in eq. (132) (Figure 4 from [61]): The errors are plotted as a function of the input value of  $a$ . The lens model Family A is considered here. The left panels are for  $\theta_{\text{in}} = 0.6'$ ,  $\theta_{\text{out}} = 2.0'$ , and the right panels for  $\theta_{\text{in}} = 0.6'$ ,  $\theta_{\text{out}} = 6.0'$ . Each panel shows three sets of five curves; the solid curves correspond to the shear method, i.e., are based on  $\langle l_\gamma \rangle$ , the dotted curves correspond to the magnification method, and the dashed curves are the errors from using the combination of shear and magnification information. Curves within each set correspond to different input values of  $q$ , with  $q = 0.6, 0.8, 1.0, 1.2, 1.4$ . For  $\langle l_\mu \rangle$  and  $\langle l_{\text{tot}} \rangle$ , the curves vary monotonically with  $q$ , the lowest corresponding to  $q = 0.6$ , the highest to  $q = 1.4$ . The dependence of the errors for  $\langle l_\gamma \rangle$  on  $q$  is more complicated and not necessarily monotonic; for small  $a$ , the uppermost curves in the left panels correspond to  $q = 0.6$ , for the upper right panel, the behaviour is monotonic, with  $\Delta a$  increasing with  $q$ . For the lower right panel, the two largest values of  $\Delta q$  correspond to the two largest values of  $q$ ; the curves for the other values of  $q$  are almost coincident.



shear alone. We will perform these investigations again with the average log-likelihood functions. Returning to eq. (110), and reinserting the density  $n_\mu$  yields

$$l_\mu = n_\mu I + (1 - \beta) \sum_{i=1}^{N_\mu} \ln \mu(\vec{\theta}_i) - N_\mu \ln n_\mu, \quad (133)$$

with

$$I := \int d^2\theta [\mu(\vec{\theta})]^{\beta-1}. \quad (134)$$

Minimising eq. (133) with respect to the unknown number density  $n_\mu$  yields the intuitive result

$$n_\mu = \frac{N_\mu}{I}. \quad (135)$$

Up to an unimportant additive constant we get for the new log-likelihood  $\hat{l}_\mu$

$$\hat{l}_\mu = (1 - \beta) \sum_{i=1}^{N_\mu} \ln \mu(\vec{\theta}_i) + N_\mu \ln I. \quad (136)$$

Following the techniques applied in section 5.4, the ensemble average of  $\hat{l}_\mu$  can be calculated,

$$\begin{aligned} \langle \hat{l}_\mu \rangle &= \bar{n}_\mu \left( \int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1} \right) \ln I \\ &+ \bar{n}_\mu (1 - \beta) \int d^2\theta [\mu_t(\vec{\theta})]^{\beta-1} \ln \mu(\vec{\theta}), \end{aligned} \quad (137)$$

where now  $\bar{n}_\mu$  is the true number density of galaxies. In Fig. 17 we have plotted contours of constant  $\langle \hat{l}_\mu \rangle$ , for two values of the outer radius of the annulus, together with the corresponding contours of  $\langle l_\mu \rangle$ ,  $\langle l_\gamma \rangle$ , and  $\langle \hat{l}_{\text{tot}} \rangle := \langle l_\gamma + \hat{l}_\mu \rangle$ . From the figure it becomes immediately clear that dropping the assumption about the knowledge of the unlensed number density leads to a drastic loss of information from the magnification method. The thin dashed contours in Fig. 17 which correspond to  $\langle \hat{l}_\mu \rangle$  are much wider than the solid contours corresponding to  $\langle l_\mu \rangle$ . In addition, whereas the contours of  $\langle l_\mu \rangle$  were seen to be substantially misaligned relative to those of  $\langle l_\gamma \rangle$ , so that the combination of shear and magnification information substantially decreases the error region in parameter space relative to any individual one of these methods, the contours of  $\langle \hat{l}_\mu \rangle$  are basically parallel to those of  $\langle l_\gamma \rangle$  (shown as heavy dashed curves in Fig. 17). Therefore, the combination of both methods, in the absence of knowledge of the unlensed number density, yields only slightly more accurate parameter estimates than the shear method alone.

We now come to the more realistic case where we know the unlensed number counts approximately up to a fractional accuracy  $\eta$ . One can then supplement the likelihood function with a prior, taken to be a Gaussian in  $n_\mu$ , with mean  $\bar{n}_\mu$ , and dispersion  $\sigma_n = \eta \bar{n}_\mu$ . Hence, the corresponding log-likelihood function becomes

$$\begin{aligned} l_\mu &= n_\mu I + (1 - \beta) \sum_{i=1}^{N_\mu} \ln \mu(\vec{\theta}_i) \\ &- N_\mu \ln n_\mu + \left( \frac{(n_\mu - \bar{n}_\mu)^2}{2(\eta \bar{n}_\mu)^2} \right). \end{aligned} \quad (138)$$

As before, this can be minimised with respect to  $n_\mu$ , yielding

$$\frac{n_\mu}{\bar{n}_\mu} = \frac{1}{2} (1 - \eta^2 \bar{n}_\mu I) + \sqrt{\frac{1}{4} (1 - \eta^2 \bar{n}_\mu I)^2 + \eta^2 N_\mu}. \quad (139)$$

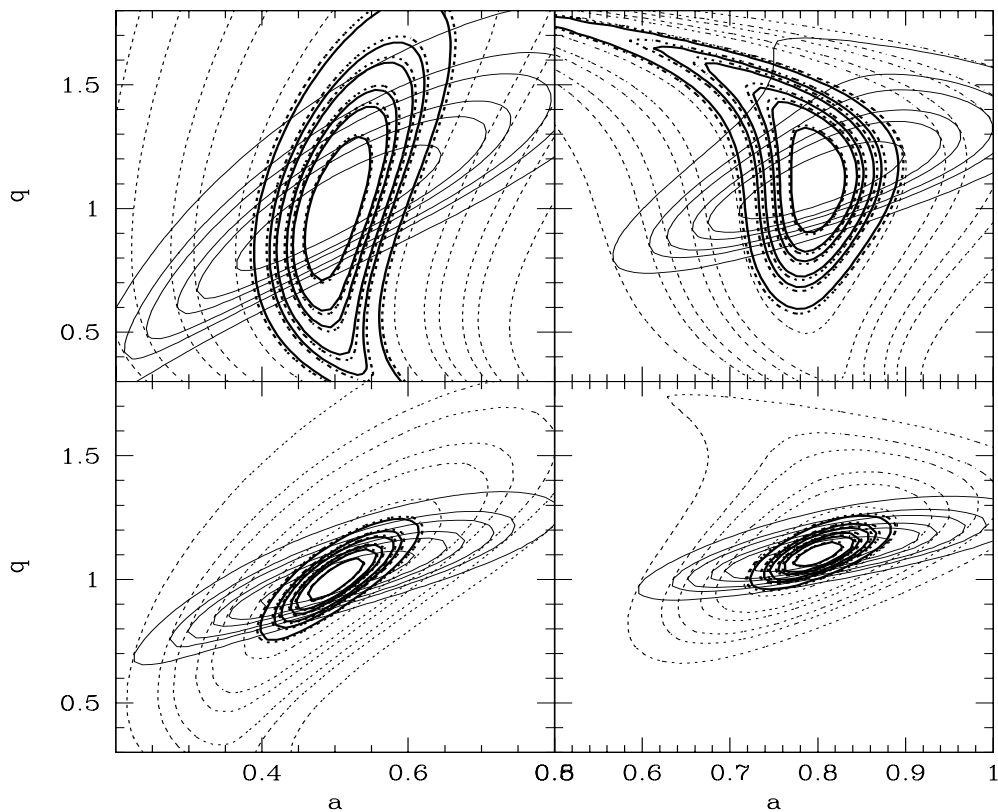


Figure 17: Parameter estimates with no calibration for the unlensed galaxy number counts (Figure 8 from [61]): Similarly to figures 11 and 12, contours of constant  $\Delta \langle l \rangle$  are plotted; the contour levels are the same as described in Fig. 11. The inner radius of the annulus is  $\theta_{\text{in}} = 0.6'$ , the outer radius is  $\theta_{\text{out}} = 4'$  (upper panels) and  $15'$  (lower panels); left (right) panels correspond to models of Family A (Family B). The thin dashed contour correspond to  $\langle \hat{l}_\mu \rangle$ , the thin solid contours to  $\langle l_\mu \rangle$ , the heavy dashed contours to  $\langle l_\gamma \rangle$ , and the heavy solid contours to  $\langle \hat{l}_{\text{tot}} \rangle := \langle l_\gamma + \hat{l}_\mu \rangle$ . The input model is  $a = 0.5$ ,  $q = 1.0$  for Family A, and  $a = 0.8$ ,  $q = 1.1$  for Family B

Inserting this value for  $n_\mu$  into eq. (138) yields the new log-likelihood function, which we shall denote as  $l_\mu^{(\eta)}$ . Taking the ensemble average in the present case is slightly more difficult, owing to the occurrence of  $N_\mu$  in the square-root. However, if we neglect the Poisson-fluctuation in the ensemble averaging in eq. (117), which is a very good approximation when  $\langle N_\mu \rangle_t$  is large, the ensemble average can be obtained as

$$\begin{aligned} \langle l_\mu^{(\eta)} \rangle &= \langle n_\mu \rangle I + \bar{n}_\mu (1 - \beta) \int d^2\theta \mu_t^{\beta-1} \ln \mu \\ &- \langle N_\mu \rangle_t \ln \langle n_\mu \rangle + \frac{(\langle n_\mu \rangle - \bar{n}_\mu)^2}{2(\eta \bar{n}_\mu)^2}, \end{aligned} \quad (140)$$

where

$$\frac{\langle n_\mu \rangle}{\bar{n}_\mu} = \frac{1}{2} (1 - \eta^2 \bar{n}_\mu I) + \sqrt{\frac{1}{4} (1 - \eta^2 \bar{n}_\mu I)^2 + \eta^2 \langle N_\mu \rangle_t}. \quad (141)$$

For  $\eta \rightarrow 0$ , this log-likelihood reduces, up to additive constants, to the one for which the unlensed number density was assumed to be known, i.e., eq. (108), whereas for large  $\eta$ , it approaches eq. (137). In Fig.18 we have plotted the 90%-confidence contours as obtained from  $\langle l_\mu^{(\eta)} \rangle$ , using models of Family A, for various outer radii of the data annulus. Four values of the fractional uncertainty  $\eta$  in the number density have been chosen, ranging from  $\eta = 0$  – for which the contours agree with the corresponding ones in Fig. 11 – to  $\eta = 0.06$ . In addition, the 90%-confidence contour for the shear and those for the combined method are plotted. As expected, by increasing  $\eta$ , the confidence region increases, and the relative increase is larger, the larger  $\theta_{\text{out}}$ . The reason for this is that for the small annulus, the relative change of the number density by magnification is much larger than the fractional uncertainty, whereas this is no longer true at larger annuli. In fact, as can be seen from the lower right panel in Fig. 18, the confidence contours seem to approach an asymptotic form quickly for large annuli, implying that the prior information becomes irrelevant unless  $\eta$  is very small, since the unlensed number density is then more accurately determined from the data set itself. Thus, the outer-most contour in the lower-right panel in Fig. 18 is very similar to the corresponding one in Fig. 17 where no prior information about the number density was assumed. Thus, the larger the region from which data are used to determine the mass profile parameters, the more accurate the unlensed number density needs to be known, for a given accuracy relative to that obtainable from the shear method. In particular, we conclude that the magnification method considered here seems to be most powerful for data sets which do not extend far from the cluster centre, whereas for wide-field imaging data, the shear method is superior.

## 5.9 Conclusions

We have developed a method to optimally combine shear and magnification information to constrain parameters of cluster mass models. The method is based on maximum likelihood techniques and on realistic prior knowledge. For massive lensing clusters, the Einstein radius is a direct observable through the location of giant arcs as well as the cluster redshift. The assumption that all background galaxies are located at the same redshift is appropriate as long as we deal with low redshift ( $z \leq 0.2$ ) lenses.

It was shown that, although the shear is easier to detect, the magnification in general complements the distortion measurement and it helps to constrain the slope of a cluster profile. The relative merits of shear and magnification depend on the combination of mass amplitude and slope though. For clusters with a higher lens strength the shear does a better job in constraining the slope than for less massive objects.

As already discussed in the introduction of this section the value of the magnification method completely depends on the accuracy with which the unlensed number counts of galaxies can be obtained. If no calibration at all can be obtained the magnification information gets lost completely and already a 10% statistical uncertainty in the number counts widen the magnification contours

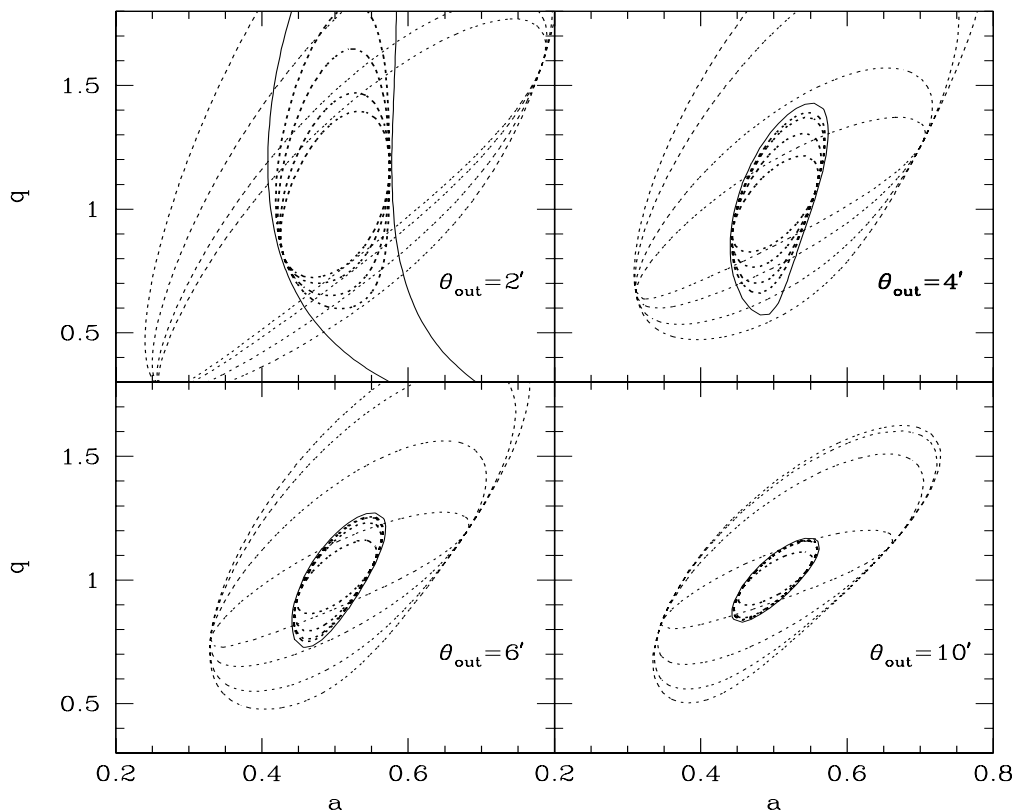


Figure 18: Parameter estimates where the unlensed number counts are known with a fractional accuracy  $\eta$  (Figure 9 from [61]): The light dashed contours plot the 90%-confidence contours of parameter estimates as obtained from the ensemble-averaged log-likelihood function in eq. (140). The model  $a = 1/2$ ,  $q = 1$  of Family A has been taken as input. The inner radius of the annulus in which data are assumed to be is  $\theta_{\text{in}} = 0.6'$ , and the outer radius is different in each panel. The solid contour is the 90%-confidence contour as obtained from the shear, and the heavy dashed curves from combining shear and magnification constraints. The sets of four dashed curves are for  $\eta = 0, 0.02, 0.04$ , and  $0.06$ .

significantly. Besides the observational difficulty to obtain this accuracy, the angular correlation function provides a fundamental limit to this measurement. Villumsen et al. 1997 (see [69]) have measured the angular correlation function to faint magnitudes in the Hubble-Deep-Field North on an arcsecond scale. If one extrapolates their power law to larger radii, an uncertainty of 10% in the number counts can be expected on a scale of  $4'$ . If this should turn out to be the case, the use of the magnification method would be very limited, indeed.

We note that the presented formalism has been used by Gray et al. 2000 (see [23]) to detect a significant magnification signal behind the cluster Abell 2219.

## 6 Technical aspects of weak lensing observations

### Summary:

*When introducing the concepts of weak gravitational lensing in section 4.2 and in the theoretical work of the last section it was assumed that we can measure, without technical difficulty, the moments of galaxy light profiles leading to ellipticities and finally to a shear estimate. In this section we introduce the problems that one encounters with these measurements in practice and how to deal with them. We have performed a large number of simulations to get confidence in our shear measurements and consequently in the scientific results that we will present in subsequent sections. The work was published in [16].*



Figure 19: A section from a CCD image

### 6.1 Practical problems for weak lensing studies

In the following we limit the discussion to observations done on optical telescopes equipped with a CCD camera. Single chip CCD cameras cover a relatively small solid angle of the sky (typically around  $5'$  to  $10'$ ) in a two-dimensional mosaic of pixels. A typical dimension for professional CCD cameras is  $2048 \times 2048$  pixels which leads to a resolution of about  $0.2'' - 0.3''$  for a single pixel. A small piece of such a pixelised CCD image is shown in Fig. 19. This setup first brings the following problems:

- As a consequence of the pixelisation our sources can only be observed in a discrete form, and all the properties we want to determine have to be inferred from the pixel values. All the integrals that we need for estimating moments of the light distribution [eq. (77) and eq. (78)] can easily be transformed into sums over discrete values so that pixelisation is not a

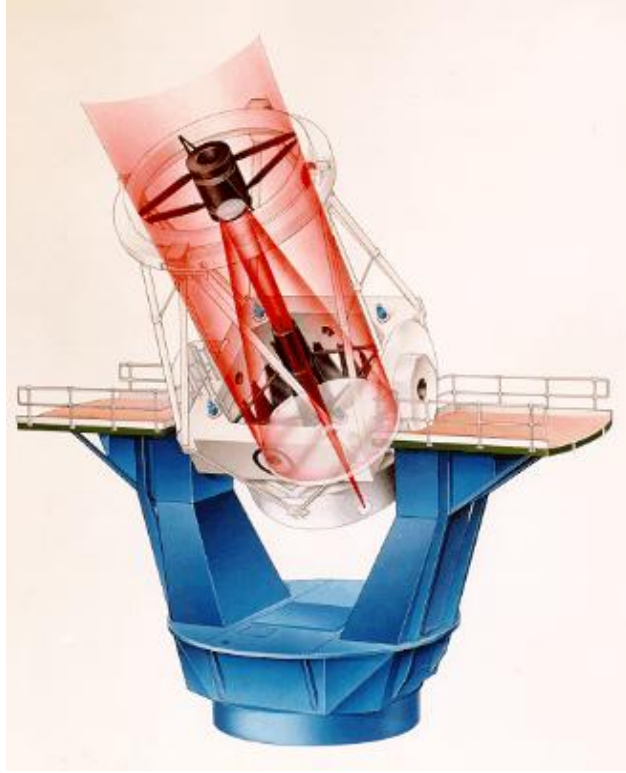


Figure 20: The 4.2m William Herschel Telescope (WHT) on La Palma (Figure obtained from <http://ing.iac.es/orm/wcray.jpg>): The light path through the optics of the telescope is shown. It comes through the entrance aperture, is deflected at the primary mirror to the secondary mirror and from there focussed to the instrument. The light paths are changed by diffractions at the telescope equipment. Note the four supports of the secondary mirror that directly cross the light path at the entrance aperture.

problem in this technical aspect. More important is that the pixel resolution may be not high enough for small, faint objects that can easily be detected but whose information is contained in only a few pixels so that accurate moment measurement cannot be performed. The only way here is to exclude such objects from the analysis.

- The integrals from eq. (77) and eq. (78) are formally integrals over the whole real plane. In practice we have to limit the calculation to a narrow region around detected objects. If we extend the calculation too far we may include the light from nearby objects and more importantly, our results are dominated by noise from the background light of the night sky at the borders of our objects. Two ways have been proposed to solve this problem. The first is to stop the integration at a fixed *isophotal threshold*, that is only to include object pixels that have a value of at least  $t \times \sigma_{\text{sky}}$ , where  $\sigma_{\text{sky}}$  is the variance of the background sky noise and  $t$  is a variable parameter. The second method is to include a weighting function into the moment calculation

$$\int d^2\theta \dots I(\vec{\theta}) \Rightarrow \int d^2\theta \dots W(\vec{\theta}) I(\vec{\theta}), \quad (142)$$

where  $W(\vec{\theta})$  is typically a circularly symmetric function, for instance a Gaussian. The large disadvantage of this weighting approach for us is that the simple relations between lensed and unlensed brightness moments given in eq. (84) are no longer valid. Nevertheless this is the method of choice for lensing studies nowadays and we will see below how lensing quantities can be inferred from these weighted moments.

- In Fig. 19 we note that sources are typically not isolated but isophotes of many objects overlap. This also severely influences measurements of object shapes, and also here the only or, at least, the safest solution is to reject images who have a too close neighbour.

Besides the above mentioned technical difficulties to select objects appropriately and to perform the necessary integrations, weak lensing investigations suffer from the, by far more severe, *point spread function* effects. Under this term we want to summarise all effects that change the shape of our objects in addition to the gravitational lensing effect we want to observe. These additional distortions need a careful consideration in the course of our analysis as they can easily mimic a shear signal.

- Light passing through the optical system of a telescope experiences diffractions changing the optical path lengths in some parts of the system. The main reasons for these diffractions are the entrance aperture, the edges of the mirrors and the supports of the mirrors. A layout of a modern telescope is shown in Fig. 20. Consequently, light coming from a point source (a star) is not focussed perfectly to a sharp point-like image in a single CCD pixel, but it results in a structured image. We call these images of a point-like source the point spread function (PSF). If there are no further sources influencing the light path from astronomical objects, we say the PSF is *diffraction limited*. These are the best observing conditions (in terms of PSF properties) we can get, and they only can be obtained in space-based observations. Although a diffraction-limited PSF can have a complex structure, it is usually still very sharp. For the Wide Field Planetary Camera (WFPC) onboard the Hubble Space Telescope more than half of the light of a stellar object is contained in one pixel (resolution: 0.1'' per pixel).
- For observations from the ground, the PSF is no longer dominated by the diffraction of the telescope's optical system, but light paths are mainly distorted by turbulent motion in the Earth's atmosphere. In this case we have *turbulence limited seeing*. Images from point-like sources are blurred and, in contrast to diffraction-limited seeing, images of stars are no longer very sharp. Ground-based observations under very good conditions lead to a *seeing* of about 0.7''. This seeing is the FWHM of a Gaussian fit to a stellar object. Thus, for a typical pixel scale of 0.2'' half of the PSF light is contained in a circular aperture with diameter  $\approx 3$  pixel. The effect is dependent on the wavelength of the incoming light. Blue light is scattered more, and lensing studies from the ground, depending on the best image quality possible, are done mostly in the red and near infrared.
- Moreover, the optics of most telescopes are not perfect. Possible defects that change the path of incoming light rays are imperfections in some mirror (deviations from a perfect parabolic surface for instance) or problems in the setup of the optical system (for instance a slight tilt of mirrors with respect to the optical axis). See Fig. 21. The PSF patterns introduced by these defects are usually summarised under the term *optical aberrations*.

Figure 22 shows a bright star in a CCD image. Also the light rays from galaxies, whose shape we want to measure, are influenced in the same way by telescope optics and the Earth atmosphere. If we refer to the observed light profile of a star as  $P$ , the intrinsic profile of a galaxy  $I$  is changed to the observed profile  $I^{\text{obs}}$  by a convolution

$$I^{\text{obs}}(\vec{\theta}) = \int d^2\vartheta I(\vec{\vartheta})P(\vec{\theta} - \vec{\vartheta}). \quad (143)$$

We assume for the following that all non-lensing shape changes in a galaxy profile can be described by this convolution, and we neglect other effects like optical distortions that would have to be described by a mapping. On the one hand the convolution leads to a circularisation of the galaxy image and the measured modulus of the galaxy ellipticity becomes smaller (see Fig. 23). On the other hand, if the PSF is anisotropic, it can introduce an artificial shear signal to our data. Figure 24 shows ellipticity measurements on stars from a globular cluster image taken at the



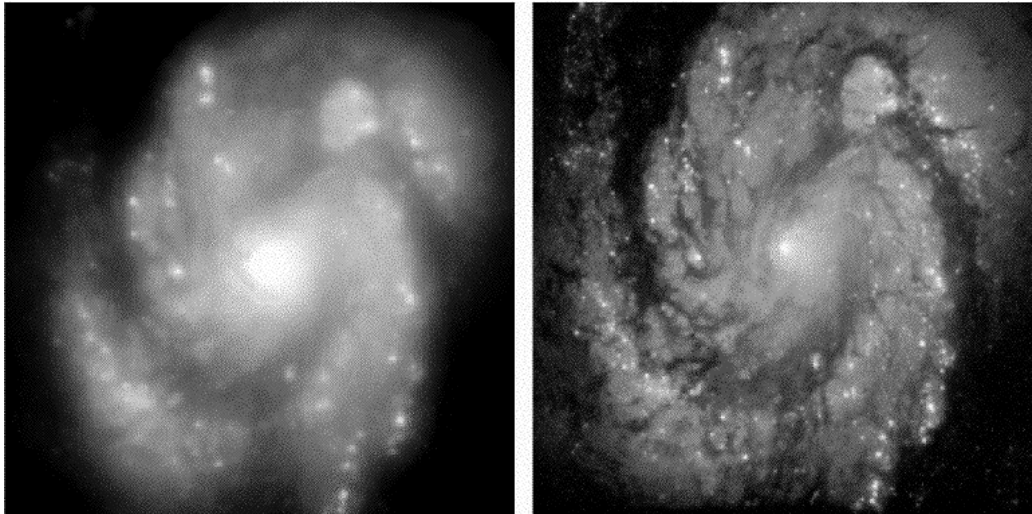


Figure 21: Optical aberrations in the The Hubble Space Telescope (HST) (Figure obtained from <http://oposite.stsci.edu/pubinfo/jpeg/WFPCM100Comp2.jpg>): The HST suffered of *spherical aberration* in the first three years of its operation (1990-1993). The parabolic form of the primary mirror was imperfect and so light reflected from it was not exactly focussed to a single point. The left image of the galaxy was taken a few days before and the right a few days after the aberration had been corrected for.

New Technology Telescope (NTT) in 1997. From this figure we see the following important characteristics for a PSF anisotropy on ground-based observations:

- The stars in this field have an ellipticity up to 10%. This value is high but not untypical for ground based observations. This could easily introduce a spurious shear signal of a few percent on extended galaxies.
- The PSF anisotropy shows large variations over the field. Fortunately, the variations are smooth, and the ellipticity structure can easily be modelled by a low-order polynomial fit. This fact will be a very important ingredient when we discuss techniques to correct for PSF effects later. It enables us to estimate the shape of the PSF everywhere on the image.

We note that the PSF anisotropy structure usually is not stable from one exposure to the next. We expect variations with varying observing conditions, like airmass, a poor tracking of the Earth rotation during a long exposure, or observing with a not ideally focussed telescope. The PSF anisotropy also depends strongly on the seeing. In very good seeing conditions, when the PSF starts to become dominated by diffraction effects, the strongest anisotropies with large variations over the field are observed (see Fig. 25 for anisotropies of the diffraction-limited HST PSF). On the other hand, a bad seeing (1.5'' or worse) leads to a strong smearing of the stellar images without any anisotropy.

The above discussion makes obvious the need for carefully taking PSF effects into account when we want to measure galaxy shapes reliably. Otherwise, a possible measured signal may be the result of systematics due to uncorrected PSF effects. In the early days of weak lensing observations, the effects of a PSF on shear measurements were mainly calibrated with simulations.

One strategy to correct for the circularisation by the PSF was to investigate by which factor a shear is underestimated. Then, all the shear estimates from individual galaxies were multiplied by this factor. This approach is not optimal since, the results of this “fudge” factor depend on the simulations (details of the PSF etc.), and it is obvious that not all galaxies are influenced by a PSF in the same way. Small galaxies are affected stronger by a PSF than large ones. Therefore, an algorithm that allows the calibration from the observed data itself would be highly desirable.

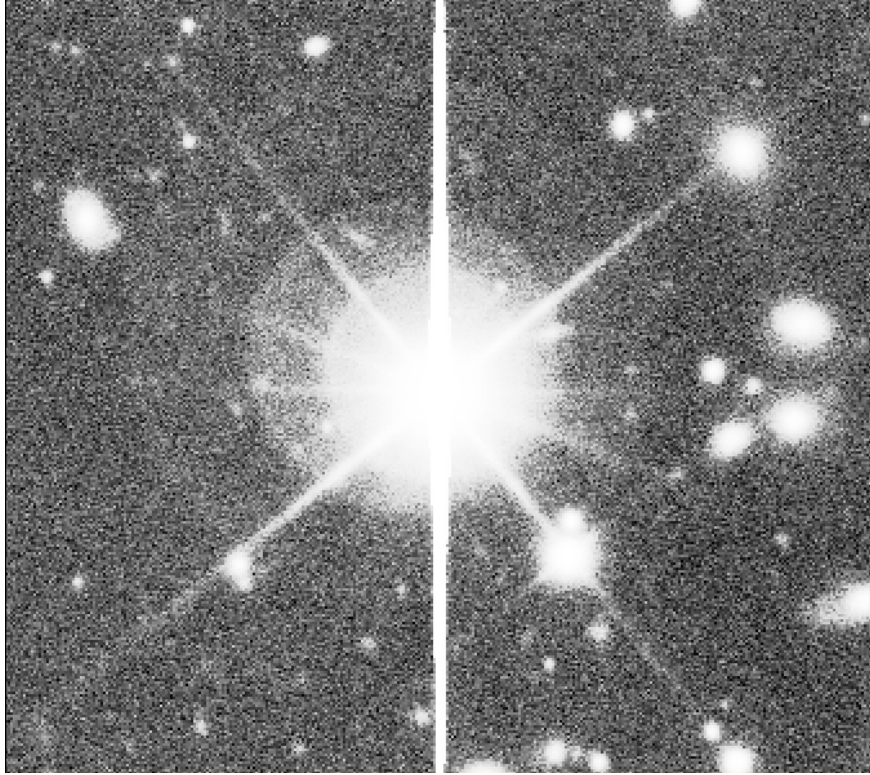


Figure 22: The image of a bright star: We note four spider arms that originate from light diffraction on the supports of the secondary mirror. Besides this structural feature, bright stars also produce *ghost images* (the haloes surrounding the star) and saturated columns (the bright, thick line across the image). One not only has to correct galaxy images for PSF effects, but also to mask out these defective areas of an image.

In the next subsection we introduce the so-called *KSB* algorithm (see [34]) for shear measurements which has this merit.

## 6.2 The KSB algorithm

In 1995 and 1997 Nick Kaiser and his colleagues [Kaiser, Squires & Broadhurst 1995 (KSB henceforth; see [34]) and Luppino & Kaiser 1997 (see [43])] showed how the reduced shear  $\mathbf{g}$  (and hence the shear  $\gamma$  for  $|\gamma| \ll 1$ ,  $\kappa \ll 1$ ) can be inferred from weighted brightness moments of the light distribution and the observed galaxy ellipticity  $\chi$  from eq. (79), i.e. they considered eq. (79)\*, where the  $Q_{ij}$  were substituted by  $Q_{ij}^W$

$$Q_{ij}^W = \int d^2\theta \theta_i \theta_j I(\vec{\theta}) W(|\vec{\theta}|), \quad (144)$$

whereby the object centre is assumed to lie at the origin of the coordinate system. It is estimated with the weighted first moment:

$$\int d^2\theta \vec{\theta} I(\vec{\theta}) W(|\vec{\theta}|) = 0. \quad (145)$$

Moreover, they considered the influence of a PSF on the light profile. For the inclusion of the PSF it was assumed that any PSF  $P$  can be written as a convolution of an isotropic PSF  $P^{\text{iso}}$

\*KSB did not consider the quantity  $\epsilon$  from eq. (81) as its usefulness was discovered one year after the work on KSB had started, in 1996. Furthermore, the derivation of the KSB formulae is easier with  $\chi$ .

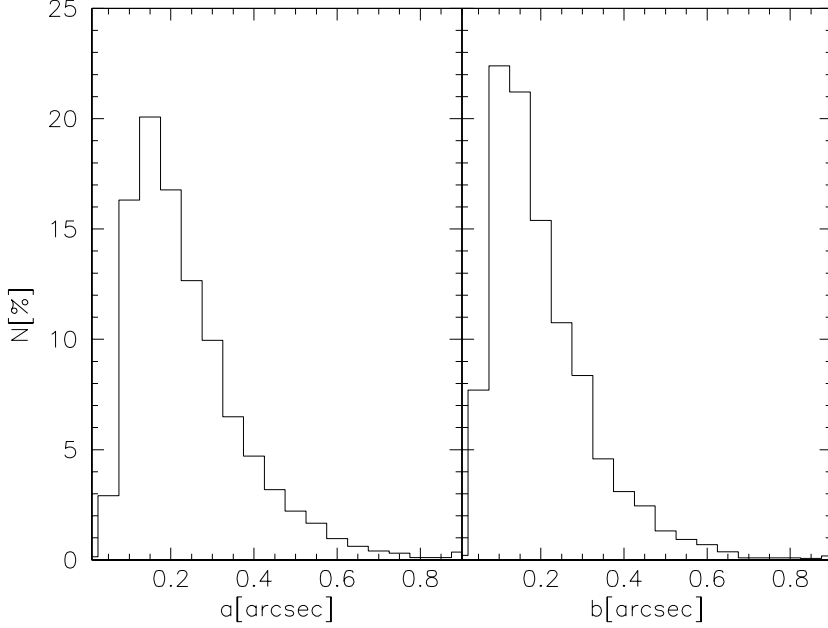


Figure 23: The distribution of intrinsic galaxy scale radii (Figure 2 from [16]): The plot shows the distribution of intrinsic scale radii for an exponential disk profile:  $I = \exp(-\sqrt{x^2/a^2 + y^2/b^2})$ . The values were obtained from the stuff program (see Appendix A). We clearly see that with a typical intrinsic scale of 0.1 – 0.2 arcsec the final size of most objects (that is after convolution with the PSF) is completely determined by the size of the PSF in ground based observations. Consequently, a measured ellipticity (like the ratio  $a/b$ ) is lower than the intrinsic one.

convolved with some anisotropic kernel  $q$

$$P(\vec{\theta}) = \int d^2\vartheta P^{\text{iso}}(\vec{\vartheta}) q(\vec{\theta} - \vec{\vartheta}). \quad (146)$$

The second moment tensor of  $q$  is traceless and its components are assumed to be small compared with those from  $P^{\text{iso}}$ . This is a pure assumption and it is not clear at all whether it is a good assumption for realistic PSFs. The derivation of the following final formula is highly technical and involved. It does not contribute to the understanding of the following and we only quote the result here. A thorough derivation can be found in Bartelmann & Schneider 2000 (see [4]).

It was shown that the total response of the galaxy ellipticity  $\hat{\chi}^s$ , which is the intrinsic ellipticity convolved with an isotropic function<sup>†</sup>, to a reduced shear  $\mathbf{g}$  and the PSF, is given by:

$$\chi - \hat{\chi}^s = -P^g \mathbf{g} + P^{\text{sm}} \mathbf{q}^*; \quad P^g = P^{\text{sh}} - P^{\text{sm}} (P^{*\text{sm}})^{-1} P^{*\text{sh}}. \quad (147)$$

$\chi$  is the observed ellipticity and the tensors  $P^{\text{sh}}$  and  $P^{\text{sm}}$  can directly be calculated from the galaxy's light profile and the weight function  $W$ . They consist of third and fourth moments of the light distribution.  $P^{\text{sh}}$  (*shear polarizability*) describes the response of the galaxy ellipticity to a gravitational shear in the absence of PSF effects. The tensor accounts for the fact that we have used weighted brightness moments.  $P^g$  modifies this tensor by a factor including the *smear polarizability* tensor  $P^{\text{sm}}$  to calibrate the shear estimate for the circular smearing by the isotropic part of the PSF. This calibration also depends on the corresponding tensors  $P^{*\text{sh}}$  and  $P^{*\text{sm}}$  from stellar objects containing the information of the PSF. The stellar anisotropy kernel  $\mathbf{q}^*$ , needed for

<sup>†</sup>This is of course not obvious but part of the involved derivation. It is only important to note that we still have  $\langle \hat{\chi}^s \rangle = 0$  if this relation holds for the intrinsic ellipticity distribution.

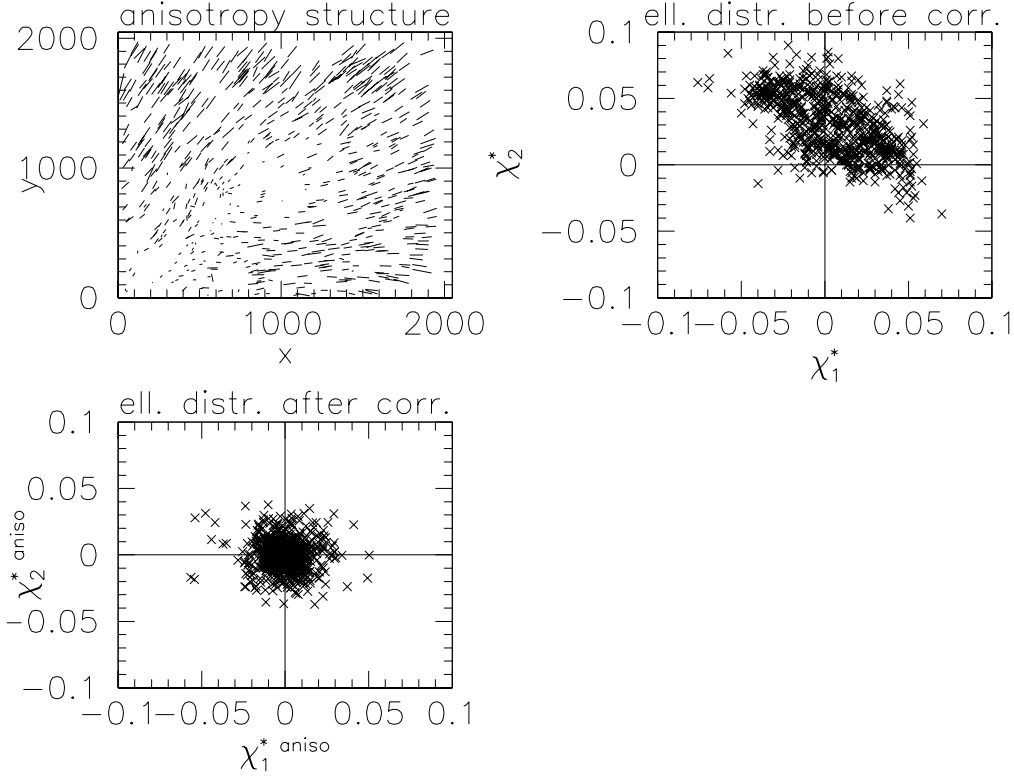


Figure 24: PSF anisotropy structure from a ground-based observation taken with the NTT: The upper left panel shows the measured weighted ellipticities (see section 6.2) from stellar objects in a globular cluster field. The direction of the sticks is parallel to the major axis of the image ellipse and the length is a measure for the modulus of ellipticity. The distribution in the  $\chi_1$ - $\chi_2$  plane is shown on the upper right. The anisotropy varies very smoothly over the field and can easily be fitted by a second order polynomial in  $x$  and  $y$ . The ellipticity distribution after the polynomial fit has been subtracted from the ellipticities is shown in the lower left panel.

the correction of the anisotropic part of the PSF, can be estimated by noting that  $\hat{\chi}^{s*} = 0$ ,  $\mathbf{g}^* = 0$  for stars, so that

$$\mathbf{q}^* = (P^{*sm})^{-1} \chi^*. \quad (148)$$

These formulae are not exact but they originate from Taylor expansions in  $\mathbf{g}$  and  $\mathbf{q}^*$  and terminating the expansions after the first order. The application of eq. (147) to estimate  $\mathbf{g}$  is a two-stage process. We first correct objects for the PSF anisotropy:

$$\chi^{\text{aniso}} := \chi - P^{sm} \mathbf{q}^* \quad (149)$$

and then consider averages over galaxy images in areas where we assume constant reduced shear:

$$\langle \chi^{\text{aniso}} \rangle - \langle \hat{\chi}^s \rangle = -\langle P^g \mathbf{g} \rangle = -\langle P^g \rangle \langle \mathbf{g} \rangle \rightarrow \langle \mathbf{g} \rangle = -\langle P^g \rangle^{-1} \langle \chi^{\text{aniso}} \rangle, \quad (150)$$

where we used  $\langle \hat{\chi}^s \rangle = 0$ . For simplicity eq. (147) is often used in the form:

$$\mathbf{g} = -(P^g)^{-1} (\chi^{\text{aniso}} - \hat{\chi}^s) \Rightarrow \langle \mathbf{g} \rangle = -\langle (P^g)^{-1} \chi^{\text{aniso}} \rangle \quad (151)$$

where we have to assume  $\langle (P^g)^{-1} \hat{\chi}^s \rangle = 0$ . With the procedure described below for applying the KSB formulae it turned out that this assumption holds, and the estimators from eq. (150) and eq. (151) are in very good agreement, although in practice the latter is easier to handle.

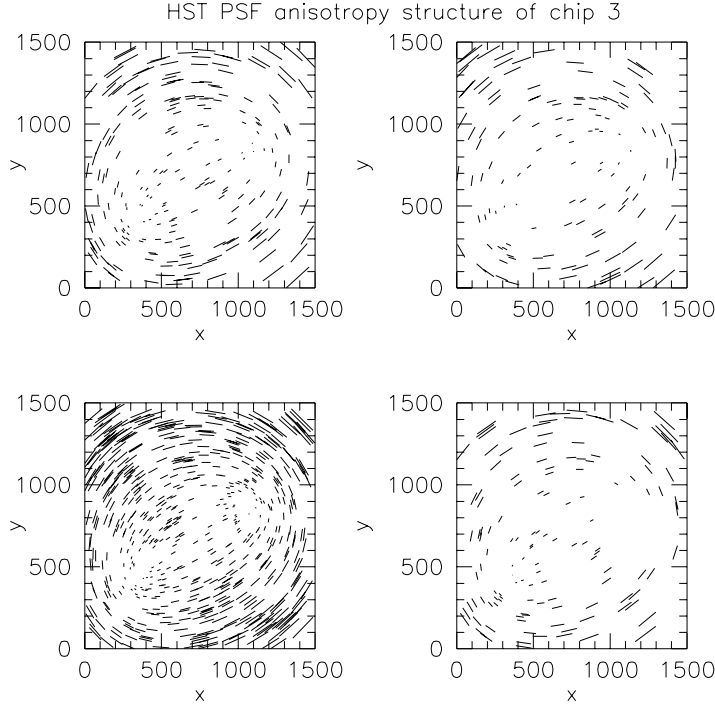


Figure 25: PSF anisotropy in HST images: The four panels show the PSF anisotropy in Chip 3 of WFC3, estimated from globular cluster fields. The panels show two observations in the F#814 (*I*-band) filter, one in the F#606 (*V*-band) and one in the F#300W (*U*-band) filter. The exposures were spread over a period of two years. As a single HST chips only covers  $\approx 1' \times 1'$ , an exposure outside a star field does not contain enough stars to estimate PSF properties appropriate on the whole field. As the Figure shows, the anisotropy pattern is very stable, though. Consequently, if one has observed a field for lensing studies (a galaxy cluster for instance), one can estimate PSF properties with a globular cluster field taken in the same period.

### 6.3 Practical problems with the KSB formulae

Although the application of the above formulae seems straightforward, there are several problems in practice:

- The upper left panel of Fig. 26 shows the trace of the  $P^g$  tensor as a function of object size. The tensor was calculated semi-analytically with objects having a Gaussian profile and no pixel noise (see also section 6.6). The shear is estimated with the inverse of this tensor and as discussed above it contains the correction for the circular smearing of the PSF. As we expect, the correction factor becomes a strong function of object size for smaller objects. For larger objects the correction tends to a constant. For stars, the tensor is zero by definition [see eq. (147)]. In the upper right panel we see the two shear estimates  $-((P^g)^{-1}\chi)$  calculated for noise free objects. The galaxies are intrinsically randomly oriented and the field contains no input shear. We note that the large corrections for very small objects lead to a small amount of objects that have an unphysical total modulus  $|g|$  of 1 or larger. Such objects would have to be excluded from the analysis. In the lower panels of Fig. 26 we repeated the same calculation with pixel noise. As expected, noise mostly influences

quantities from small objects. The spread in  $P^g$  becomes fairly large for small object sizes and it can become unphysically negative. Also the shear estimates are strongly affected by noise and the number of objects with unphysically large shear estimates is increased tremendously. One of the major issues in the application of the KSB formulae is to find a reliable set of objects. For this we have to find appropriate cuts in some parameter space (besides the object size, we observe dependencies on the magnitude  $m$ , the signal-to-noise ratio (see below) or the ellipticity). A large problem here is that we would have to find the best possible cuts for every observation individually. Finally we realised that we can solve the problem in an elegant way by introducing a weighting scheme in the shear estimates from eq. (150) and eq. (151).

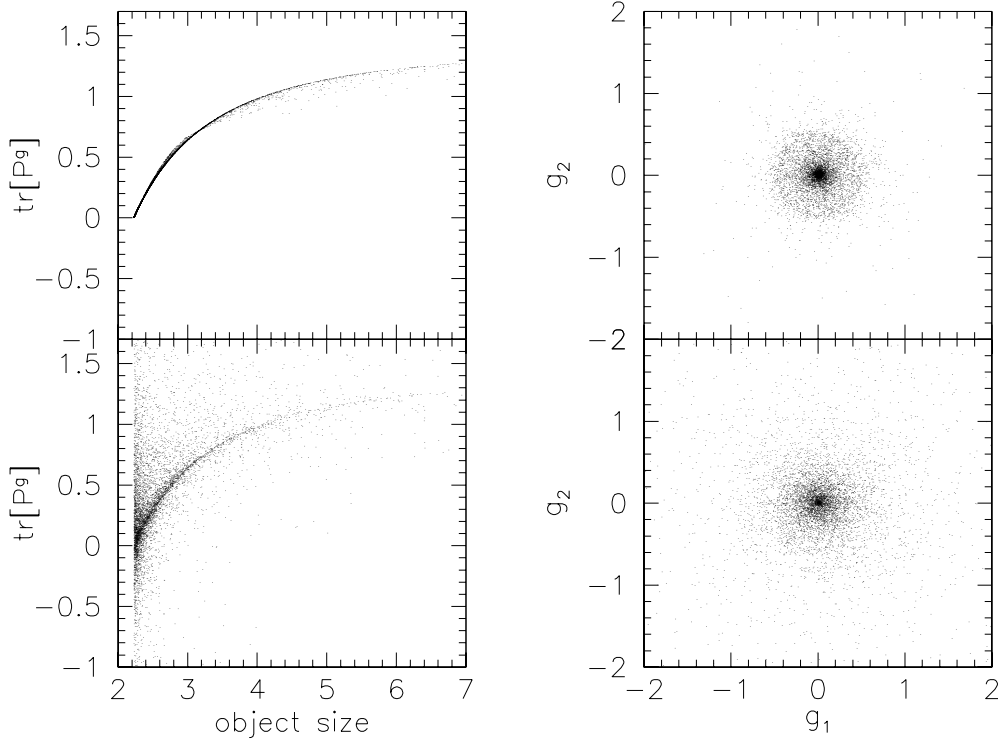


Figure 26: Noise in shear estimates: The upper left panel shows the trace of the  $P^g$  tensor calculated for noise free objects. As expected this tensor, responsible for the correction of PSF smearing in shear estimates, has a strong dependence of object size. The upper right panel shows shear estimates  $(P^g)^{-1}\chi$  for noise-free objects. The lower panels show the same with pixel noise. For more details see the text.

- When looking at calculated values of the components for the  $P^g$  tensor we note that the diagonal elements are very similar and that the off-diagonal terms are typically at least one order of magnitude smaller than the diagonal elements [this is not unexpected; for the special case, where we would calculate the unweighted surface brightness moments, and in the absence of a PSF we would get  $P_{11}^g = P_{22}^g = 2$  and  $P_{12}^g = P_{21}^g = 0$ . This follows from a Taylor expansion of eq. (85)]. The use of the tensor equations (150) and (151) is therefore often simplified by estimating  $P^g$  with the scalar quantity  $P_s^g$

$$P_s^g = 0.5 \text{tr}[P^g] ; \langle g \rangle = - \left\langle \frac{\chi^{\text{aniso}}}{P_s^g} \right\rangle. \quad (152)$$

This, on the one hand, leads to a more stable estimation of the dominating diagonal part of

the tensor, but on the other hand, it may introduce systematic biases due to the neglect of the off-diagonal terms.

- For the application of eq. (147) we need to know the quantities  $(P^{*sm})^{-1}P^{*sh}$  and  $\mathbf{q}^*$ , that have to be calculated from stellar objects, at the position of every galaxy that we want to correct. As discussed earlier, the anisotropy of stars is a smooth function over an image. On our galaxy fields we typically have about 30-40 isolated, randomly distributed, foreground stars so that we can accurately fit  $\mathbf{q}^*$  with a low-order polynomial, and thus estimate it everywhere on the image. The quantity  $(P^{*sm})^{-1}P^{*sh}$  is typically nearly constant over the field so that it is estimated by means over all stars. If this assumption does not hold, it can be fitted with a polynomial as well.
- So far the weight function  $W$  for the calculation of our moments has not been specified. In the original work, KSB used a Gaussian with a width direct proportional to the size of the object under consideration

$$W(|\vec{\theta}|) = \frac{1}{2\pi r_g^2} e^{-\frac{|\vec{\theta}|^2}{2r_g^2}}, \quad (153)$$

where  $r_g$  is a measure for the size of the object (see below). So, in the expression for  $P_g^s$ , the quantities for the galaxy and for the stars (and the polynomial fits of this quantity) are calculated with weight functions of different scale. Hoekstra et al. 1998 (see [26]) used simulations, and were the first to show that all quantities should be calculated with the scale of the object we want to correct. Formally, the KSB formalism should be free of any scale.

## 6.4 Motivation for our simulations

Although the complete KSB algorithm was finally written down in 1997 and it had become the state-of-the art method for weak lensing studies, it was not tested thoroughly until 1999. At this time we were just about to start our studies on wide-field-imaging data and we started collecting data for our cosmic shear project (see the next two sections). For these studies we had to be sure that gravitational shear can be measured reliably at a one percent level. At the same time, Emmanuel Bertin had just finished his *SkyMaker* program that produces simulated CCD images with realistic PSFs for ground-based observations. Therefore, we started extensive simulations with the KSB algorithm with the intention to clarify the following questions:

- Konrad Kuijken showed in 1999 (see [41]) that the KSB PSF anisotropy correction already breaks down for very simple PSF profiles (he used the superposition of two Gaussian profiles). This is not surprising considering the assumption that a PSF can always be decomposed into an isotropic part and a compact, anisotropic kernel [eq. (146)]. We have to be sure that the proposed correction procedure is valid for PSF profiles that are observed in ground-based observations.
- What is the best way to apply the KSB formulae? Is it true that the scalar application of eq. (152) gives the better result, or does the omission of the off-diagonal terms (although they are very noisy) introduce systematic biases?
- What influence does pixel noise have on the shear estimation?
- The KSB formulae are valid to first order in the shear ( $|g| \ll 1$ ). Is this expansion valid for typical weak-lensing applications with ground-based observations (from  $|g| = 0.01$  for cosmic shear up to  $|g| = 0.2$  for cluster mass reconstructions), or should we expand the formalism to higher order?
- Can we set up a fully automatic procedure, from astronomical images to an object catalog, for reliable shear measurements?

One of the major issues for the last point is the selection of objects that can be used for shear estimation (see above). This problem has two aspects. First, the practical problems discussed above to select well-separated objects after detection in an astronomical image. And second, what signal-to-noise do we need for an object so that the measurement is not dominated by noise. The latter issue is considered first.

## 6.5 The introduction of a weighting scheme in shear estimates

During our tests it turned out that it is difficult to find appropriate parameter cuts for every single image. Figure 26 for instance suggests that we should discard very small objects, and objects that have an unphysically high shear estimate. To find ideal limits for an arbitrary observation is difficult. As our applications mainly deal with *empty field*<sup>‡</sup> observations, we can use the following fact to exclude noisy objects by a weighting mechanism rather than rejecting them completely:

For studies in empty fields, where we always deal with the weak lensing case, the observed ellipticity distribution is to first order identical with the intrinsic ellipticity distribution. Furthermore it is a reasonable assumption that this distribution does not change over a few arcminutes, the typical scale of a ground-based observation. We also expect that noise in the shear estimates depends strongly on some multidimensional parameter space  $\pi$ . As discussed above, it is expected that very small objects with a low S/N are mostly influenced by noise. Figure 27 clearly confirms this. It shows the variance  $\sigma_g^2$  of the shear estimate as a function of object size and S/N ratio. The basic idea of our weighting scheme is to weight every galaxy with  $U = 1/\sigma_g^2$ . To estimate the variance  $\sigma_g^2 = \frac{1}{N} \sum |\mathbf{g}|^2$  one has to bin galaxies according to the parameter space  $\pi$ . Obviously, galaxies are not uniformly distributed in parameter spaces tracing the noise properties. There are much more small, low S/N objects than large and bright ones. So, binning galaxies in a regular grid in  $\pi$  is not optimal. Instead of using an adaptive grid, the *next neighbour* approach is applied: For every galaxy we consider its  $N$  nearest neighbours (typically  $N \simeq 20$ ) and assign the inverse of  $\sigma_g^2$  from these neighbours as weight to that galaxy. Thereby, the distance  $d$  from a galaxy  $k$  to its neighbour  $l$  in a parameter space  $\pi$  consisting of  $i = 1..M$  elements is defined by

$$d = \sqrt{\sum_{i=1}^M (\pi_i^k - \pi_i^l)^2}. \quad (154)$$

Obviously, objects do not have a unique set of next neighbours in an arbitrary space (binning our data to half its resolution for instance leads to a scaling of the object size but would leave quantities like  $m$  unchanged). We did not investigate possible consequences of this here. For finding a reasonable parameter space  $\pi$  for this weighting scheme and to investigate other aspects, not depending on the details of the PSF, we first performed simulations with Gaussian profiles that we describe in the next section.

## 6.6 Semi-analytical simulations with Gaussian profiles

Gaussian profiles do not fit real galaxies very well and in particular they are not a good estimate for realistic PSFs. Nevertheless, they allow a quick investigation of potential biases connected with the KSB procedure:

- We can investigate the influence of pixel noise on the shear estimates in detail.
- Pixel noise also influences the estimate of the object centre relative to which all lensing quantities are calculated. So, we have to investigate the accuracy with which the object centre can be determined, and how errors in this quantity influence the final shear estimate.

---

<sup>‡</sup>We want to understand under this term a piece of the sky not dominated by massive galaxy clusters.



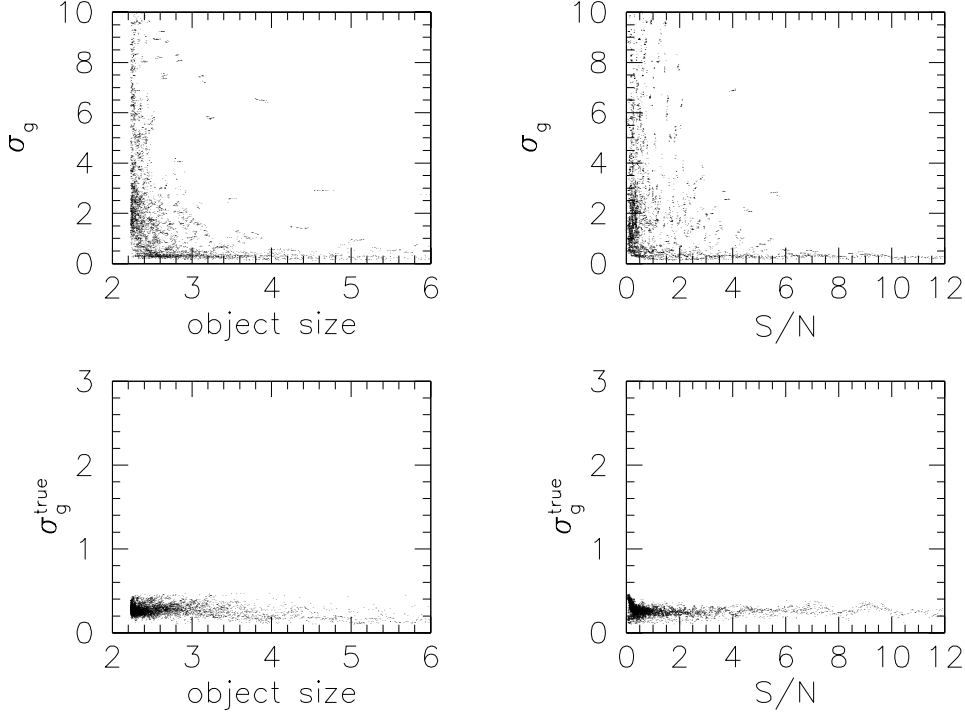


Figure 27: The variance of shear estimates as a function of object size (left panels) and of S/N ratio (right panels). For a definition of the S/N ratio, see eq. (157). In the ideal, noise free case, the variance should always be the same (lower panels), whereas it becomes a function of object size and S/N in the presence of noise. Consequently, this variance can be used to weight individual objects.

- From eq. (149), eq. (150) and eq. (151), we see that the final shear estimates through the tensor  $P^g$  is calculated with different profiles. The tensor is calculated with the observed profile. But this profile depends on the PSF anisotropy present. In the ideal case, the tensor should be calculated from a profile not affected by a PSF anisotropy. We can investigate the influence of this drawback.

Moreover, most of the quantities defined in KSB can be calculated analytically with Gaussian profiles, which permits to check that the numerical simulations are done properly. Our simulations with Gaussian profiles were performed in the following way:

1. We used the *Stuff* program to generate a galaxy catalog with intrinsic magnitudes, sizes and ellipticities. The program is part of the SkyMaker tool and is described in more detail in Appendix A. See Fig. 28 for a comparison of object properties from the stuff program and a real CCD image.
2. We put Gaussian galaxy profiles that were convolved with other Gaussian profiles (mimicking the stars) on a pixel grid:

$$I = A \exp\left(-\frac{x'^2}{a^2} - \frac{y'^2}{b^2}\right), \quad (155)$$

where  $x' = x \cos(\theta) - y \sin(\theta)$ ,  $y' = x \sin(\theta) + y \cos(\theta)$  for an object with semi-major (semi-minor) axes  $a$  and  $b$  and a position angle  $\theta$  with respect to the abscissa. The dimensions of the pixels and the PSF profiles were adapted to a CCD with  $0.2''$  resolution and a seeing (FWHM) of  $0.7''$  (This means  $a_* = b_* = 2.012$  pixel units for an isotropic PSF). We chose  $A_* \approx 1000\sigma_{\text{sky}}$  to get high S/N measurements for all stellar quantities. The intrinsic scales

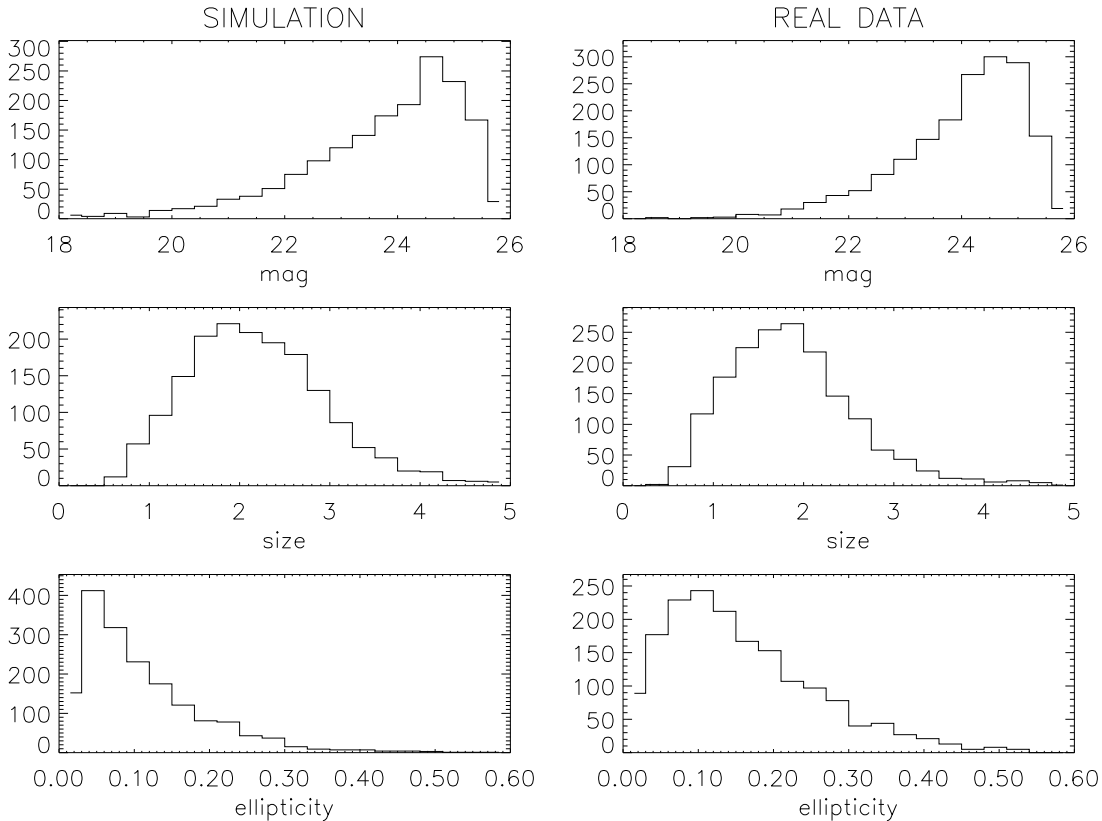


Figure 28: A comparison of galaxy properties in one of our SkyMaker images with real data in one of our  $I$  band CFHT fields (Figure 14 from [16]): The galaxy properties of the input catalog were generated by the stuff program. Both images have a seeing of  $0.7''$  and corresponded to 12000 sec. exposures. Objects have been detected and the quantities have been calculated with SExtractor. The two upper panels show a comparison of the magnitude distribution, the two middle panels of the object size defined as  $\sqrt{AB}$ , where  $A$  and  $B$  are the semi major and semi minor axis of the image ellipse. The quantities are formally defined in eq. (206). The lower panels compare the ellipticity parameter  $\epsilon = (1 - A/B)/(1 + A/B)$ . The only noticeable strong difference is that the ellipticity distribution in the real data is slightly broader. The shift of the peak in this distribution is caused by a strong PSF anisotropy in the CFHT data.

for the galaxies  $a$  and  $b$  were taken from the output of the Stuff program (see Fig. 23) as well as the apparent magnitudes fixing the amplitude  $A$ . The centres of the objects were chosen randomly within a pixel. We put every object separately on the pixel grid and immediately analysed it. So, in these simulations we have the ideal case of isolated objects and also very low S/N objects can be studied in detail.

3. Before calculating lensing quantities we estimated the object centre by iteratively solving the equation:

$$\vec{\theta} = \frac{\int \vec{\theta} I(\vec{\theta}) W(|\vec{\theta} - \vec{\theta}|) d^2\theta}{\int I(\vec{\theta}) W(|\vec{\theta} - \vec{\theta}|) d^2\theta}, \quad (156)$$

where the pixel centre of the true position was used as starting point. As weight function we used the Gaussian from eq. (153), where we chose as object size  $r_g = 1.107 \frac{\sqrt{2} a_s b_s}{\sqrt{a_s^2 + b_s^2}}$ , where  $a_s$  and  $b_s$  are the scale radii of our objects after smoothing with the PSF. This is not the same  $r_g$  as that used in the KSB algorithm. The quantity used there is described in Appendix B. Our definition is the same as the KSB  $r_g$  for a circular object. It was checked that the

$g_1$	0.01	0.03	0.15	0.25	0.08	0.2	0.14
$g_2$	0.007	0.05	0.1	0.2	0.27	0.14	0.2

Table 2: The shear combinations we investigated in this section with Gaussian profiles for the galaxies.

results from this section do not depend significantly on this choice. Moreover we defined the

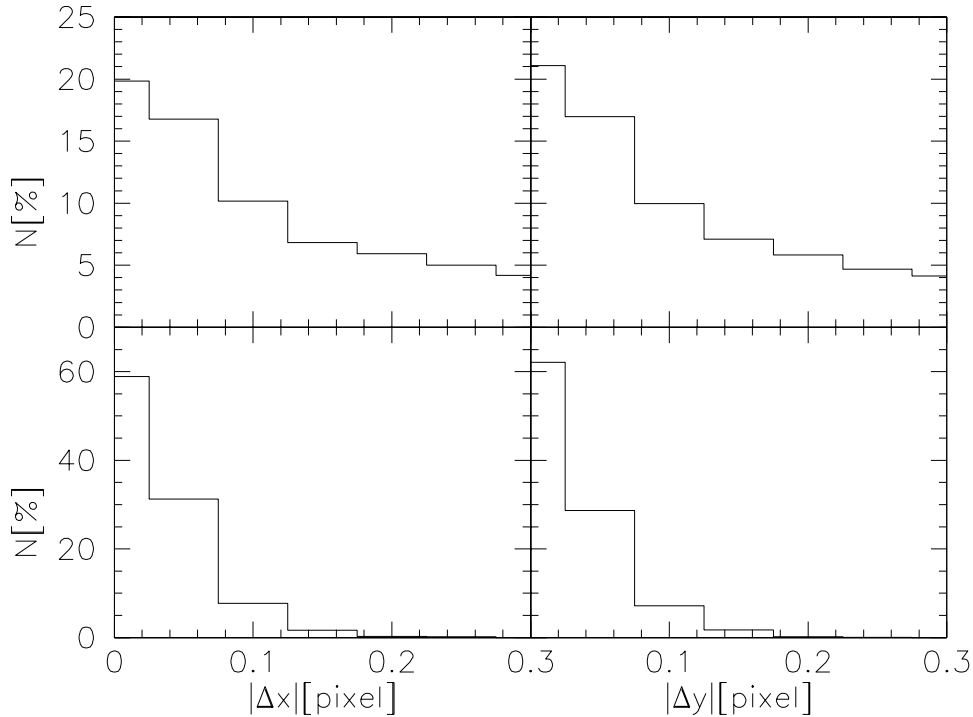


Figure 29: The distribution of the errors  $\Delta x$  and  $\Delta y$  of galaxy centres (Figure 3 from [16]): The upper panels show the distribution of all the objects, the lower ones show those with  $S/N > 2$  (see text).

following weighted flux as  $S/N$  for an object:

$$\frac{S}{N} = \frac{\int I(\vec{\theta}) W(|\vec{\theta}|) d^2\theta}{\sigma_{\text{sky}} \sqrt{\int W^2(|\vec{\theta}|) d^2\theta}}. \quad (157)$$

The accuracy with which we can determine the object centre is shown in Fig. 29. For the sample with  $S/N > 2$  we can determine the object centre with an accuracy of  $1/20$  of a pixel.

4. We then performed the KSB measurements. All quantities were calculated with respect to the estimated object centre from the last step. We note that all the calculations were done two times: First, without noise, and second, for a Gaussian noise model with  $\sigma_{\text{sky}} = 13.0$  (this adapts these calculations to the characteristics of the SkyMaker simulations from section 6.7.1). Hereby we added the noise only to the galaxies but not to the stars.

We performed all the calculations for the shear combinations listed in Table 2. For each combination we generated a final catalog with about 75000 objects having  $15.0 < m < 28.0$ . If not stated otherwise our galaxies have the intrinsic ellipticity generated by the Stuff program.

### 6.6.1 Estimation of $P^g$

Before searching for the best possible parameter space for our weighting scheme in the shear estimates (see above), we investigate whether we can lower the noise in the quantity  $P^g$  and improve the shear estimate in this way. This is motivated by the observation that this quantity itself is a strong function of the object size, as discussed above. Consequently, we can hope that estimating this quantity by means or medians in some parameter bin lowers its noise. As all quantities were calculated with and without sky noise we can clarify this issue here.

For this we took the  $g_1 = 0.25; g_2 = 0.2$  realisation and the scalar  $P^g$  estimator from eq. (152). We smoothed  $P^g$  with the next neighbour approach described in section 6.5 in several parameter spaces to obtain  $P_{\text{smooth}}^g$  and compared these values with the noise-free  $P_{\text{true}}^g$ . We finally used the median of the next neighbour  $P^g$  estimates as  $P_{\text{smooth}}^g$ . This turned out to give better results than the mean. The result is shown in Fig. 30. We see that  $P^g$  is mainly a function of the object size  $r_g$  and the modulus of the raw, noise-free but unobservable ellipticity  $\chi_{\text{true}}$ . In this case we could really significantly lower the noise in this quantity. In practice, we only observe noisy values for  $\chi$  and a smoothing as a function of  $r_g$  alone seems as good as smoothing in  $r_g$  and  $|\chi|$ . Including quantities like  $m$  or S/N is clearly worse.

In order to see whether these smoothings improve the final shear estimates, we calculated the estimator  $\mathbf{g} = \frac{\mathbf{X}}{P^g}$  and  $\mathbf{g}^{\text{smooth}} = \frac{\mathbf{X}}{P_{\text{smooth}}^g}$  for the various smoothings. Figure 31 shows that even for the best  $r_g - |\chi_{\text{true}}|$  we find  $\Delta\mathbf{g} \approx \Delta\mathbf{g}^{\text{smooth}}$ . Also other tests we performed show that smoothing  $P^g$  does not improve the shear estimation. We conclude that noise in the quantity  $\chi$  is the major contribution to the noise in the final shear estimate.

### 6.6.2 Weighting scheme

We now search for the best possible parameter space to calculate  $\sigma_g^2$ , and thus the weighting of the galaxy ellipticities. For this investigation a catalog with intrinsically round objects was used to isolate pixel noise errors from those of the intrinsic ellipticity distribution. From Fig. 32 we see that using  $\pi_U = (r_g, m)$  and  $\pi_U = (r_g, \text{S/N})$  as parameter spaces for the  $\sigma_g^2$  calculation give good and very comparable results in the final shear estimation. They improve the estimation significantly in comparison to no weighting. Therefore, weights from the  $\pi_U = (r_g, \text{S/N})$  smoothing are now used throughout.

### 6.6.3 Accuracy of shear estimates

In the following, we present shear estimates with our Gaussian simulations. They cannot represent real observations, and as discussed above, they were not put on CCD images. They represent the ideal case where we investigate isolated objects only. This fact, though, allows us to evaluate the value of our object detection and selection later. We can compare the results quoted below with those when we finally put the Gaussians on images and run them through our complete procedure. We have investigated the  $g_1, g_2$  combinations listed in Table 2. Our measurements were done first without PSF anisotropy and then with a stellar axis ratio  $b_*/a_* = 0.9$ , oriented such that  $\mathbf{q}_2^* = 0$ . To get the results quoted below, we have calculated averages  $\langle \mathbf{g} \rangle$  and errors  $\sigma_{\langle \mathbf{g} \rangle}$  over all galaxies in our catalogs. Hereby, we estimated means, standard deviations and mean uncertainties from a weighted quantity  $x$  by

$$\langle x \rangle = \frac{\sum_k x_k U_k}{\sum_k U_k}; \sigma_x^2 = \frac{\sum_k U_k}{(\sum_k U_k)^2 - \sum_k U_k^2} \sum_k U_k (x_k - \langle x \rangle)^2; \sigma_{\langle x \rangle}^2 = \frac{\sum_k U_k^2}{(\sum_k U_k)^2} \sigma_x^2. \quad (158)$$

The results for the noise-free calculations are shown in Fig. 33. Note here that we have also used the same weighting scheme for the noise-free calculations, although this would not be necessary (see Fig. 27). The plot shows the following trends for the calculations *without* PSF anisotropy:

- For small input shear values  $|\mathbf{g}| \leq 0.1$  we can clearly recover the input shear within  $|\Delta\mathbf{g}| \approx 0.01$  with the tensor as well as with the scalar shear estimator. For larger shear values we

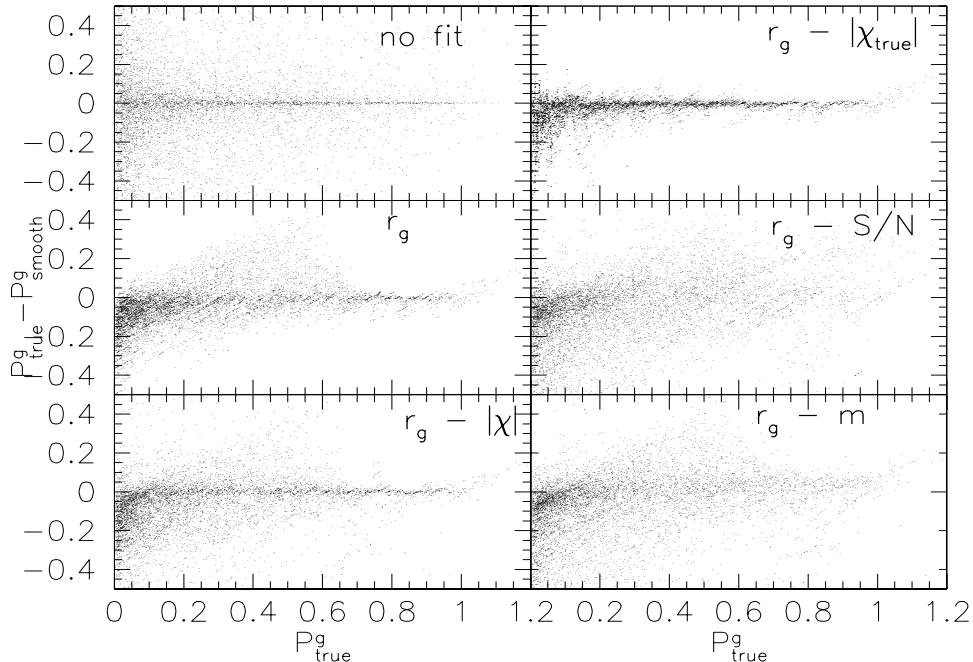


Figure 30: The difference between the true and smoothed  $P^g$  as function of  $P^g_{\text{true}}$  with  $P^g_{\text{smooth}}$  calculated in different parameter spaces (Figure 4 from [16]): In the upper left with no smoothing, the errors of  $P^g$  are randomly distributed around the true value. In the middle left panel (smoothing as function of  $r_g$ ) we see a systematic tail for intermediate values of  $P^g_{\text{true}}$  indicating the dependence of  $P^g$  on at least one second parameter. Otherwise this smoothing reduces the noise very well. The middle right and lower right panels show that  $P^g$  does not depend on S/N or  $m$ . From the upper right panel we see that  $P^g$  depends on the raw ellipticity  $|\chi|$  as well as on  $r_g$ . Unfortunately our measured, noisy ellipticities lead to significant noise in the smoothing (lower left panel).

systematically underestimate the shear, with a maximal underestimation of 0.02 to 0.03 for the highest input shear of 0.2 – 0.25. We note that for both shear estimators the underestimation as a function of input shear can be very well represented by a straight line indicating that the underestimation is a constant fraction of the input shear over the whole range of input shears considered here. This fraction is about 10% – 15% for the scalar and 5% – 10% for the tensor shear estimate.

- We also note that the results are completely equivalent in both shear components.

For the simulations *with* PSF anisotropy in the  $g_1$ -direction we conclude the following:

- Considering the  $g_2$  component the results are very similar to the isotropic PSF case.
- For the  $g_1$  component the scalar correction underestimates the shear by an amount comparable to the isotropic PSF case. In contrast, the tensor correction overestimates the input shear in the direction of the PSF anisotropy. This relative overestimate is about 20% for the high  $|g|$  and can reach up to 100% for the lowest input shears. In Fig. 36 we show analytic results with intrinsically round objects and an input shear of  $g_1 = 0.25$ . These calculations confirm that for Gaussian profiles the tensor correction overestimates the final shear in the direction of a PSF anisotropy. From our analysis of the SkyMaker simulations in section 6.7, where we have encountered similar problems with the tensor correction, we conclude that the problems are not connected with the anisotropy correction of KSB but with the smearing correction.

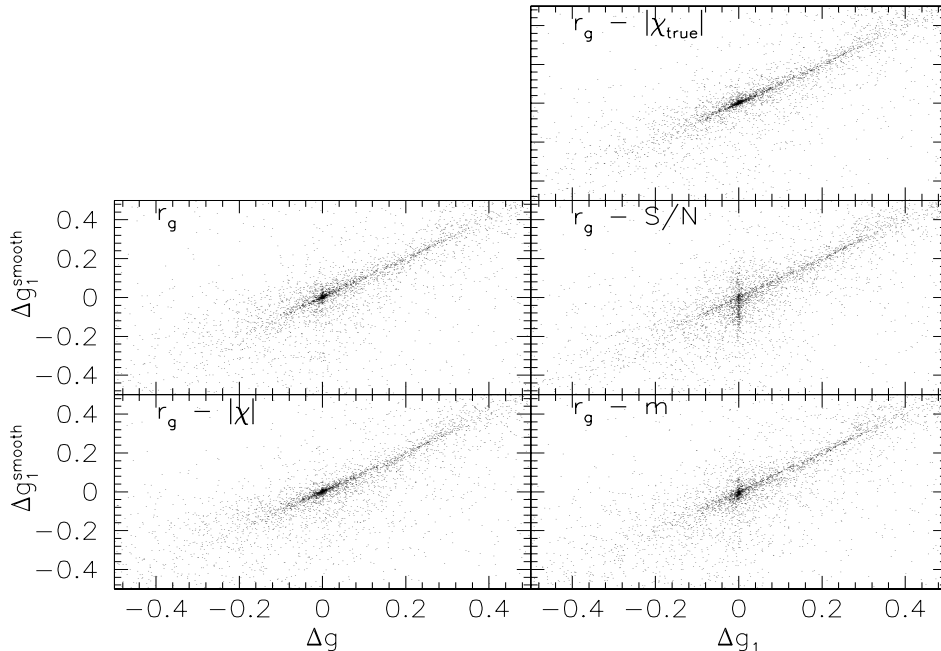


Figure 31: The errors  $\Delta g_1$  vs.  $\Delta g_1^{\text{smooth}}$  for the various smoothings of  $P^{\mathcal{E}}$  (Figure 5 from [16]): Even for the best smoothing of this quantity with  $r_g$  and the true raw ellipticity  $\chi_{\text{true}}$  we have  $\Delta g \approx \Delta g^{\text{smooth}}$  indicating that smoothing  $P^{\mathcal{E}}$  does not improve the shear estimation. The parameter spaces  $r_g$  and  $r_g - |\chi|$  give similar results while the  $r_g - S/N$  smoothing especially shows a spread of  $\Delta g_1^{\text{smooth}}$  where  $\Delta g_1 = 0$ .

For the calculations including sky noise we have investigated the whole sample, and two subsamples with  $S/N < 2$  and  $S/N > 2$ . Figure 34 shows that the  $S/N > 2$  sample results are similar to the noise-free case, i.e. in the absence of PSF anisotropy we have very good results with the tensor estimate (however, there is no underestimation in the tensor case at all in contrast to the noise-free case; in fact, the underestimation is lowered to 4% in the noise-free case for the same subsample of galaxies), and about a 10% – 15% underestimate for the scalar calculations. For the low signal-to-noise objects ( $S/N < 2$ ), the shear is underestimated by about 30% in the scalar and up to about 50% in the tensor case, where the underestimate in the tensor case shows an increasing trend for increasing shear. The error in the object centre determination is not the main reason for the underestimate: for  $S/N < 2$  we repeat the same calculations using the true object centres. The results from this test are very comparable to those shown here. For the whole sample, the scalar correction is better with a relative underestimate of about 10% – 15% over the tensor one which has an underestimate of about 20% – 30%. If PSF anisotropy is present, noise-free and noisy cases give the same trends. Figure 35 shows the noise amplitude of the estimators. We considered the ratio of the shear uncertainty from noisy objects  $\sigma_{\langle \mathbf{g} \rangle}$  and the noise-free calculations  $\sigma_{\langle \mathbf{g} \rangle}^{\text{true}}$ . For  $S/N > 2$ , pixel noise increases the errors in the shear estimates by about a factor of 2, for  $S/N < 2$ , by a factor 5-10, and by a factor of 5 for the whole sample.

#### 6.6.4 Summary

With our proposed way to apply the KSB technique to measure gravitational shear we come to the following conclusions when dealing with Gaussian profiles:

- Our analysis shows that the scalar estimator from eq. (151) systematically underestimates shear by about 10% – 15%. For low- $S/N$  objects, the underestimate is about 30%. The

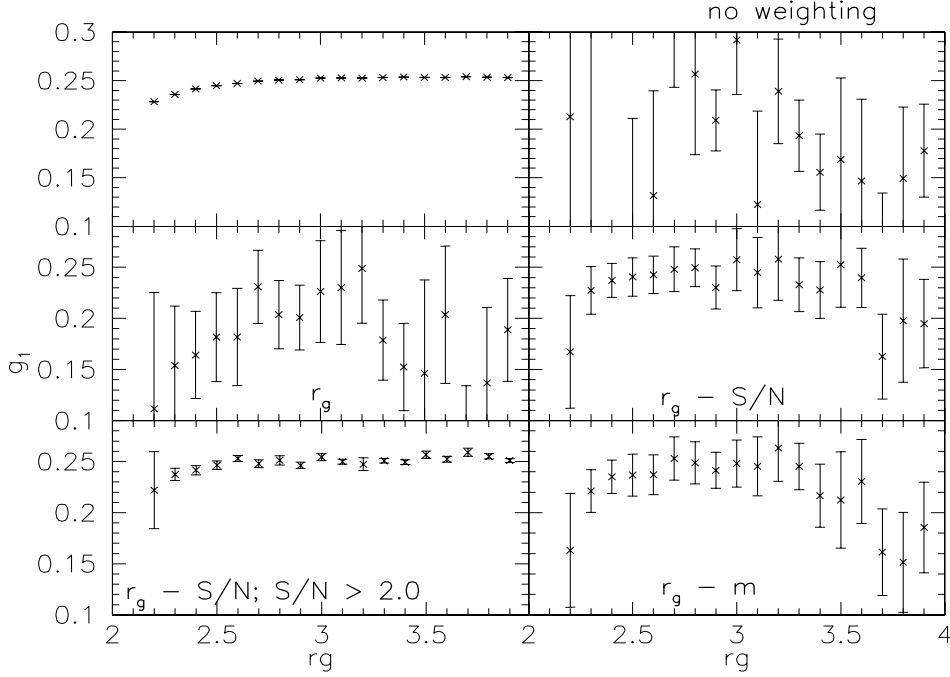


Figure 32: The measured reduced shear  $g_1$  as a function of the object size for various parameter spaces used in our weighting scheme (Figure 6 from [16]): The upper left panel shows the result for noise free objects. The upper right panel shows results if no weighting at all is applied. The middle left panel is for the parameter space  $r_g$ , the middle right for  $r_g - S/N$  and the lower right for  $r_g - m$ . The lower left panel shows the result for  $r_g - S/N$  when rejecting all  $S/N < 2$  objects. We see that  $r_g$  and  $S/N$  are largely independent quantities.

tensor estimator from eq. (150) is better for high- $S/N$  objects, showing no systematic over- or underestimation in the final result. This estimator is less stable for low- $S/N$  objects when there is a systematic underestimation that can reach more than 50%. When including PSF anisotropy, the scalar estimator still gives very stable and comparable results both in the direction of the PSF anisotropy and perpendicular to it. In contrast, we overestimate the shear in the direction of PSF anisotropy in the tensor case. We conclude that the scalar estimator is more stable and conservative. Both estimators show very similar noise properties.

- We have shown that smoothing  $P^g$  does not improve the final shear estimates over taking raw, noisy values.
- We can give an objective parameter cut  $S/N > 2$  for which we can measure shear with about the same accuracy as with no sky noise. Our  $S/N$  parameter [see eq. (157)] is very convenient as we do not have to find the threshold for every observation individually (as would be the case for quantities like  $m$ ).

As a final step we have repeated the analysis done here putting our Gaussian galaxies and stars on FITS images and analysing them as the SkyMaker simulations described in the next section. In addition to the steps presented so far, we first have to perform object detection and selection for the shear determination. This was done with a modified version of the so-called *imcat* software, Nick Kaiser’s implementation of the KSB algorithm. The package basically consists of two parts: A hierarchical peak finder for object detection (the *hfindpeaks* program), and a separate tool to calculate the relevant quantities from the detected objects (the *analyse* program). These tools and our exact procedure from the images to final object catalogs are described in more detail

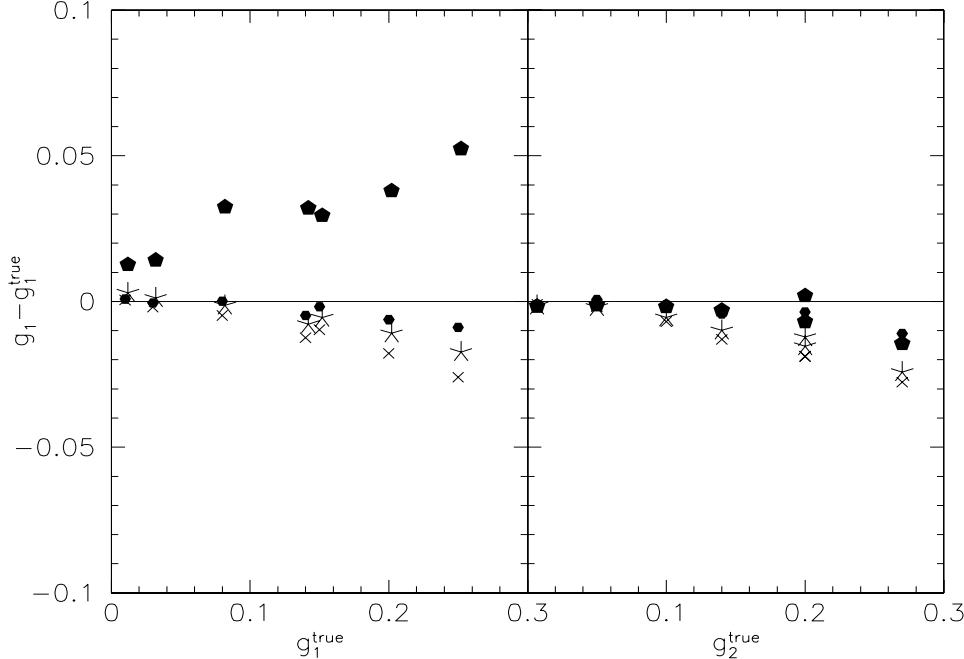


Figure 33: The errors on shear recovery for noise-free objects (Figure 7 from [16]): The crosses and filled hexagons show results for the scalar and tensor estimates with no PSF anisotropy present, the stars and filled pentagons for calculations with a 10% PSF anisotropy in the  $g_1$ -direction. No object selection was done, and our weighting scheme was used. The points were obtained by averaging over all  $\approx 75000$  galaxies that we have for every shear realisation in our catalog. The error bars are less than 0.1 percent and so smaller than the symbols. They have been omitted for clarity of the plot.

in Appendix B. An important difference with the previous semi-analytical calculations is the fraction of  $S/N < 2$  objects. Previously they constituted more than 60% of all the objects, but for the simulations where objects have to be detected first, this fraction is very small, about 5% – 10% in the final catalogs. It turned out that shear estimates with the subsample  $S/N > 2$  and the complete sample are now similar, so that the small subsample of low- $S/N$  objects becomes unimportant. Therefore, only the results with all objects without a cut in  $S/N$  will be shown from now. Figure 37 displays the results from the simulations. They are comparable to those obtained from the  $S/N > 2$  subsamples in Fig. 34. This check ensures that our shear analysis pipeline gives results consistent with analytical prediction, and that it is ready to be used for the final analysis involving the realistic simulations done with SkyMaker. These will be presented in the next section.

## 6.7 The SkyMaker simulations

### 6.7.1 The image characteristics of our SkyMaker simulations

For every combination of the shear and PSF, ten SkyMaker images were produced and analysed. Each image has the characteristics of a 12000 sec.  $I$  band exposure taken at a telescope with a 3.5m primary and a 1m secondary mirror, and has a dimension of  $2048 \times 2048$  pixels with a scale of  $0.206''$  per pixel. The atmospheric seeing in every image was  $0.7''$ . For every image we detected our objects and analysed them as described in Appendix B. A comparison of magnitude distribution, object size and ellipticity distribution between detected objects in one of the images and a 12000 sec.  $I$  band CFHT image having a measured seeing of  $0.7''$  was shown in Fig. 28. A



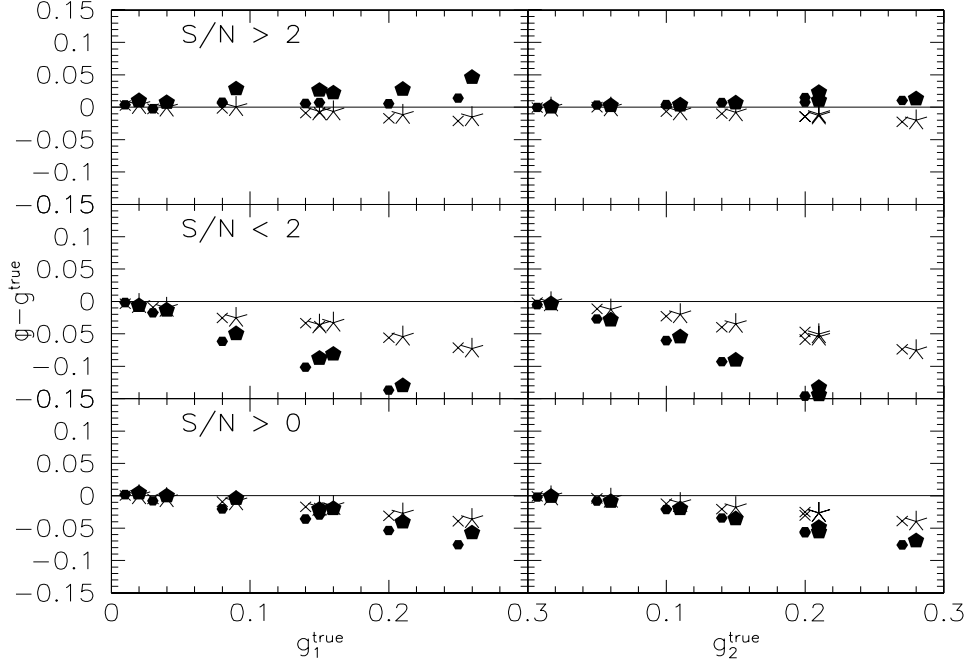


Figure 34: The same as Fig. 33 for noisy objects (Figure 8 from [16]): The two upper most panels show results for  $S/N > 2.0$ , the middle panels for  $S/N < 2.0$  and the lower panels for the whole sample. For high  $S/N$  objects, the tensor estimation is more accurate but in general the scalar calculation is much more stable towards low signal objects. For more details see the text.

	PSF_1	PSF_2	PSF_3	PSF_4	PSF_5	PSF_6	PSF_7	PSF_8
$\chi_1$	0.00	0.00	0.00	0.00	0.1	0.06	0.04	0.00
$\chi_2$	0.00	0.01	0.00	0.00	0.07	0.03	-0.03	0.00

Table 3: The anisotropies of the PSFs described in Fig.39. The values quoted are the raw ellipticity measurements from KSB where the  $r_g$  from the stars was used as filter scale.

description how the PSF is constructed in SkyMaker is given in the Appendix of [16].

Before we can do any analysis, the detections of a real image have to be separated into stars and galaxies. This is done with a half-light-radius  $r_h$  vs.  $m$  plot and is summarised in Fig. 38.

### 6.7.2 Results of the SkyMaker simulations

Figure 39 shows contours of the outer and core parts of the PSFs that we have used in our simulations. The outer parts of the profiles all look the same (except for the quadratic PSF) and they differ mostly in the cores. The anisotropies caused by these PSFs are given in Table 3. For every PSF we have analysed seven sets of images with the shear combinations given in Table 4 where a few sets are slightly different from those in Table 2. The results for our SkyMaker simulations are shown in Figures 40, 41 and 42. As for the semi-analytical and simulated Gaussian profiles, we now discuss the accuracy of the shear estimators for the different shear amplitude and PSF anisotropies:

- In the case of no PSF anisotropy at all (PSF\_1), the tensor correction gives the correct shear on the whole range of  $|g| = 0.012$  up to  $|g| = 0.32$ . The scalar correction underestimates the shear by about 10% – 15%. The  $1\sigma$  error bars are about at  $\pm 0.005$ .

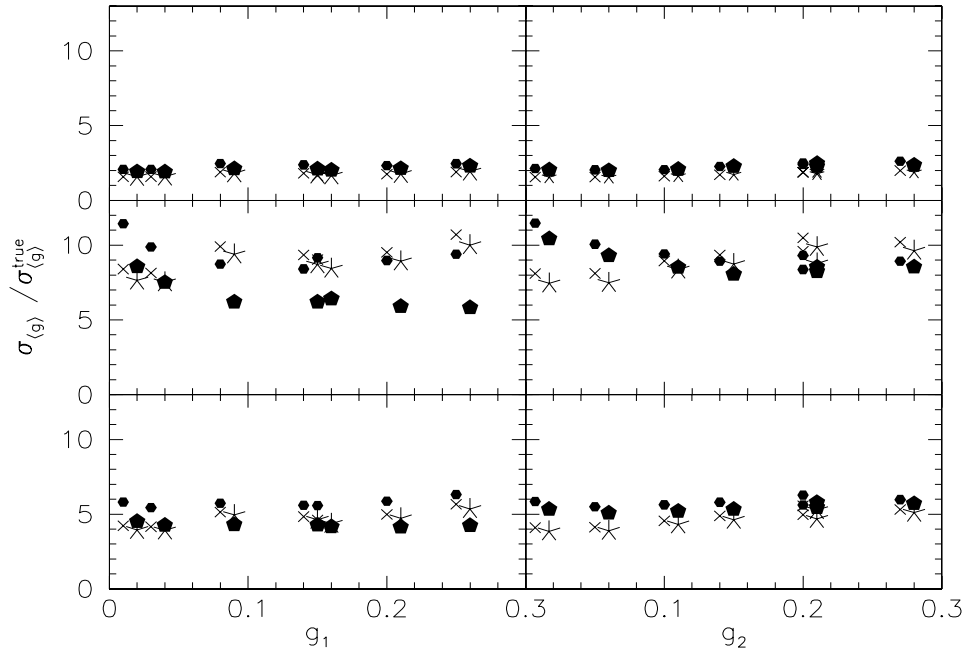


Figure 35: The noise properties of the shear estimations from Fig. 34 (Figure 9 from [16]): For high S/N objects, sky noise increases the noise in the shear estimates by a factor of 2, for low S/N objects by a factor 5-10 and about a factor of 5 for the whole sample.

$g_1$	0.01	0.03	0.25	0.08	0.2	0.2	0.14
$g_2$	0.007	0.05	0.2	0.14	0.27	0.14	0.2

Table 4: The shear combinations we investigated with our SkyMaker simulations.

- In the presence of PSF anisotropy, the tensor correction can over- or underestimate the true shear, while it is always underestimated in the scalar case. In the sense of deviation from the true value, the tensor case always gives the better result. The relative underestimates with the scalar correction can reach up to 30% in the worst case. As the anisotropy of stars is perfectly corrected by our polynomials for  $\mathbf{q}^*$  we investigated whether there is nevertheless an anisotropy residual in the galaxies. For this we compared the position angle from the input shear with the recovered one in Fig. 42. There we see that in the scalar case the position angle is nearly perfectly recovered. This means that the PSF anisotropy is corrected very well but the calibration factor  $P^g$  is too low. In the tensor case the recovery of the position angle is not as good as in the scalar case but also very acceptable. The reason for the over- or underestimates of the magnitude of the shear is probably that the elements of  $P^g$  that are used for the isotropic correction should be calculated from profiles that are already corrected for PSF anisotropy effects.
- Also for all the SkyMaker simulations, the over- or underestimates are nearly a constant fraction of the input shear. For the tensor estimator we can recover the input shear to an accuracy between 10% – 15%. This means that we can measure a shear below 0.1 nearly within 0.01, the accuracy required for accurate measurements of the cosmic shear on scales  $\leq 10'$  (see section 8).
- For all the SkyMaker simulations, the recovery of the shear is better than in the anisotropic

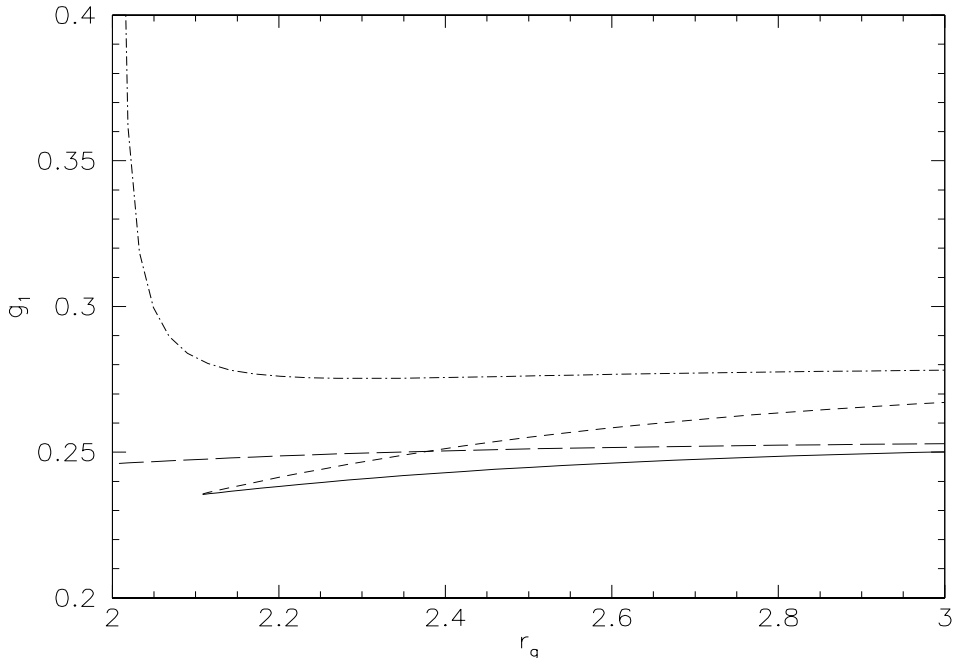


Figure 36: Analytical results for  $g_1 = 0.25$ , round intrinsic objects, no PSF anisotropy and a 10% PSF anisotropy in the  $g_1$  direction (Figure 10 from [16]): The solid and short dashed lines are scalar and tensor shear estimate for the PSF anisotropy free case, the long dashed and dashed dotted line represent scalar and tensor correction when PSF anisotropy is present. The figure confirms that the resulting shear is overestimated significantly in the tensor case with PSF anisotropy.

case with Gaussian profiles (see Fig. 37). It seems that the more realistic SkyMaker PSF profiles better reflect the assumptions in the KSB algorithm than Gaussian profiles, namely that PSF anisotropy mainly comes from the central core. While a Gaussian has its anisotropy on all scales, Fig. 39 shows that the anisotropy details of the SkyMaker PSFs indeed lie in the core.

- Regarding the issue of scalar vs. tensor correction, we can conclude that, although the scalar case always underestimates the true shear, it provides the more conservative answer and it should be used when one is interested mainly in the position angle of objects, while the tensor should be given preference when the amplitude of the shear is important.

## 6.8 Conclusions

We have investigated how well we can recover weak gravitational shear with realistic simulations for ground-based observations. With respect to the main motivation for this work, to test the reliability of our scientific results presented in the following sections, the results are very encouraging. In all our simulations we could recover weak shear up to  $|\mathbf{g}| = 0.1$  with an accuracy of 0.01 or better and we can significantly exclude the detection of a  $|\mathbf{g}| = 0.04 - 0.05$  false signal based purely on uncorrected PSF effects. We conclude that we clearly can trust the KSB algorithm for the detection of cosmic shear requiring accurate measurements of very weak shear on the percent level. Although the PSFs we have investigated certainly do not cover all possible defects mimicking a lens signal, other independent studies like Hoekstra et al. 1998 (see [26]), who looked at diffraction-limited model PSFs for the HST, or Bacon et al. (see [1]), indicate that the KSB algorithm works better than one could expect looking at its assumptions. We could show that for the

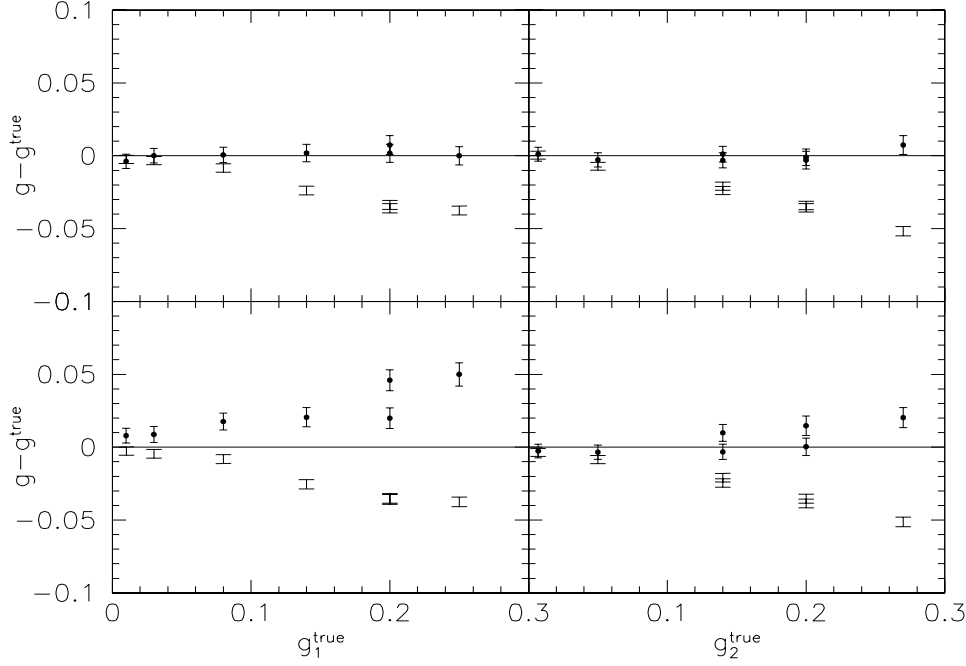


Figure 37: Results for Gaussian profiles put on FITS images and analysed with the procedure outlined in Appendix B (Figure 11 from [16]): Upper panels are for the PSF anisotropy-free case, and the lower panels for a PSF anisotropy of 10% in the  $g_1$  direction. All galaxies were used and no selection according to S/N was done. Dots with errorbars represent measurements with tensor estimation, lines with errorbars those with scalar estimation. The final estimates were done with about 13500 galaxies for every realisation and error bars have a typical size of 1%. The results are completely comparable to the  $S/N > 2$  case of Fig. 34

PSFs investigated here, the anisotropy correction proposed in the original KSB work is working very well, but that there are problems with the isotropy correction if an anisotropy is present. The reason for this is probably that in such a case, the boost factor  $P^g$  is calculated with the wrong, not anisotropy-corrected, galaxy profile. Moreover, we have presented a fully automatic procedure leading from the image frames to an object catalog for reliable shear measurements.

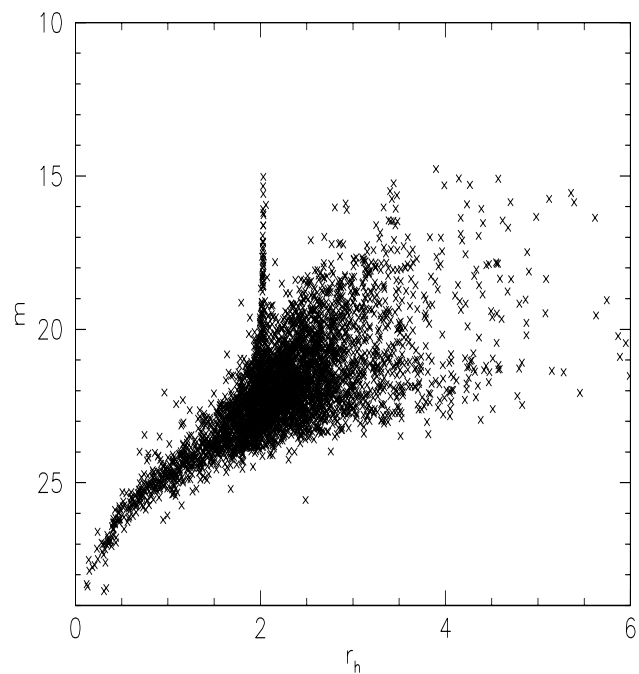


Figure 38: Half light radius vs. the magnitude (in arbitrary units) from all the detections in one of our SkyMaker simulations from section 6.7 (Figure 1 from [16]): Bright, unsaturated stars are clearly localised in a small region around  $r_h \approx 2$  and in  $15 < m < 20$ . Larger objects are considered as galaxies, while the smaller, very faint objects are probably noise detections. For objects around the stellar locus with faint magnitudes, a clear classification between stars and galaxies cannot be done and we typically exclude this region from our analysis.

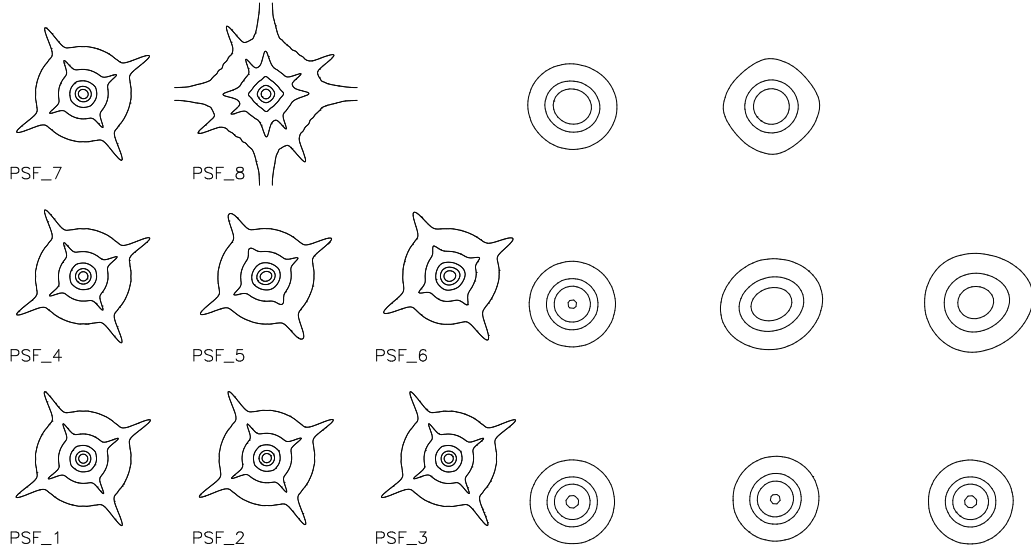


Figure 39: Outer part (left) and the core (right) of the PSFs used in the SkyMaker simulations (Figure 15 from [16]): All the PSFs contain four spider arms produced by the support of the secondary mirror. From left to right and from bottom to top the PSFs contain the defects [For the tracking errors, the first value in the bracket gives the rms drift in the direction of the second value. The third value is the rms of a Jitter in the orthogonal direction. For the optical aberrations the “d80” diameters introduced are quoted (“d80” is the diameter within which 80% of the light from a point source is enclosed because of the aberration, omitting the effects from diffraction; see the Appendix of [16] for a detailed description of tracking errors and aberrations). We note that the “d80” diameters of the individual aberrations do not add quadratically as the aberration polynomials are not orthogonal.]: (1) None; (2) tracking error ( $0.0''$ ,  $97.0^\circ$ ,  $0.05''$ ), defocusing ( $0.05''$ ), coma ( $0.26''$ ); (3) tracking error ( $0.05''$ ,  $0.0^\circ$ ,  $0.01''$ ), defocusing ( $0.07''$ ) and triangular aberration ( $0.015''$ ); (4) tracking error ( $0.02''$ ,  $34.0^\circ$ ,  $0.02''$ ), defocusing ( $0.1''$ ) and spherical aberration ( $0.3''$ ); (5) tracking error ( $0.3''$ ,  $31.0^\circ$ ,  $0.2''$ ), defocusing ( $0.3''$ ) and astigmatism ( $0.72''$ ); (6) tracking error ( $0.3''$ ,  $62.0^\circ$ ,  $0.3''$ ), defocusing ( $0.3''$ ), coma ( $0.31''$ ) and triangular aberration ( $0.72''$ ); (7) tracking error ( $0.5''$ ,  $-20.0^\circ$ ,  $0.11''$ ), defocusing ( $0.07''$ ) and triangular aberration ( $0.122''$ ) (8) defocusing ( $0.05''$ ) and quadratic aberration ( $0.61''$ )

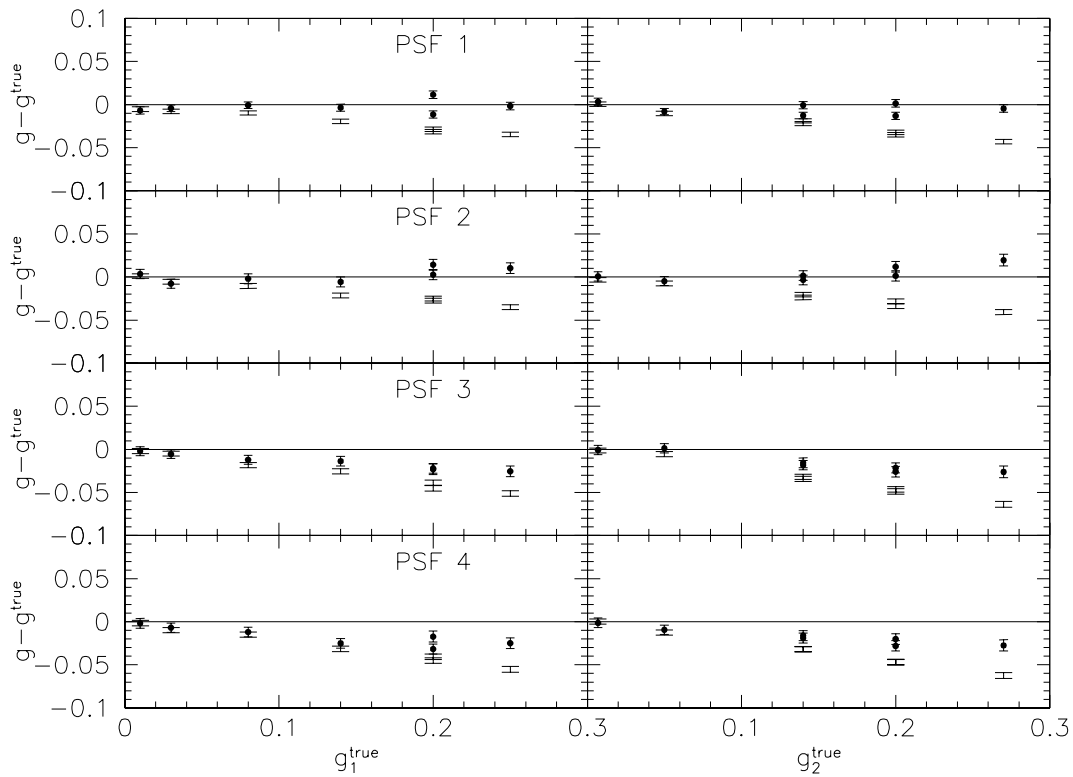


Figure 40: SkyMaker results I (Figure 16 from [16]): From top to bottom the figure shows SkyMaker results for the first four PSFs of Fig.39. The objects from the ten images for every shear/PSF combination were pasted together and the means and errorbars shown here were obtained from this merged catalog.

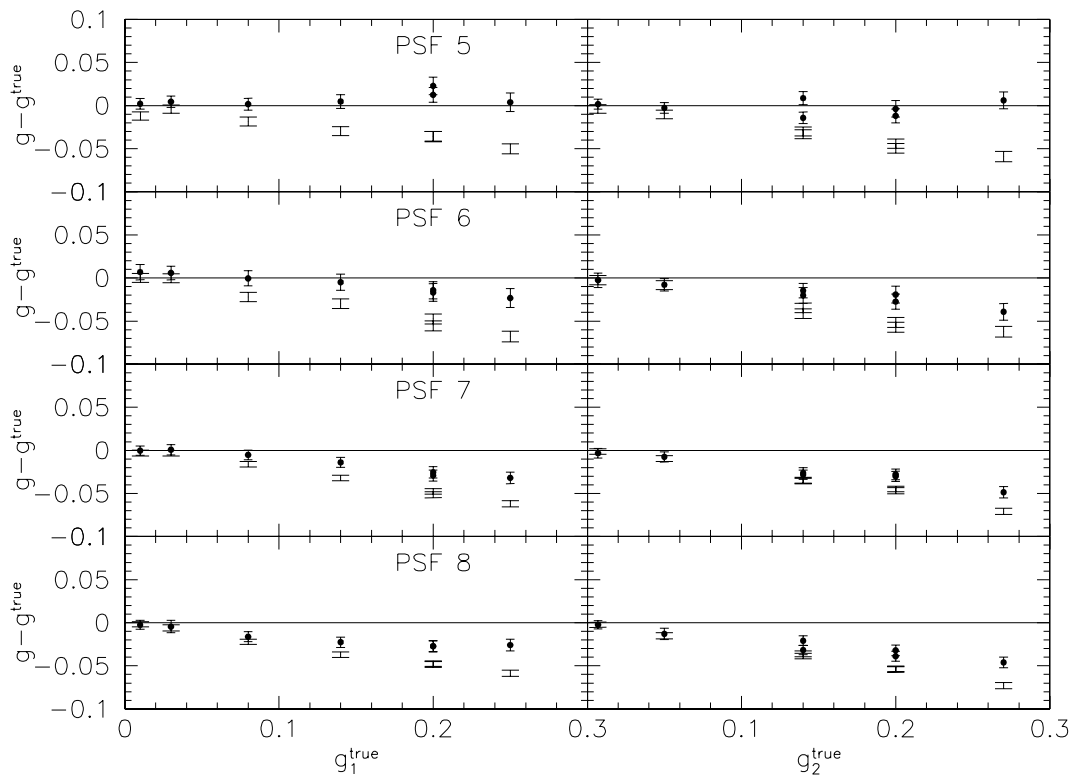


Figure 41: SkyMaker results II (Figure 17 from [16]): The same as Fig. 40 for the last four PSFs of Fig. 39.



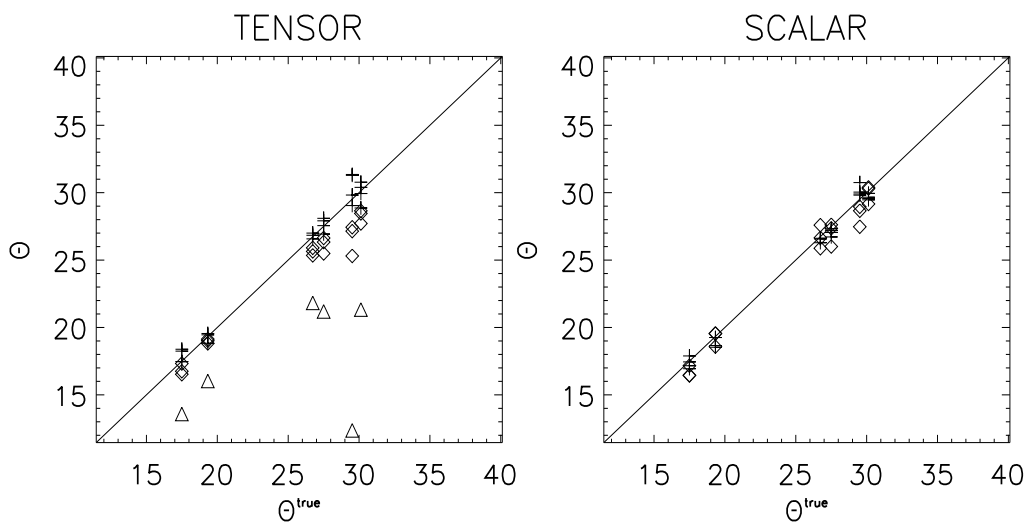


Figure 42: Position angles of the input shear vs. the recovered shear (Figure 18 from [16]): Crosses correspond to the results from Fig. 40 and diamonds to those from Fig. 41. Triangles are for shear values where no PSF anisotropy correction has been done (PSF\_4). In this plot we omitted the very weak shear case ( $g_1 = 0.01$ ,  $g_2 = 0.007$ ) as the position angle is badly determined here.

## 7 Detection of a dark matter concentration near the cluster Abell 1942

### Summary:

*In this section we report on the first detection of a dark matter concentration based on weak lensing techniques only. The discovery was made by chance when analysing deep wide-field imaging data around the cluster Abell 1942. We had high-quality lensing data around the cluster centre taken with two different cameras in two different years but in the same optical band ( $V$ ) at the CFHT telescope. This data set is ideal to investigate the stability of shear measurements (see the previous section) on real data and also to test different mass reconstruction techniques in detail. We routinely applied a weak lensing cluster finder developed by Peter Schneider to our data and discovered a massive object to which no obvious light is associated, about  $7'$  South of Abell 1942. As the detection was very close to the border of our data field we were not sure whether our signal could originate from border effects. The field was reobserved in the  $I$  band and our detection was confirmed in the new data. In the first part of the section we introduce the weak lensing cluster finder, the so-called  $M_{\text{ap}}$  statistics. We then turn to a more detailed description of our analysis and finally to physical properties of our dark mass candidate. The work presented in this section has been published in [17].*

### 7.1 Finding matter concentrations with weak gravitational lensing

We have seen in section 4.2 that the dimensionless surface mass density  $\kappa$  of a lens can be reconstructed with the observable  $\gamma$  [see eq. (91)] in the weak lensing regime. If we have a blank field with image ellipticities of our sources we could reconstruct the complete two-dimensional mass density and analyse this  $\kappa$  map for peaks that we associate with mass concentrations later. Although this strategy is feasible, there are more easy-to-apply techniques if we are primarily interested in detecting a mass concentration rather than analysing its distribution in detail. The basic idea is to go to an arbitrary point  $\vec{\theta}_0$  in our data field and investigate whether we have a significant masspeak in a circular aperture with radius  $\theta_{\text{in}}$  surrounding  $\vec{\theta}_0$ . Thus, we would like to calculate an integral of the form

$$\bar{\kappa}(\theta_{\text{in}}) = \frac{1}{\pi\theta_{\text{in}}^2} \int_A d^2\vartheta \kappa(\vec{\vartheta}) + \kappa_0. \quad (159)$$

Here,  $\bar{\kappa}(\theta_{\text{in}})$  denotes the mean surface mass density within the circular aperture with radius  $\theta_{\text{in}}$  and  $A$  is the area of it. For simplicity we have put  $\vec{\theta}_0$  to the origin of our coordinate system. Unfortunately,  $\bar{\kappa}$  cannot be calculated with eq. (159) directly, as the integral on the right hand side is only determined up to a constant ( $\kappa_0$ ), making the result meaningless. To eliminate this constant, we consider the quantity

$$\zeta(\theta_{\text{in}}, \theta_{\text{out}}) := \bar{\kappa}(\theta_{\text{in}}) - \bar{\kappa}(\theta_{\text{in}}, \theta_{\text{out}}). \quad (160)$$

Here,  $\bar{\kappa}(\theta_{\text{in}}, \theta_{\text{out}})$  is the mean mass density in the annulus between  $\theta_{\text{in}}$  and  $\theta_{\text{out}}$  with  $\theta_{\text{in}} < \theta_{\text{out}}$ . As  $\bar{\kappa}(\theta_{\text{in}}, \theta_{\text{out}})$  is non-negative,  $\zeta(\theta_{\text{in}}, \theta_{\text{out}})$  provides a lower bound on  $\bar{\kappa}(\theta_{\text{in}})$ . We cannot obtain more than this lower limit.

As in section 4.2, also this so-called  $\zeta$ -statistics (introduced by Kaiser; see [35]) for a lower limit of the mean mass within an aperture  $\theta_{\text{in}}$  can be expressed in terms of the observable  $\gamma$ . For this, we first use the Gauss Integral theorem in its two-dimensional form

$$\int_A d^2x \left( \frac{\partial \vec{B}}{\partial x} + \frac{\partial \vec{B}}{\partial y} \right) = \oint_{\partial A} B_x dy - B_y dx, \quad (161)$$

where  $\vec{B}$  is a two-dimensional vector field,  $\partial A$  denotes the border of our aperture and  $B_x$  and  $B_y$  are the two components of  $\vec{B}$ . We see from eq. (51) and eq. (52) that  $\kappa = 0.5\nabla\vec{\alpha}$ , so choosing

$\vec{B} = \vec{\alpha}$  in eq. (161), we get  $\int_A 2\kappa(\vec{\theta}) d^2\theta$  on the left hand side. For the right hand side, we get after introducing polar coordinates, using the relations of eq. (54) and performing partial integration

$$\oint_{\theta_A} \alpha_1 dy - \alpha_2 dx = \theta_{\text{in}} \left( \int_0^{2\pi} \kappa(\theta_{\text{in}}, \varphi) d\varphi - \int_0^{2\pi} [\gamma_1(\theta_{\text{in}}, \varphi) \cos(2\varphi) + \gamma_2(\theta_{\text{in}}, \varphi) \sin(2\varphi)] d\varphi \right). \quad (162)$$

Equation (161) then reads with  $\vec{B} = \vec{\alpha}$

$$\begin{aligned} \bar{\kappa} &= \frac{1}{\theta_{\text{in}}^2 \pi} \int_A d^2\theta \kappa(\theta) = \frac{1}{2\pi\theta_{\text{in}}} \left( \int_0^{2\pi} \kappa(\theta_{\text{out}}, \varphi) d\varphi - \int_0^{2\pi} [\gamma_1(\theta_{\text{in}}, \varphi) \cos(2\varphi) + \gamma_2(\theta_{\text{in}}, \varphi) \sin(2\varphi)] d\varphi \right) \\ &=: \langle \kappa(\theta_{\text{in}}) \rangle + \langle \gamma_t(\theta_{\text{in}}) \rangle, \end{aligned} \quad (163)$$

where the last equation defines the so-called *tangential shear*. This equation relates the total mean mass within  $\theta_{\text{in}}$  to the mean boundary mass  $\langle \kappa(\theta_{\text{in}}) \rangle$  and the mean tangential component of the shear  $\langle \gamma_t(\theta_{\text{in}}) \rangle$  on the border of the aperture. As  $\langle \kappa(\theta_{\text{in}}) \rangle > 0$ , a lower bound for  $\bar{\kappa}(\theta_{\text{in}})$  is already given by the mean tangential shear on the boundary  $\langle \gamma_t(\theta_{\text{in}}) \rangle$ .

The dependence on  $\langle \kappa(\theta_{\text{in}}) \rangle$  can be eliminated by noting

$$\bar{\kappa}(\theta_{\text{in}}) = \frac{2}{\theta_{\text{in}}^2} \int_0^{\theta_{\text{in}}} \theta d\theta \langle \kappa(\theta) \rangle \Rightarrow \frac{d\bar{\kappa}}{d\theta} = -\frac{2}{\theta} \bar{\kappa} + \frac{2}{\theta} \langle \kappa(\theta) \rangle; \quad \theta \in [0; \theta_{\text{in}}]. \quad (164)$$

Inserting eq. (163) into the last equation, we derive the following differential equation

$$-\frac{2 \langle \gamma_t(\theta) \rangle}{\theta} = \frac{d\bar{\kappa}(\theta)}{d\theta}, \quad (165)$$

relating  $\bar{\kappa}$  directly in terms of  $\gamma_t$ . Integrating from  $\theta_{\text{in}}$  to  $\theta_{\text{out}}$ , finally an expression for  $\zeta(\theta_{\text{in}}, \theta_{\text{out}})$  in terms of the observable  $\gamma$  is obtained

$$\zeta(\theta_{\text{in}}, \theta_{\text{out}}) = \bar{\kappa}(\theta_{\text{in}}) - \bar{\kappa}(\theta_{\text{out}}, \theta_{\text{in}}) = -\frac{2\theta_{\text{out}}^2}{\theta_{\text{out}}^2 - \theta_{\text{in}}^2} \int_{\theta_{\text{in}}}^{\theta_{\text{out}}} \frac{d\theta}{\theta} \langle \gamma_t(\theta) \rangle. \quad (166)$$

The dependence of the  $\bar{\kappa}$  on the tangential shear is very intuitive when we take another look at Fig. 7. We saw there that weak lensing tends to align galaxies preferentially tangential to a massive object.

The  $\zeta$ -statistics can also be written as

$$\zeta(\theta_{\text{in}}, \theta_{\text{out}}) = \int_0^{\theta_{\text{out}}} d\theta W_\zeta(\theta) \kappa(\vec{\theta}), \quad (167)$$

with

$$W_\zeta(\theta) = \begin{cases} \frac{1}{\pi\theta_{\text{in}}^2} & : 0 \leq \theta < \theta_{\text{in}} \\ -\frac{1}{\pi(\theta_{\text{out}}^2 - \theta_{\text{in}}^2)} & : \theta_{\text{in}} \leq \theta \leq \theta_{\text{out}} \end{cases}. \quad (168)$$

Hereby,  $W_\zeta$  is a *compensated filter*, i.e.  $\int_0^{\theta_{\text{out}}} d\theta \theta W_\zeta(\theta) = 0$ . It can be shown (Schneider 1996; see [59]) that the weight function  $W_\zeta$  can be replaced by an arbitrary compensated filter function  $U$ . This leads to the so-called  $M_{\text{ap}}$  statistics

$$M_{\text{ap}}(\theta_{\text{out}}) := \int d^2\theta \kappa(\vec{\theta}) U(|\vec{\theta}|) \quad (169)$$

$$= - \int d^2\theta \gamma_t(\vec{\theta}) Q(|\vec{\theta}|). \quad (170)$$

Once  $U(\eta)$  is specified,  $Q(\eta)$  is given by

$$Q(\eta) = \frac{2}{\eta^2} \int_0^\eta dx x U(x) - U(\eta). \quad (171)$$

For the current work we have used the filter

$$U(\eta) = u_l(\eta/\theta_{\text{out}})/\theta_{\text{out}}^2; \quad u_l(\rho) = \frac{(l+2)^2}{\pi}(1-\rho^2)^l \left( \frac{1}{l+2} - \rho^2 \right), \quad (172)$$

which gives

$$Q(\eta) = q_l(\eta/\theta_{\text{out}})/\theta_{\text{out}}^2; \quad q_l(\rho) = \frac{(l+1)(l+2)}{\pi}\rho^2(1-\rho^2)^l. \quad (173)$$

We used  $l = 1$  throughout. This results in  $\theta_{\text{out}} = \sqrt{3} \theta_{\text{in}}$ .

- The above formulae can easily be applied to real data where we have measurements of the reduced shear  $\mathbf{g}$ , and so for  $\gamma$  in the weak lensing regime, only on discrete points. We consider the quantity

$$\hat{M}_{\text{ap}}(\theta_{\text{out}}) = -\frac{\pi\theta_{\text{out}}^2}{N} \sum_i g_t(\vec{\theta}_i) Q(|\vec{\theta}_i|), \quad (174)$$

where the sum extends over all  $N$  galaxies within our aperture  $\theta_{\text{out}}$ , and the tangential component  $g_t(\vec{\theta})$  of the image ellipticity is defined in analogy to  $\gamma_t$  (see eq. (163)). As we deal with observations, we use from now on  $-\mathbf{g}$  (see eq. (151)) as the unbiased estimator for the reduced shear instead of the theoretical  $-\epsilon$ .  $\hat{M}_{\text{ap}}$  is an unbiased estimator for  $M_{\text{ap}}$ . But even if we calculate  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  in a region with strong lensing features, and since we are primarily concerned about the detection of mass, we can consider  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  in its own right, representing the tangential alignment of galaxy images with respect to the point  $\vec{\theta}_0$ . Our criterion for a mass detection is then how probable the tangential alignment we are observing comes from a random distribution of galaxy ellipticities, i.e. if there is no massive object present. This representation also remains valid if the aperture is centred on a position which is less than  $\theta_{\text{out}}$  away from the boundary of the data field, so that part of the aperture is located outside the data field.

- As mentioned in the last paragraph, we want to evaluate the significance of a mass detection by the probability that the observed value for  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  is drawn from a random distribution of galaxy ellipticities. This significance can be evaluated by randomising the position angles of the galaxies, calculate  $\hat{M}_{\text{ap}}(\theta_{\text{out}})_{\text{rand}}$ , repeating this exercise  $N_{\text{rand}}$  times and see how often  $\hat{M}_{\text{ap}}(\theta_{\text{out}})_{\text{rand}}$  exceeds the original measurement  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$ . We call this number  $M$  and the *errorlevel*  $\nu$  of a detection then reads

$$\nu := \frac{M}{N_{\text{rand}}}. \quad (175)$$

This quantity directly gives the probability to observe  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  with a random distribution of galaxy ellipticities. On an empty wide-field observation we can calculate  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  on a regular grid and determine the errorlevel of a mass detection at these points with the above-described procedure.

We now describe the data to which we applied this analysis.

## 7.2 The data

As already mentioned in the summary we have used two different data sets in different optical bands  $V$  and  $I$  that were taken with two different cameras, in different years, for our analysis. Both data sets were obtained with the CFHT telescope on Hawaii.

- The  $V$ -band observations were done with the MOCAM camera in June 1995. The MOCAMer consists of a quadratic mosaic of  $4 \times 2K \times 2K$  chips with a pixel scale of  $0.206''$ . This results in a total field of view of  $14' \times 14'$  and MOCAM was the largest mosaic camera at that time. The data set consists of 9 exposures a 1800 sec. The final coadded 4h30min image has an exceptional  $V$  band seeing of  $0.74''$ .

- The  $I$ -band data were taken with the UH8K camera consisting of a rectangular mosaic of 8  $2K \times 4K$  CCDs, in May 1998. For our analysis we used only one single chip out of the  $28' \times 28'$  mosaic. The final image, consisting of 9 exposures a 1200 sec., has a seeing of  $0.67''$ .

Figure 43 shows the relative geometry of the data. We also have a  $V$ -band image from the Faint

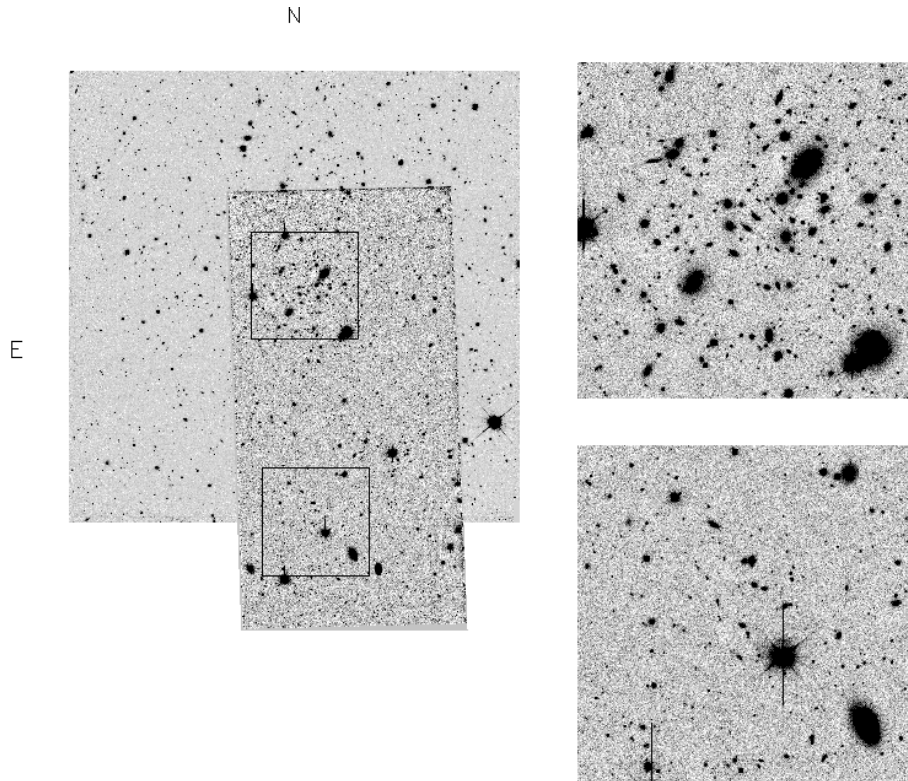


Figure 43: The geometry of the optical data we had around A1942 (Figure 1 from [17]): The left-hand side shows the area of the  $V$ -band MOCAM field (square) and the  $I$ -band UH8K-chip data (rectangle). The framed regions are  $3.3 \times 3.3$  cutouts around the cluster centre of A1942 and around our ‘dark clump’ candidate that we will discuss in section 7.3. These regions are zoomed in on the right-hand side. The ‘dark clump’ region is centred around  $\alpha(J2000)=14^{\text{h}} 38^{\text{m}} 22.59^{\text{s}}$ ;  $\delta(J2000)=03^{\circ} 32' 32.22''$ .

Object CAMera (FOCAM) but as it only covers the very central part of the cluster Abell 1942 (total field-of-view:  $7' \times 7'$ ) and it did not contain our dark clump candidate, we did not use it for this work. From the images we produced catalogs containing positions and ellipticities in a similar way as described in section 6. As our work on the simulations of the KSB algorithm only had started, we used a slightly modified way to arrive at the final catalogs:

- We created first an initial catalog with “hfindpeaks” and calculated lensing quantities with the “analyse” program. To get rid of problematic objects these catalogs were crosscorrelated with those produced by the SExtractor package\*. This program is optimised for the photometric analysis of objects, but not for lensing studies. Moreover it analyses objects for possible defects like blending. These objects were removed from the SExtractor catalog before the crosscorrelations.

---

\*SExtractor is an object detection and analysis tool developed by Emmanuel Bertin. It detects objects by searching for  $N$  connected pixels having  $k\sigma_{\text{sky}}$ .  $N$  and  $k$  can be defined by the user. The program is extremely fast and only requires very moderate computer resources. See [6] for a more detailed description of the software.

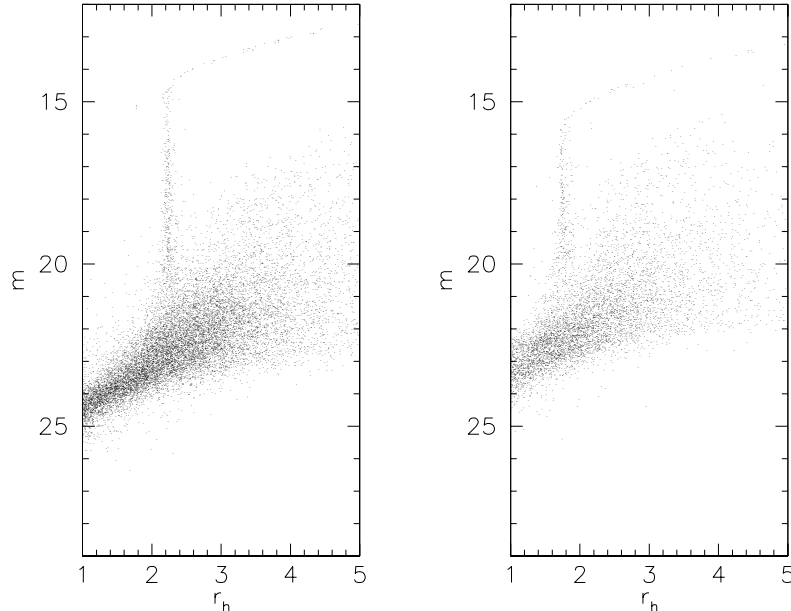


Figure 44: Magnitude as a function of half-light radius  $r_h$  for all objects detected in the MOCAM frame (left) and the UH8K-chip (right): The object tail on top of the stellar locus contains saturated stars. The magnitudes are in an arbitrary system. In both cases we can clearly identify a prominent sequence of stellar objects at about  $r_h = 2.2$  for MOCAM, and  $r_h = 1.75$  for UH8K-chip.

- From this “cleaned” catalog we selected stellar candidates and we performed the corrections for PSF anisotropy and PSF smearing as described in section 6. See also Figs. 44, 45 and 46.
- We did not introduce a weighting scheme into our shear calculations but rejected objects whose final corrected modulus of the tensor ellipticity was larger than unity. We ended up with 4190 objects ( $V > 22.0$ ) for the MOCAM image (21 objects per square arcmin) and 1708 objects ( $I > 21.0$ ; 17 objects per square arcmin) for the UH8K-chip.

We note that the final catalogs from the MOCAM and the UH8K-chip were not crosscorrelated at any stage of the analysis but everything was done completely independent for both data sets.

### 7.3 The detection of a dark cluster candidate

When we first analysed the MOCAM field in May 1997 we started with the comparison of several cluster mass reconstructions. One of these reconstructions is shown in Fig. 47. We clearly pick up the mass of A1942, but the map also reveals several other peaks. As most of these peaks are lying close to the border of the data field they are very likely to originate from the higher noise expected due to boundary effects. Nevertheless, the application of the  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  statistics should also clarify the significance of these peaks lying near the boundary of the data (see above). The application of the  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  statistics to the MOCAM field, where adjacent gridpoints had a separation of  $13''$ , led to the significance maps shown in Fig. 48. Displayed are contours of constant  $\nu$  for four different filter scales  $\theta_{\text{out}}$ . As can be seen, the cluster centre shows up prominently in the  $\nu$ -map on all scales. In addition, two highly significant peaks show up, one at the upper right corner, the other  $\sim 7'$  South of the cluster centre, close to the edge of the MOCAM field. We have verified the robustness of this Southern peak by using SExtractor ellipticities instead of those from imcat, and found both the cluster components and the Southern peak also with that catalog (although

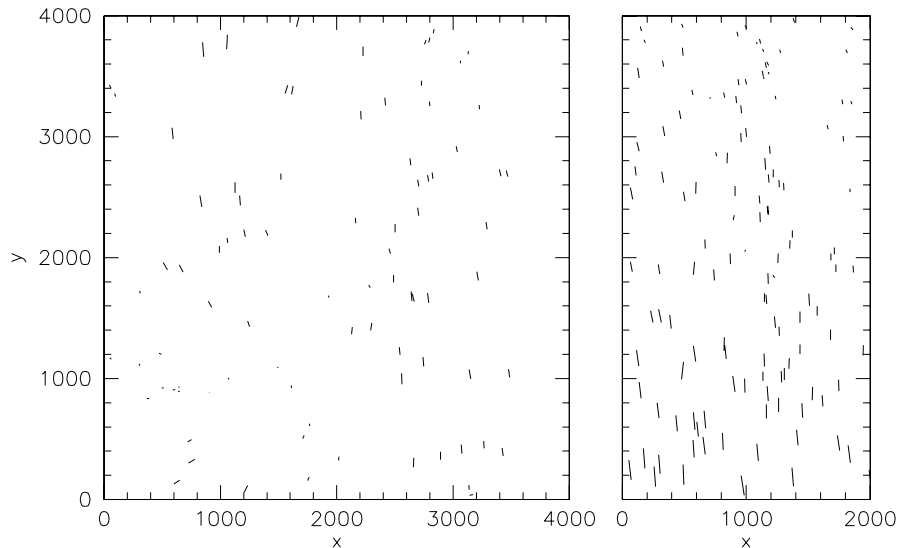


Figure 45: The ellipticity fields for stars for the  $V$ -band MOCAM field (left panel) and the  $I$ -band UH8K-chip containing the cluster (Figure 3 from [17]): Both fields show a smooth variation and can be easily modelled by a low-order polynomial. The maximal ellipticity is about 5% for the MOCAM and 8% for the UH8K-chip.

these ellipticities were not corrected for any PSF effects; a sign for the superb quality of the data). Despite these findings we were still sceptical about the physical origin of the significant Southern peak. Nearly two years after these first findings we analysed the  $I$ -band UH8K-chip. Abell 1942 and the Southern peak from the MOCAM field are both located on one chip of the mosaic. A cluster mass reconstruction for this chip is shown on the right of Fig. 47. In this reconstruction the cluster and our dark clump candidate show up. Here, our candidate is located in the middle of the data field so that boundary effects play no role. The application of the  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  statistics to this data led to the errorlevels in Fig. 49. Hereby the same grid-spacing as in the MOCAM data ( $13''$ ) was used. After this confirmation of our dark clump candidate in the  $I$ -band data we became confident in it, and we started investigating its physical properties.

## 7.4 Properties of our dark clump candidate

- Figures 48 and 49 show significance contours for  $N_{\text{rand}} = 5000$  randomisations of the background galaxies. In the MOCAM data, the  $\hat{M}_{\text{ap}}(\theta_{\text{out}})_{\text{rand}}$  was never larger than the measured value. So we first calculated the true significance of our clump candidate. Concentrating on the position with the lowest  $\nu$  value (that we consider as the centre of our clump) we performed  $2 \times 10^6$  randomisations. The complete probability distribution of these randomisations is shown on the left-hand-side of Fig. 50 as solid (MOCAM) and dashed curve (UH8K) in Fig. 50. These curves nearly coincide and can be very well approximated by a Gaussian. The value of  $\hat{M}_{\text{ap}}(160'')$  at the dark clump is 0.0395 for MOCAM, and 0.0283 for the UH8K-chip. The fact that these two values are different is not problematic for two reasons: First, for the UH8K-chip, the whole aperture fits inside the data field, whereas it is partially outside for MOCAM; hence, the two values of  $\hat{M}_{\text{ap}}$  measure a different tangential alignment. Second, the two data sets use galaxies selected in different wavebands, their redshift distribution can be different, yielding different values of the resulting lens strength. The probability that a randomisation of image orientations yields a value of  $\hat{M}_{\text{ap}}$  larger than the observed

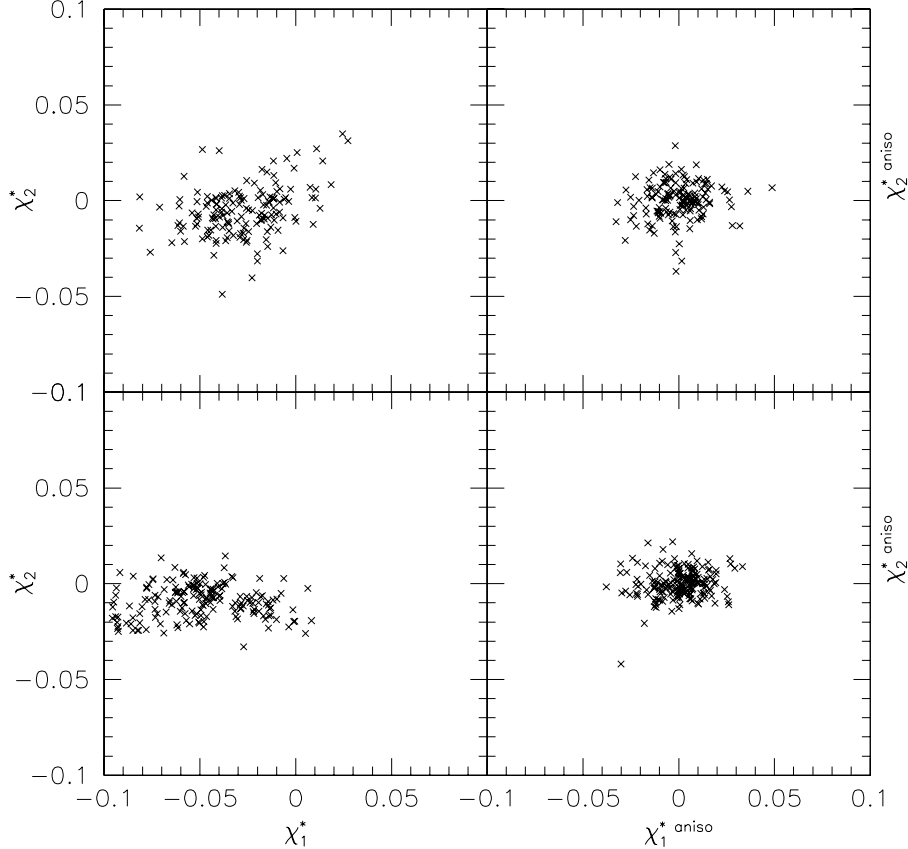


Figure 46: Stellar ellipticities (Figure 4 from [17]): The left panels show the raw imcat ellipticities from bright, unsaturated foreground stars in our fields (upper panels: MOCAM field; lower panels: UH8K-chip). The right panels show the ellipticities after they have been corrected with a second-order polynomial as described in the text. The rms of the ellipticities after correction is typically 0.015.

one is  $\sim 10^{-6}$  for the MOCAM field, and  $4.2 \times 10^{-4}$  for the UH8K-chip. Next we investigate whether the highly significant value of  $\hat{M}_{\text{ap}}$  at the dark clump comes from only a few galaxy images that have a high ellipticity and just happen to be tangentially aligned to the clump centre. For this, the sample of galaxy images inside the aperture was bootstrap resampled<sup>†</sup>, to obtain the probability  $p_{\text{boot}}(\hat{M}_{\text{ap}})$  that this resampling yields a particular value of  $\hat{M}_{\text{ap}}$ . This probability is also shown in Fig. 50. The probability that the bootstrapped value of  $\hat{M}_{\text{ap}}$  is negative is  $3.8 \times 10^{-4}$  for the UH8K-chip, and  $< 10^{-6}$  for the MOCAM field.

- The fact that the signal does not originate from a few galaxies only is supported by the radial dependence of the tangential shear component shown in Fig. 51. It is reassuring that the radial behaviour of  $\langle g_t \rangle$  is very similar on the two data sets. In fact, owing to the different wavebands of the two data fields and the fact that the aperture does not fit inside the MOCAM field, this agreement is better than one might expect. The mean tangential ellipticity is positive over a large angular range; except for one of the inner bins (for which the error bar is fairly large),  $\langle g_t \rangle$  is positive in all bins for  $\theta \lesssim 150''$ . This figure thus shows

<sup>†</sup>Out of the sample of  $N$  galaxies lying within the  $\hat{M}_{\text{ap}}$  aperture we randomly took  $N$  galaxies allowing for galaxies to be drawn several times. The fraction of galaxies that appear in the sample more than one time is about 17%. We recalculated  $\hat{M}_{\text{ap}}$  with this sample and repeated the whole procedure  $N$  times.



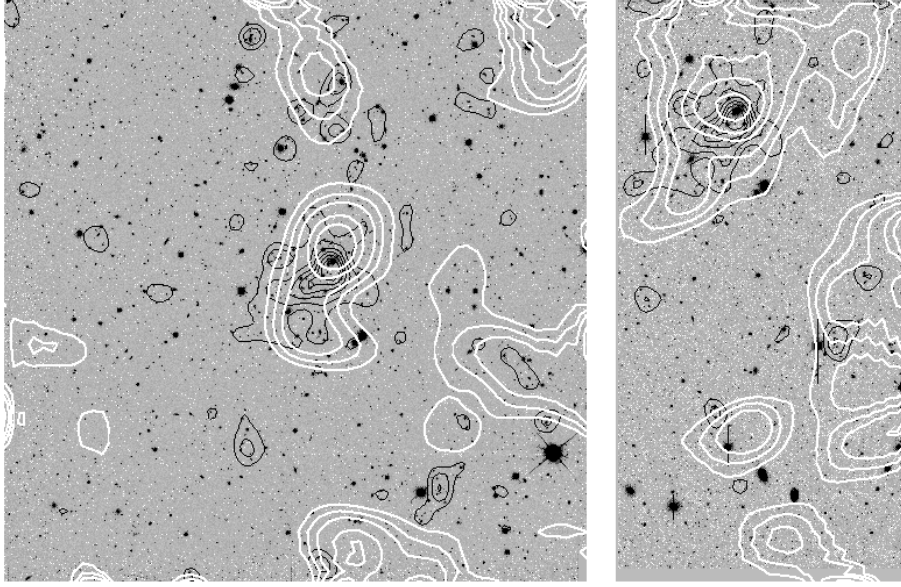


Figure 47: Lens mass reconstruction and light distribution (Figure 5 from [17]): The figure shows the mass reconstructions and the galaxy number densities from the MOCAM field (left panel) and the UH8K-chip (right panel). The white contours show  $\kappa = 0.03, 0.05, 0.07, 0.1, 0.12, 0.15, 0.17$  and  $0.2$ . The black contours show the smoothed galaxy light distribution from all galaxies brighter than  $V = 21.0$  and  $I = 20.0$

that the large and significant value of  $\hat{M}_{\text{ap}}$  at the dark clump is not dominated by galaxy images at a particular angular separation.

- We next argue that it is very unlikely for our object to lie at a redshift larger than unity if the lensing signal comes from a massive object at a well determined redshift. In the  $I$ -band we used galaxies in a magnitude interval  $21.0 < I < 24.5$ . In this magnitude range we would expect about 30 galaxies/arcmin<sup>2</sup>. With 17 galaxies/arcmin<sup>2</sup> we have used approximately half of them as putative background galaxies in our analysis. The median of simulated redshift distributions that extend the CFRS data (Lilly et al. 1995; see [42]) to fainter magnitude limits (see Fig. 52) is at about  $z \approx 0.7 - 0.8$ . If we assume that all our galaxies lie in the tail of this distribution, then  $z = 1.0$  represents a good upper limit for the redshift of our clump.
- Next we determine a lower limit for the mass-to-light ratio. For a crude mass estimate, we describe the mass profile by a singular, isothermal sphere and we calibrate it with Fig. 51. The velocity dispersion  $\sigma_v$  of the clump is then given by

$$\left(\frac{\sigma_v}{c}\right)^2 = \frac{1}{2\pi} \langle \gamma_t \theta \rangle \left\langle \frac{D_{\text{ds}}}{D_s} \right\rangle^{-1}, \quad (176)$$

where the product  $\gamma_t \theta$  would be independent of  $\theta$  for an isothermal sphere model, and the final term is the ratio lens-source to observer-source distance, averaged over the background galaxy population. Although the tangential shear appears to be fairly small close to the centre position of the clump (Fig. 51), there is a region between  $\sim 50''$  and  $\sim 150''$  where the tangential shear is clearly positive and decreases smoothly with radius. Thus we use the value of  $\langle \gamma_t \rangle$  at  $100''$  for our calibration of  $\gamma_t \theta$  in eq. (176). Inserting numerical values into eq. (176) yields:

$$\sigma_v = 1135 \sqrt{\frac{\gamma_t(100)}{0.06}} \sqrt{\frac{1}{3 \langle D_{\text{ds}}/D_s \rangle}} \text{ km/s}, \quad (177)$$

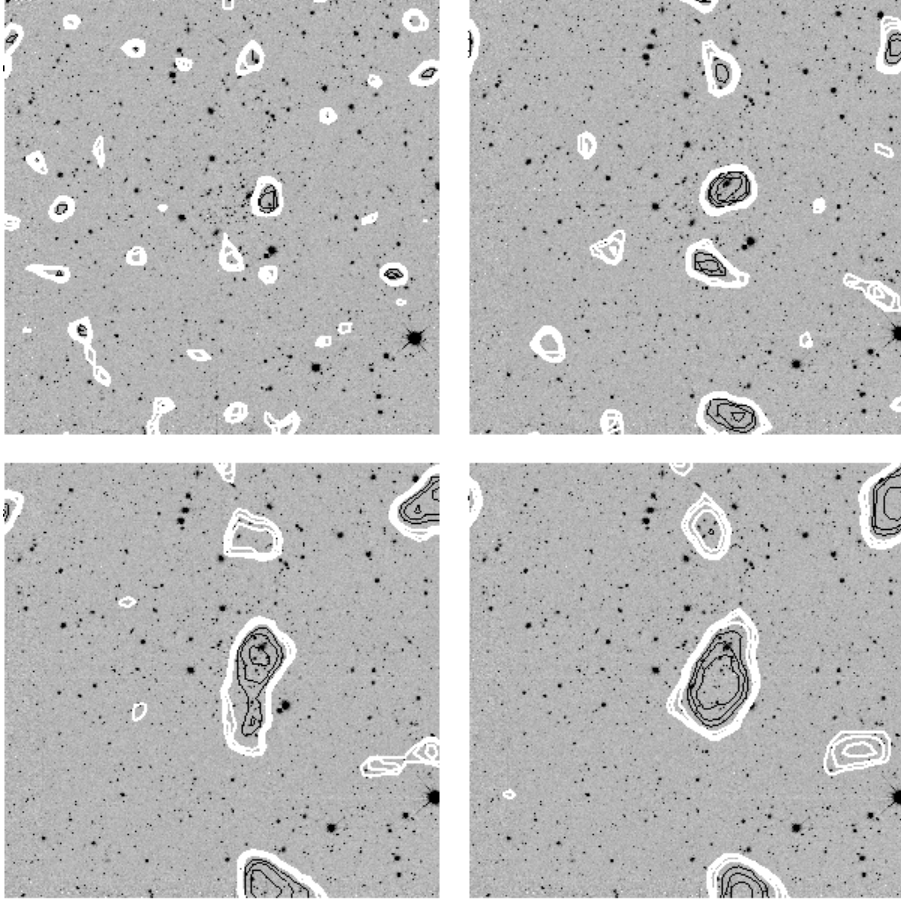


Figure 48: Error levels  $\nu$  for the MOCAM data (Figure 6 from [17]): The four panels show the significance  $\nu$  (see text) of the  $\hat{M}_{\text{ap}}(\theta_{\text{out}})$  maps of the MOCAM field. We chose  $N_{\text{rand}} = 5000$ , the black contours mark areas with  $\nu = 1, 10, 30/5000$  and the white contours  $\nu = 100, 180, 260/5000$ . The filter scales  $\theta_{\text{out}}$  are  $80''$  (upper left panel),  $120''$  (upper right panel),  $160''$  (lower left panel) and  $200''$  (lower right panel). For the larger scales the cluster components and the dark clump are detected with a very high significance.

where  $\gamma_t(100)$  is the tangential shear at  $100''$  from the mass centre. Alternatively, we can express this result in terms of the mass within a sphere of radius  $R$ ,  $M(< R) = 2\sigma_v^2 R/G$ ; for example, within  $R = 0.5h^{-1}$  Mpc, we find

$$M(< 0.5h^{-1} \text{ Mpc}) = 2.9 \times 10^{14} h^{-1} M_{\odot} \frac{\gamma_t(100)}{0.06} \frac{1}{3 \langle D_{\text{ds}}/D_s \rangle}. \quad (178)$$

Whereas this model is quite crude, the largest uncertainty in a quantitative mass estimates comes from the unknown redshift of the dark clump and the unknown redshift distribution of the background galaxy population. The mass is a monotonically increasing function of the lens redshift and depends very strongly on the assumed mean source redshift, in particular for values of  $z_d \geq 0.5$  for the clump. We show this in the upper left panel of Fig. 53. With the  $I$ -band data we now estimated the light coming from the dark clump. For this we created a SExtractor catalog counting every connected area with at least 3 pixels  $0.5\text{-}\sigma$  above the sky background as a potential object. The flux of all these objects (except from obvious stars) in a circle of  $100''$  radius around the clump centre was summed up. We did the same in 32 control circles around ‘empty’ regions in the other UH8K-chips. It turned out that the flux within the clump region is compatible with the mean flux of the control annuli, i.e., there

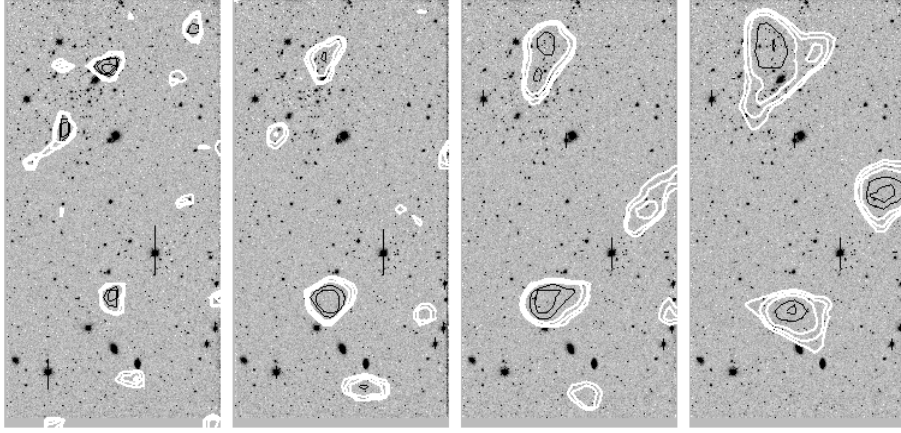


Figure 49: Error levels  $\nu$  for the UH8K-chip (Figure 7 from [17]): The same as Fig. 48 for the UH8K-chip  $I$ -band data. The filter scales are, from left to right:  $80''$ ,  $120''$ ,  $160''$  and  $200''$ . A1942 and the dark clump are also detected here with a very high significance.

is no overdensity of light at the position of the dark clump. See also the light distributions plotted in Fig. 53. Therefore, we took the  $1\text{-}\sigma$  fluctuation of the fluxes in the control circles as a reasonable upper limit for the light coming from the dark clump. For converting the apparent flux (or equivalently the apparent magnitude  $m_I$ ) into a total  $I$ -band magnitude  $M_I$ , we used the formula

$$m_I = M_I + 5 \log_{10} \left[ \frac{D_{\text{lum}}(z)}{10\text{pc}} \right] + 2.5 \log_{10}(1+z) - K_{\text{corr}} + ZP + \epsilon. \quad (179)$$

We assumed that we are dominated by elliptical galaxies, using  $K_{\text{corr}}$  corrections for this galaxy type calculated with the latest version of the Bruzual & Charlot stellar population synthesis models for the spectrophotometric evolution of galaxies (Bruzual & Charlot 1993; see [11]).  $D_{\text{lum}}(z)$  is the luminosity distance and  $ZP$  our magnitude zeropoint. We did not consider luminosity evolution ( $\epsilon = 0$ ). From the total  $I$ -band magnitude we derived a bolometric magnitude and a bolometric luminosity using

$$M_{\text{bol}} = M_I + \alpha(V - I). \quad (180)$$

The value of  $\alpha$  is obtained by the corresponding relation for the Sun which gives  $\alpha = 0.901$ . For estimating  $V - I$  we also calculated the total flux in  $100''$  annuli of the MOCAM data in the same way as above.  $V - I$  was then estimated by  $\langle V \rangle - \langle I \rangle$ , where the means were taken over all annuli. This resulted in  $V - I = 1.98$  and finally in an estimate for the bolometric luminosity. With a lower limit for the mass and an upper limit for the luminosity we can give lower limits for the mass-to-light ratio of our object. This is shown in Fig. 53 for different source redshift distributions and two cosmologies. We see that the EdS universe gives fairly high  $M/L$  estimates in comparison to a  $\Omega_0 = 0.3$ ,  $\Omega_\Lambda = 0.7$  model. When we assume a redshift of  $z_d \approx 0.8$  for our clump we obtain a lower limit of  $M/L \approx 300$  in the  $\Omega_\Lambda$  cosmology. This is a conservative lower limit which could be lowered significantly only if one assumes that the redshift distribution of the faint galaxies extends to substantially higher redshift.

- Since our dark clump candidate has a total mass typical of a massive cluster which are typically strong X-ray emitters we retrieved ROSAT data around A1942 from the NASA/GSFC HEASARC archive. This patch of the sky was observed with the ROSAT HRI in August 1995 with a total integration time of 44515s. Figure 54 shows the X-ray flux overlayed on the UH8K-chip. We see a prominent X-ray emission around A1942 and a faint  $3.2\sigma$  X-ray

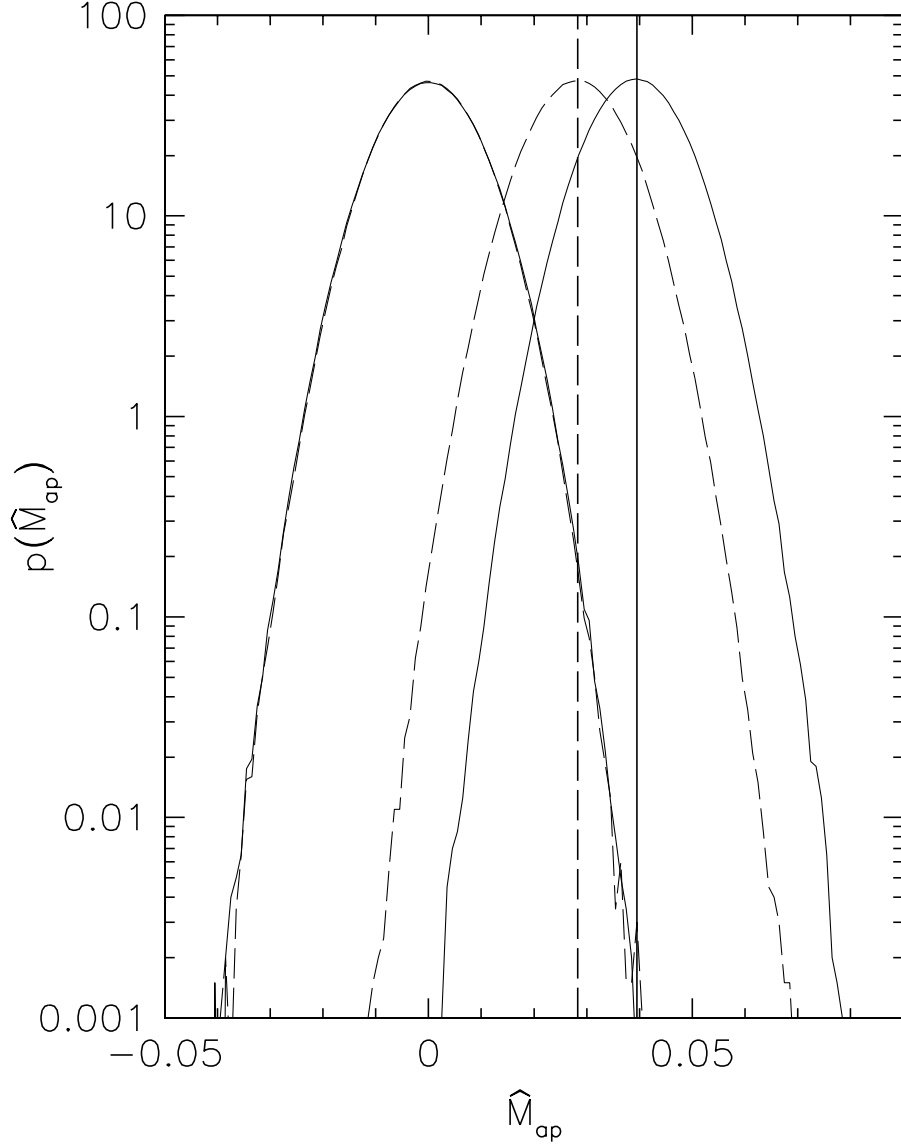


Figure 50: Probability distributions for  $\hat{M}_{\text{ap}}(160'')$  at the clump centre (Figure 8 from [17]): Solid (dashed) curves correspond to the MOCAM (UH8K-chip) data set. For an aperture of  $160''$ , the left of the two curves shows the probability distribution  $p_0(\hat{M}_{\text{ap}})$  for values of  $\hat{M}_{\text{ap}}(160'')$ <sub>rand</sub> obtained by randomising the position angles of the galaxy images. These two curves nearly coincide. The two curves on the right-hand side show the probability distribution  $p_{\text{boot}}(\hat{M}_{\text{ap}})$  obtained from bootstrap resampling of the galaxy images inside the aperture. The two vertical lines show the measured values of  $\hat{M}_{\text{ap}}$

detection about  $1'$  away from the centre of our dark lens. It is not clear whether this X-ray emission is connected with our clump or whether it originates from another source like a galaxy group. Indeed, in the light distribution from the right panel of Fig. 47, we see an overdensity at the position of the X-ray emission. If we assume that the X-ray emission is connected with our clump candidate, crude mass estimates from an X-ray flux-temperature

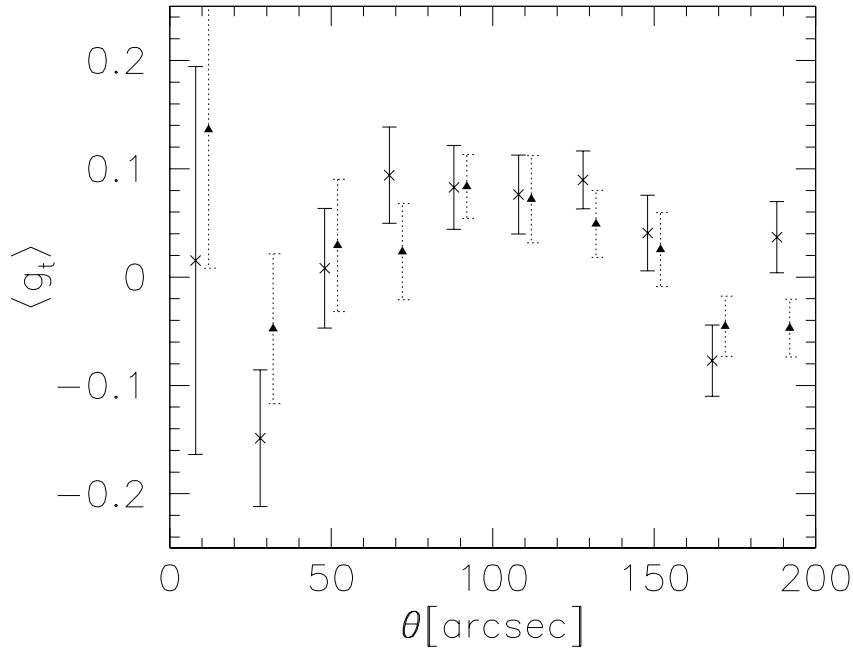


Figure 51: Mean tangential image ellipticity in independent bins of width  $20''$  around the dark clump (Figure 9 from [17]): Triangles (Crosses) show the mean, solid (dashed) error bars the 80% error interval obtained from bootstrapping, using the MOCAM (UH8K-chip) data. For better display, the points and error bars are slightly shifted in the  $\theta$  direction.

mass relation range from  $1.5 \times 10^{13} h^{-1} M_{\odot}$  if the clump is at the same redshift as A1942 ( $z = 0.223$ ) to  $1.6 \times 10^{14} h^{-1} M_{\odot}$  at  $z = 1$  (for  $q_o = 0.5$  and within a radius of  $0.5 h^{-1}$  Mpc). A beta profile fit yielded estimates from  $9.2 \times 10^{12} h^{-1} M_{\odot}$  at  $z = 0.223$  and  $2.3 \times 10^{13} h^{-1} M_{\odot}$  at  $z = 1$ . The discrepancies of these values indicate the large uncertainties involved in the mass estimate from the weak X-ray source. In any case, it seems that the X-ray and lensing mass disagree assuming that both come from the same source.

## 7.5 Conclusions

Purely based on the assumption of weak lensing studies, that background galaxies are oriented randomly in the absence of a foreground mass, we have detected a mass concentration with a very high significance level about  $7'$  South of the cluster Abell 1942. The analysis was done independently in two high-quality data sets that were obtained in two different years, with two different cameras and in two optical bands at the CFHT telescope. This makes it very unlikely that the signal originates from problems in the optics or uncorrected PSF effects.

Assuming it is real, our discovery seems to have very unusual physical properties if the lensing signal comes from a massive object at a well defined redshift. While the lensing signal, seen in a range of  $50'' - 150''$  around the clump centre, indicates a very massive object with  $M \geq 10^{14} M_{\odot}$ , the lack of observed light discourages the existence of a typical cluster. We do not see any obvious overdensity of optical light at the centre of the lens signal and we estimate  $M/L \geq 500 M_{\odot}/L_{\odot}$  for an EdS universe, where mass-to-light ratios are typically quoted. This only could be lowered to the ratio of a typical galaxy-clusters ( $200 - 300 M_{\odot}/L_{\odot}$ ) by putting the clump to substantially higher redshift. This on the other hand would result in a tremendous increase of the mass estimate (see Fig. 53). Also the lack of strong X-ray emission rules out a usual galaxy cluster. Only further observations may elucidate the nature of our discovery. After our findings, the detection of a second dark clump in HST data was reported (see [68]).

But whatever the interpretation at this point, one must bear in mind that weak lensing opens

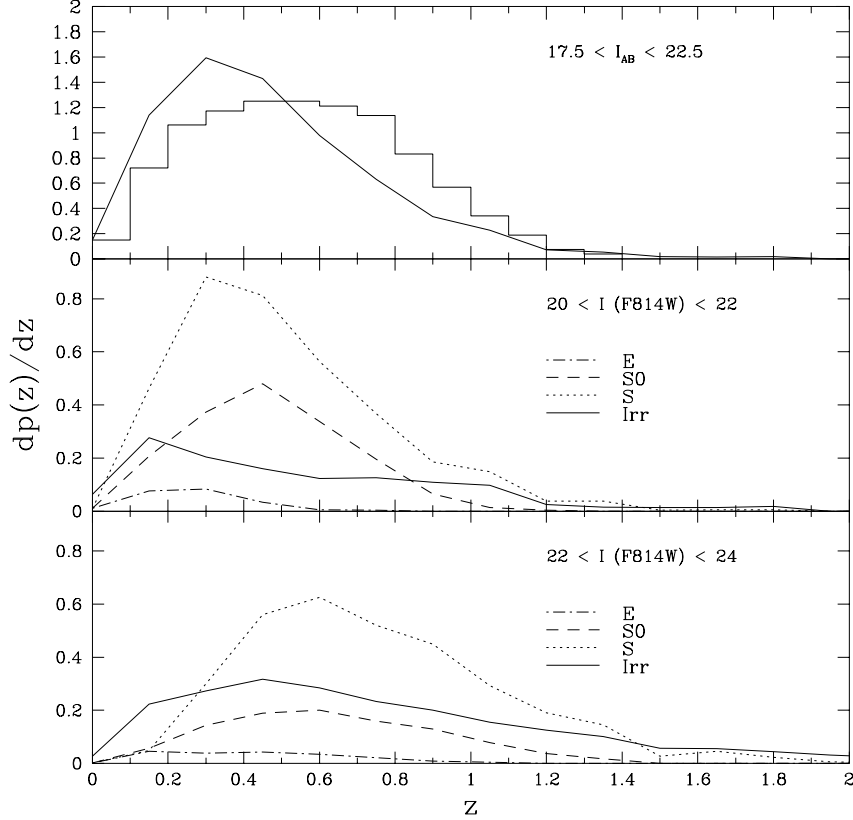


Figure 52: Redshift distributions (Figure 3 from [3]): The figure shows predictions for the redshift distribution of galaxies selected in different magnitude bins. The predictions were obtained with a semi-analytical *hierarchical clustering* approach for galaxy-formation. The uppermost panel shows a comparison of the model calculations (line) with a sample from the Canada-France-redshift survey (histogram). The middle and bottom panels show predictions for fainter magnitude limits. The bottom panel represents approximately our galaxy sample in the UH8K-chip. Assuming that we used about half the galaxies of this sample in our analysis,  $\langle z \rangle = 1$  is a good upper limit for the mean redshift of our background sources.

up a new channel for the detection of massive haloes in the Universe, so that one should perhaps not be surprised to find a new class of objects, or members of a class of objects with unusual properties. As already mentioned in section 3.2 the potential consequences of the existence of such highly underluminous objects may be far reaching: if, besides the known optical and X-ray luminous clusters, a population of far less luminous dark matter haloes exist, the normalisation of the power spectrum may need to be revised, and the estimate of the mean mass density of the Universe from its luminosity density and an average mass-to-light ratio may change.

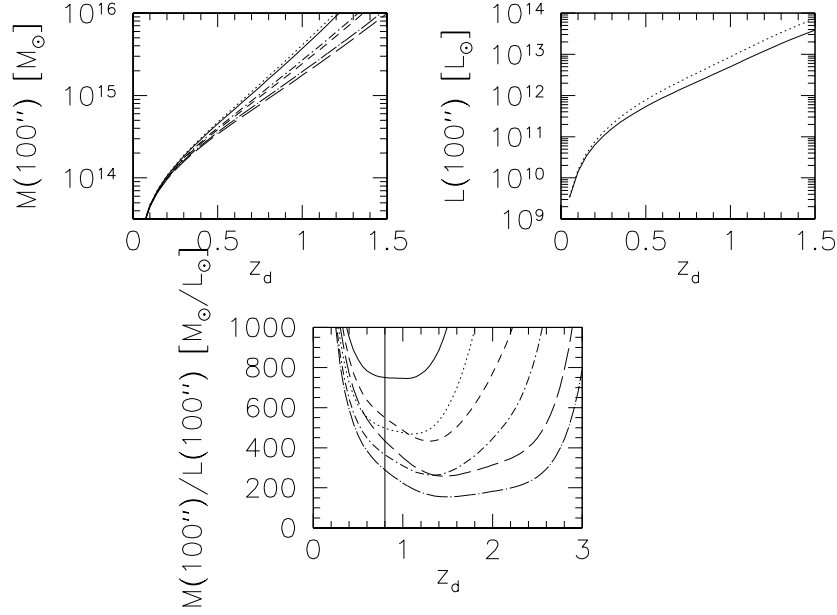


Figure 53: Mass, luminosity and  $M/L$  estimates for the clump (Figure 10 from [17]): Estimate of the lensing mass (upper left panel), an upper bound for the luminosity of the lens (upper right panel), and a lower limit on the mass-to-light ratio (lower panel), as a function of assumed lens redshift. All estimates are for an aperture size of  $100''$ . The solid, short dashed and long dashed curves show the  $M/L$  ratio in an EdS universe ( $\Omega_0 = 1, \Omega_\Lambda = 0$  for  $\langle z_s \rangle = 0.8, \langle z_s \rangle = 0.9$  and  $\langle z_s \rangle = 1.0$ ). The dotted, dot-short dashed and dot-long dashed curves show the same in an  $\Omega_0 = 0.3, \Omega_\Lambda = 0.7$  universe. We have assumed a redshift distribution  $\propto z^2 \exp[-(z/z_0)^{3/2}]$  for the source galaxies; hence  $\langle z_s \rangle \approx 1.5z_0$ . A value of  $\gamma_t(100) = 0.06$  was used, which corresponds to the measured average tangential shear on Fig. 13 between  $50''$  and  $150''$ .

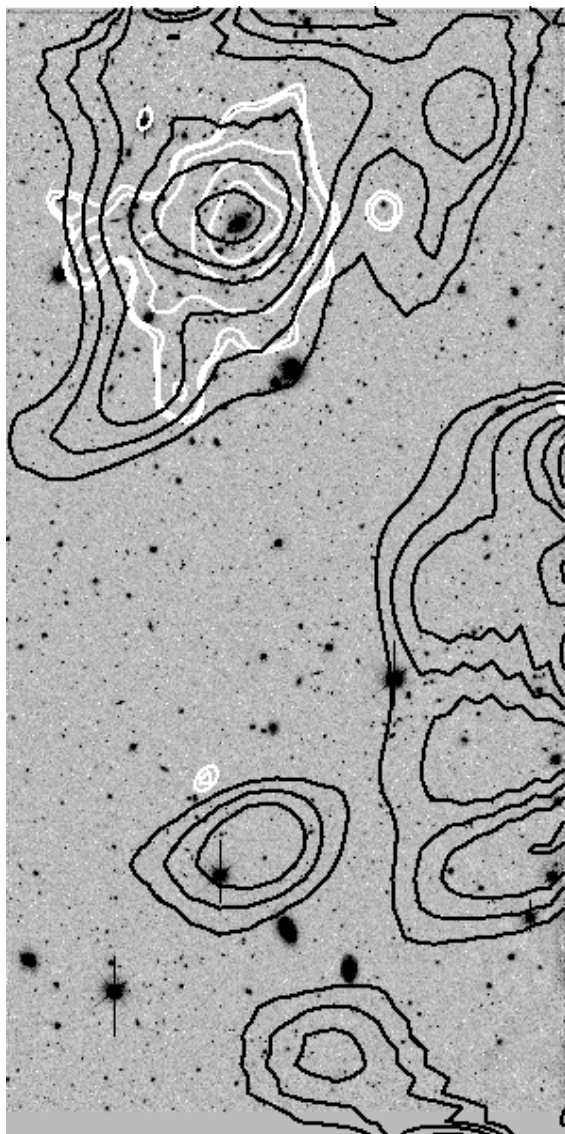


Figure 54: Lensing mass (black) and X-Ray emission (white) in the UH8K-chip (Figure 11 from [17]): The surface density contours are the same as in Fig. 47, whereas the X-ray contours correspond to  $1.5 \times 10^{-5}$ ,  $1.6 \times 10^{-5}$ ,  $2.0 \times 10^{-5}$ ,  $3.0 \times 10^{-5}$  and  $4.0 \times 10^{-5}$  counts/s/pixel. The cluster A1942 itself is clearly seen in X-rays, centred on the brightest cluster galaxy. In addition, slightly extended X-ray emission near the dark clump is detected.



## 8 First detection of cosmic shear

### Summary:

*Cosmic Shear aims at the direct measurement of the large-scale matter distribution by investigating the density contrast  $\delta$ . Theoretical work about these measurements with weak lensing have started as early as 1991 (see [8]) and we now have a very clear picture how the effect can be detected and what we can expect from these measurements. These studies predict that large-scale structures cause a shear on the percent level. Besides the large amount of high-quality data that these studies require, the challenge to measure shear at this accuracy and thereby controlling possible systematic effects that could mimic a signal, have been the reason that it took until the year 2000 before a high signal-to-noise measurement has been obtained. Our extensive work on the KSB shear measurement algorithm and its positive result make us confident that the measured signal is of cosmological origin. We start this section with theoretical predictions for the measurements we performed, and then give a detailed description of our results and our tests for possible systematics. The work was published in [71].*

### 8.1 Theoretical predictions for cosmic shear

We have shown in eq. (73) that we can define a surface mass density  $\kappa$  for the projected density field  $\delta$  in complete analogy to the case where all the matter was concentrated at a well defined comoving distance  $w$  (or at a well defined redshift  $z_d$ ). Consequently, the projected  $\delta$  (i.e.  $\kappa$ ) could directly be reconstructed. Theoretical work has been done on this approach by Van Waerbeke et al. 1999 (see [70]) and Fig. 55 shows some of their results. We could analyse these maps to infer the statistical properties of  $\kappa$  and thus the cosmological parameters of our universe.

Similarly to the previous section, there are easier ways to infer the statistical properties of  $\kappa$  if we do not need a full two-dimensional mass map of it. The basic idea here is (also in close analogy to the previous section) to measure moments of  $\gamma$  in circular apertures. Because of the linear connection between  $\gamma$  and  $\kappa$  these measurements directly reflect moments of  $\kappa$  and hence the density contrast  $\delta$ .

We first look again at eq. (76). In contrast to the single lens case where the redshift distribution of the sources was neglected up to the point where we wanted to estimate physical values for the mass we will incorporate it here immediately into the formalism. Let us assume that a redshift distribution for our galaxies  $p_z(z)dz = p_w(w)dw$  is given. Then eq. (76) becomes after averaging over the source distances

$$\kappa(\vec{\theta}) := \int_0^{w_h} dw p_w(w) \kappa(\vec{\theta}, w) = \frac{3}{2} \left( \frac{H_0}{c} \right)^2 \Omega_0 \int_0^{w_H} dw g(w) f_K(w) \frac{\delta(f_K(w)\vec{\theta}, w)}{R(w)}, \quad (181)$$

where

$$g(w) := \int_w^{w_H} dw' p_w(w') \frac{f_K(w' - w)}{f_K(w')} \quad (182)$$

is the source-averaged distance ratio  $D_{ds}/D_s$  for a density fluctuation at distance  $w$ , and  $w_H$  is the comoving distance to the horizon ( $z(w_h) = \infty$ ).

In the previous section we reported on the detection of a massive object by calculating a filtered first moment of  $\kappa$  (via the tangential shear) in a circular aperture. Doing the same with the projected surface mass density of eq. (181) we would calculate a filtered first moment of the density contrast  $\delta$  that is zero by definition (at least if the field under consideration is large enough). On the other hand considering the correlation function  $\langle \kappa(\vec{\theta})\kappa(\vec{\theta}') \rangle$  we also consider the two point correlation function of the density contrast and hence the Fourier transform of the power spectrum [see eq. (29)]. With the Fourier transform

$$\tilde{\kappa}(\vec{s}) := \int d^2\theta \kappa(\vec{\theta}) \exp(-i\vec{\theta}\vec{s}) \quad (183)$$

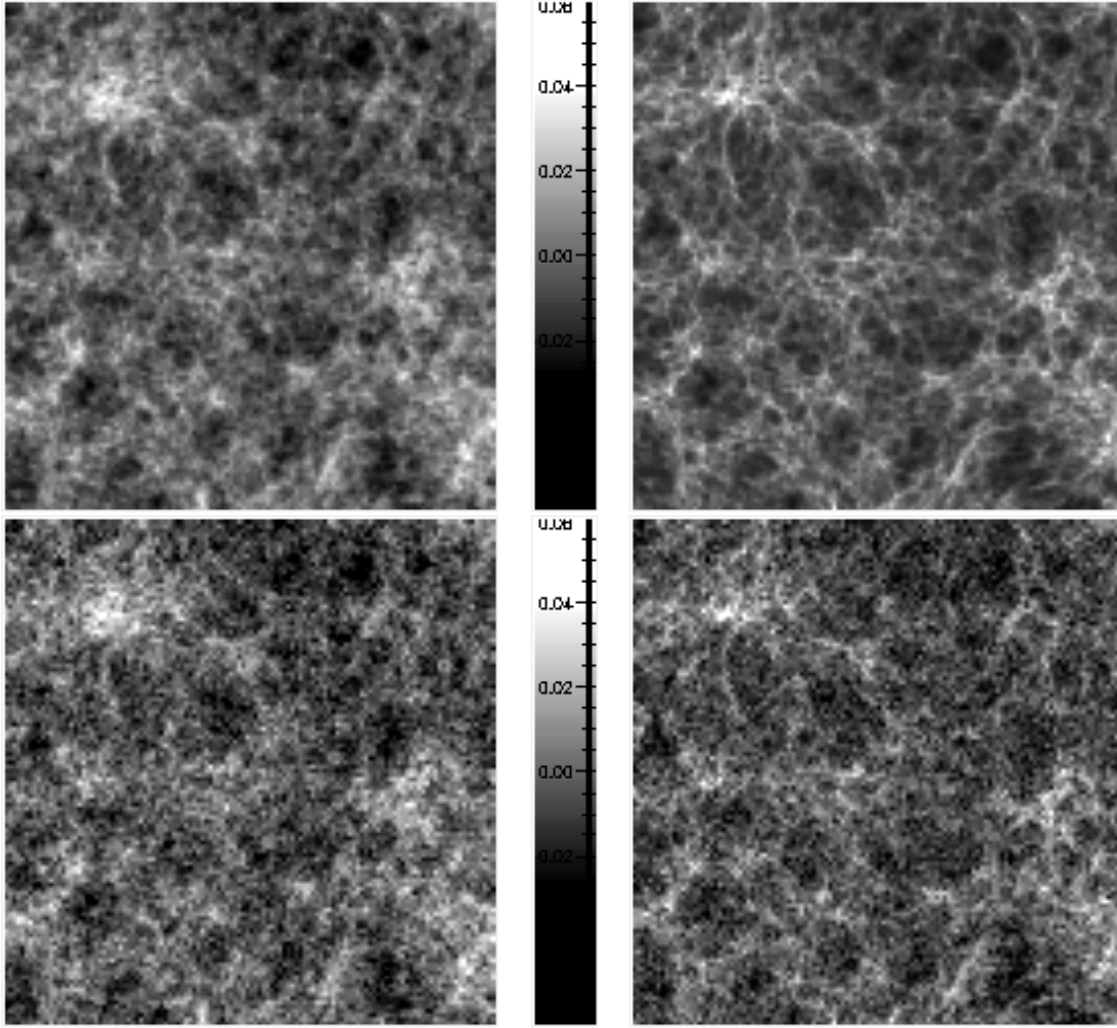


Figure 55: Reconstructed  $\kappa$  maps from large-scale structures simulations (Figure 1 from [70]): On the upper panels noise free  $\kappa$  maps from the density contrasts in two cosmologies [left panel:  $\Omega_0 = 1.0$  (flat model); right panel:  $\Omega_0 = 0.3$  (open model)] are shown. The lower maps show reconstructed maps from background galaxies at  $z_s = 1.0$  (including realistic noise). The total area of the simulations is 25 square degrees. These reconstructions would provide us with the full probability distribution of the projected density contrast  $\kappa$ .

of  $\kappa(\vec{\theta})$  we have

$$\langle \tilde{\kappa}(\vec{s}) \tilde{\kappa}^*(\vec{r}) \rangle = (2\pi)^2 \delta_D(\vec{s} - \vec{r}) P_\kappa(|\vec{s}|), \quad (184)$$

where  $P_\kappa(|\vec{s}|)$  is related to the three-dimensional power spectrum by the projection

$$P_\kappa(|\vec{s}|) = \frac{9}{4} \left( \frac{H_0}{c} \right)^4 \Omega_0^2 \int_0^{w_H} dw \frac{g^2(w)}{R^2(w)} P \left( \frac{|\vec{s}|}{f_K(w)}; w \right) \quad (185)$$

(see [32]). For relating  $\langle \tilde{\kappa}(\vec{s}) \tilde{\kappa}^*(\vec{r}) \rangle$  to the observable shear, Blandford et al. 1991 (see [8]) have shown the important relation

$$\langle \kappa(\vec{\theta}) \kappa(\vec{\vartheta}) \rangle = \langle \gamma(\vec{\theta}) \gamma^*(\vec{\vartheta}) \rangle. \quad (186)$$

Thus the two point correlation function of  $\kappa$  is the same as the two point correlation function of

the shear. We now consider the mean shear within a circular aperture with radius  $\theta_{\text{out}}$

$$\bar{\gamma}(\theta_{\text{out}}) = \frac{1}{\pi\theta_{\text{out}}^2} \int d^2\theta \gamma(\vec{\theta}). \quad (187)$$

The dispersion  $\langle \bar{\gamma} \bar{\gamma}^* \rangle =: \langle \gamma^2 \rangle$  of this quantity is given by

$$\langle \gamma^2 \rangle = \frac{1}{\pi^2 \theta_{\text{out}}^4} \int d^2\zeta \int d^2\vartheta \langle \gamma(\vec{\zeta}) \gamma^*(\vec{\vartheta}) \rangle \quad (188)$$

$$= \frac{1}{\pi^2 \theta_{\text{out}}^4} \int d^2\zeta \int d^2\vartheta \langle \kappa(\vec{\zeta}) \kappa(\vec{\vartheta}) \rangle \quad (189)$$

$$= \frac{1}{\pi^2 \theta_{\text{out}}^4} \int d^2\zeta \int d^2\vartheta \int \frac{d^2s}{(2\pi)^2} \exp(i\vec{s}(\vec{\zeta} - \vec{\vartheta})) P_\kappa(|\vec{s}|) \quad (190)$$

$$= \frac{2}{\pi \theta_{\text{out}}^4} \int_0^\infty ds s P_\kappa(s) \left( \int_0^{\theta_{\text{out}}} d\vartheta \vartheta J_0(s\vartheta) \right)^2 \quad (191)$$

$$= 2\pi \int_0^\infty ds s P_\kappa(s) [I(s\theta_{\text{out}})]^2, \quad (192)$$

where  $I(\eta)$  was defined as

$$I(\eta) = \frac{J_1(\eta)}{\pi\eta} \quad (193)$$

and  $J_0$  and  $J_1$  are the first two Bessel functions of the first kind. In the calculations above we have used eq. (186), eq. (184), the fact that the angular integration in  $\vec{\zeta}$  and  $\vec{\vartheta}$  leads to the Bessel function  $J_0$  and the relation

$$\frac{d}{d\theta}(\theta J_1(\theta)) = \theta J_0(\theta) \quad (194)$$

between the Bessel functions  $J_0$  and  $J_1$ .

- We have just shown that by measuring the variance of the shear in circular annuli we get direct information about the filtered projected power spectrum at this scale. Schneider et al. 1999 (see [63]) showed that we can use the  $M_{\text{ap}}$  statistics as a measure for cosmic shear also here. From the definition in eq. (169) it is immediately obvious that the weighted correlation function of  $\kappa$  is proportional to the weighted correlation function of the tangential shear. For the variance  $\langle M_{\text{ap}}^2(\theta_{\text{out}}) \rangle$  of  $M_{\text{ap}}$  we get with similar calculations as above

$$\langle M_{\text{ap}}^2(\theta_{\text{out}}) \rangle = 2\pi \int_0^\infty ds s P_\kappa(s) [I'_l(s\theta_{\text{out}})]^2, \quad (195)$$

where the weight function  $I'_l(s\theta_{\text{out}})$  reads

$$I'_l(s\theta_{\text{out}}) = \int_0^1 dx x u_l(x) J_0(x\eta) = \frac{2^l \Gamma(l+3)}{\pi} \eta^{-(l+1)} J_{3+l}(\eta). \quad (196)$$

$u_l$  was defined in eq. (172) and  $\Gamma(x)$  is the Gamma function. Figure 56 shows a comparison between the filters  $I(\eta)$  and  $I'_l(\eta)$ . Besides the filter function, the two measures also differ in their noise properties. It turned out that a significant measure of  $M_{\text{ap}}$  requires more data than a measurement of the shear variance. For the rest of this thesis we will only deal with the shear variance that we finally have measured. The term “measurement of cosmic shear” is most often understood with “the measurement of the shear variance on some scale”. In Fig. 57 we show theoretical calculations for this measure in different cosmological models. We clearly see that we have to make accurate measurements of the shear variance to distinguish different cosmologies.

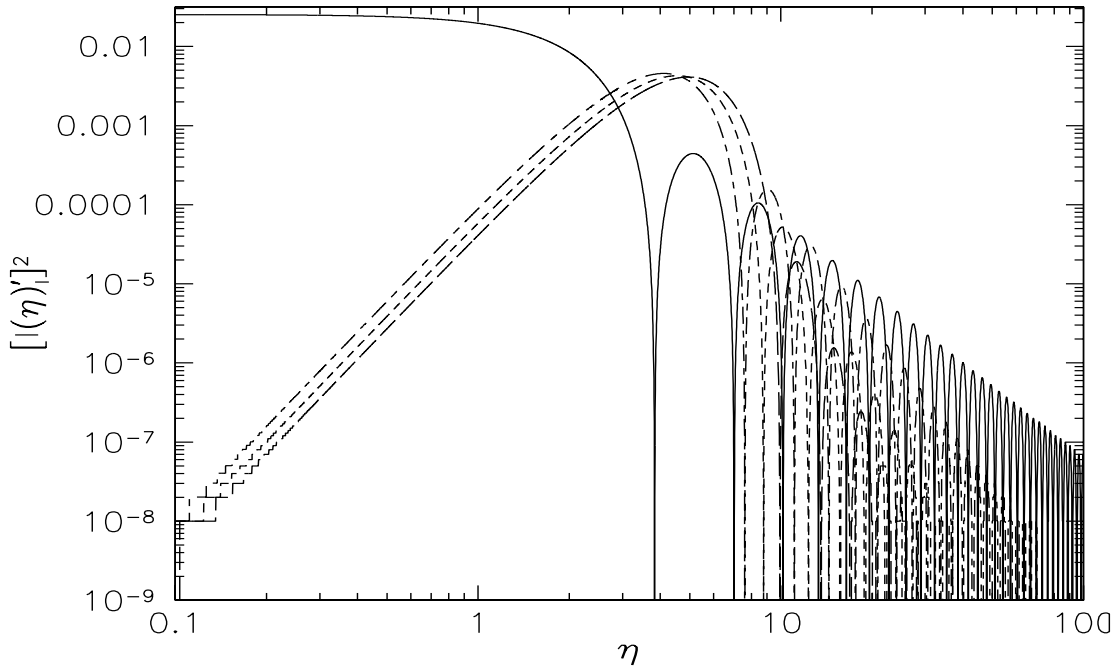


Figure 56: Filters for the shear variance calculation (Figure 2 from [63]): The plot shows the filter functions  $[I_l(\eta)]^2$  for  $l = 1$  (long-dashed short-dashed curve),  $l = 2$  (short dashed curve) and  $l = 3$  (long dashed curve) from the  $M_{\text{ap}}$  statistics. The filter  $[I(\eta)]^2$  (solid curve) from the shear variance is also given. We note that the  $M_{\text{ap}}$  filters are quite peaked. Consequently we can hope to accurately measure the projected power spectrum on different scales. In contrast, the shear variance only gives us the mean of the power spectrum on large scales.

- The shear variance can easily be estimated by the observable image ellipticities. We consider the quantity

$$G(\theta_{\text{out}}) := \sum_{\alpha=1}^2 \left( \frac{1}{N} \sum_{k=1}^N g_{\alpha}(\vec{\theta}_k) \right)^2, \quad (197)$$

where the inner summation is performed over the  $N$  galaxies at positions  $\vec{\theta}_k$  within our aperture, and the outer summation over the ellipticity components. The expectation value  $\langle G \rangle$  is

$$\langle G \rangle = \frac{\sigma_{\epsilon}^2}{N} + \langle \gamma^2 \rangle. \quad (198)$$

Hence, an unbiased estimator for  $\langle \gamma^2 \rangle$  is given by  $G - \frac{\sigma_{\epsilon}^2}{N}$ . The term  $\frac{\sigma_{\epsilon}^2}{N}$  can easily be estimated by using ellipticities with randomised position angles. We note that eq. (197) directly estimates  $\langle \gamma^2 \rangle$  if we leave out the diagonal terms.

- The noise in the calculation of  $\langle \gamma^2 \rangle$  consists of two parts: On the one hand, the intrinsic ellipticity distribution as in all weak lensing studies, on the other hand we have the so-called *cosmic variance* in addition. This is the variance we get when we measure  $\langle \gamma^2 \rangle$  in different directions of the sky. While the errors coming from the intrinsic ellipticity distribution are getting small very quickly, simulations showed that we need several square degrees of data to beat the cosmic variance and to get significant constraints on cosmological models.

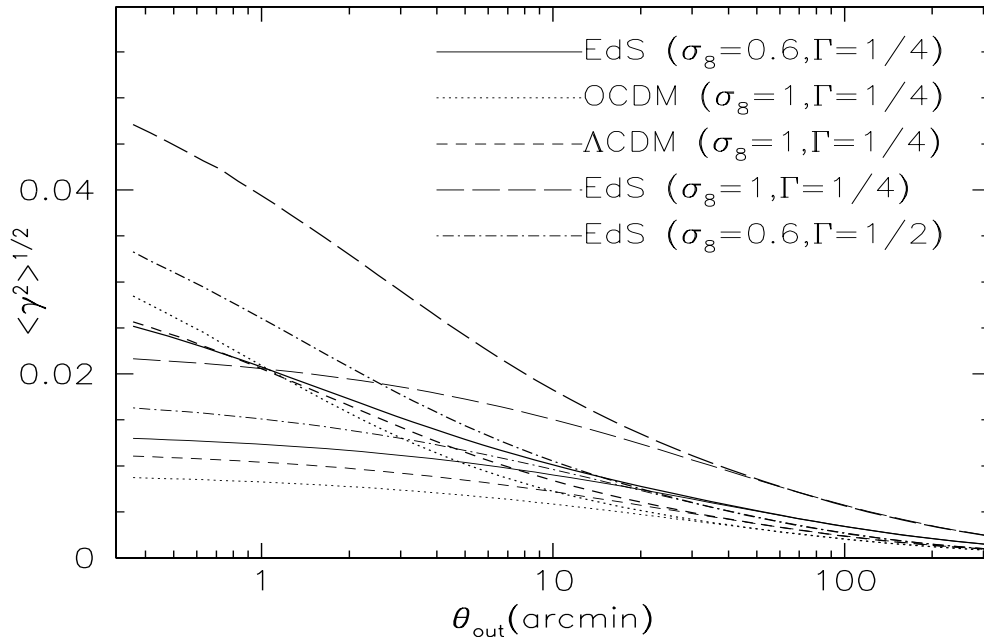


Figure 57: Theoretical predictions for the shear variance: Shown are results from eq. (188) for the cosmological models of Table 1. All sources are assumed to lie at a redshift of unity (the assumed mean redshift of our data). Thin curves correspond to the linear, thick ones to the nonlinear evolution of the power spectrum. We note that the non-linear evolution raises the prediction on small angular scales by about a factor of 2, but the expected shear requires accurate measurements on the percent level, especially if we want to distinguish between cosmological models.

## 8.2 The data

As mentioned above, the accumulation of high-quality data is one of the major obstacles for the detection of cosmic shear. Thus we mixed data taken from 1996 to 1999 at the CFHT telescope. The observations were done with the UH8K camera (see last section) and the new CFH12K camera that was installed in 1999. It is a 12Kx8K camera covering  $28' \times 42'$  of the sky with a single exposure (0.206'' pixel scale). A summary of all the image sets is given in Table 5. Because the observations were initially done for different scientific purposes they are in two different passbands  $V$  and  $I$  and they differ in exposure time. This makes the whole set somewhat heterogeneous but this was unimportant for the primary goal: the detection of cosmic shear.

The set consists of eight different pointings spread over five statistically independent areas, each separated by more than 10 degrees. The total field covers about 1.7 square degrees (the total coverage should be 2.05 square degree but some of the chips of the UH8K and CFH12K cameras have a poor image quality and were excluded from the analysis). All the data were taken in a very compact dither pattern so that we finally worked only with individual chips of  $7' \times 14'$ . For every chip an object catalog was generated similar to the procedure described in Appendix B differing in minor details that are not important for the following. We have masked regions around very bright stars and other image defects and simply rejected all the objects falling within these masked areas. Finally, we were left with 191000 galaxies (26 galaxies per square arcmin) for our analysis.

## 8.3 Analysis and measured signal

We have measured  $\langle \gamma^2 \rangle$  on scales from  $\theta_{\text{out}} = 0.66'$  to  $\theta_{\text{out}} = 3.33'$ . The upper limit is given by the dimension of single chips (we did not paste our single chip catalogs to get data on the full

Target	Name	Camera	Used area	Filter	Exp. time	Period	seeing
F14P1	F1	CFH12K	764 arcmin <sup>2</sup>	V	5400 sec.	May 1999	0.9''
F14P2	F2	CFH12K	764 arcmin <sup>2</sup>	V	5400 sec.	May 1999	0.9''
F14P3	F3	CFH12K	764 arcmin <sup>2</sup>	V	5400 sec.	May 1999	0.9''
CFDF-03	F4	UH8K	669 arcmin <sup>2</sup>	I	17000 sec.	Dec. 1996	0.75''
SA57	F5	UH8K	669 arcmin <sup>2</sup>	I	12000 sec.	May 1998	0.75''
A1942	F6	UH8K	573 arcmin <sup>2</sup>	I	10800 sec.	May 1998	0.75''
F02P1	F7	CFH12K	1050 arcmin <sup>2</sup>	I	9360 sec.	Nov. 1999	0.8''
F02P4	F8	CFH12K	1050 arcmin <sup>2</sup>	I	7200 sec.	Nov. 1999	0.9''

Table 5: List of the fields for our cosmic shear detection: Most of the exposures were taken in the *I* band at CFHT. The total area is 1.7 square degrees, and the 8 fields are uncorrelated.

dimension of the image mosaics). On each chip we measured  $\langle \gamma^2 \rangle$  on a grid of  $10 \times 20$  pixels so that the distance between two grid-points is 200 pixels (41.2'). A grid-point was rejected if it crossed the boundary of the chip or if more than 10% of its area was in a masked region. As we have used a weighting scheme for our galaxies, we modified the estimator  $G(\theta)$  to

$$G(\theta) = \sum_{\alpha=1,2} \left( \frac{\sum_{k=1}^N U_k g_{\alpha}(\vec{\theta}_k)}{\sum_{k=1}^N U_k} \right)^2. \quad (199)$$

Figure 58 shows the measured signal. The error bars and the diagonal term of eq. (199) were obtained by 1000 randomisations of the galaxy ellipticity orientation. We checked that we get the same result for the signal if we leave out the diagonal terms in eq. (199) and do not subtract the estimate for it. We see that the measured signal is consistent in shape and amplitude with theoretical predictions. In order to have a better idea of the significance of the signal, we can compare for each scale the histogram of the shear variance in the randomised samples and the original measurement. This is shown on Fig. 59, for all the smoothing scales shown in Fig. 58. The significance is up to  $5.5\sigma$  for individual scales. Note that the measurement points at different scales are correlated, and that an estimate of the overall significance of our signal would require the computation of the noise correlation matrix between the various scales.

## 8.4 Analysis of possible systematics

After the measurement of our signal we have to do a careful analysis about possible systematics. It was the first time that a significant measurement of  $\langle \gamma^2 \rangle$  has been done and we have to be convinced that it is of cosmological origin. The following tests were done:

- We rejected galaxies that have a next neighbour within  $d$  pixels to see whether overlapping isophotes produce a spurious signal. The result is shown in Fig. 60. We have applied cuts of  $d = 0$  (no rejection),  $d = 5, 10, 20$  pixels. We see that very close pairs are producing a spurious signal and the  $d = 0$  and 5 cases show an excess of power in particular on the small scales. The signal is stable for  $d = 10$  and does not change when going to  $d = 20$ . So we used the  $d = 10$  rejection finally.
- The next tests concern residuals from the PSF anisotropy correction. The star ellipticity reached up to 20% in our data fields. Maps of the uncorrected and the corrected star ellipticities are shown in Fig. 61 to Fig. 68. These figures demonstrate the variability of the anisotropy structure and the necessity to correct every observation individually as discussed in section 6. Also the discontinuities between the structures of neighbouring chips are seen. Figure 69 shows that the stellar ellipticities are efficiently corrected. We investigate now whether galaxies nevertheless show a preferred elongation in the direction of the PSF

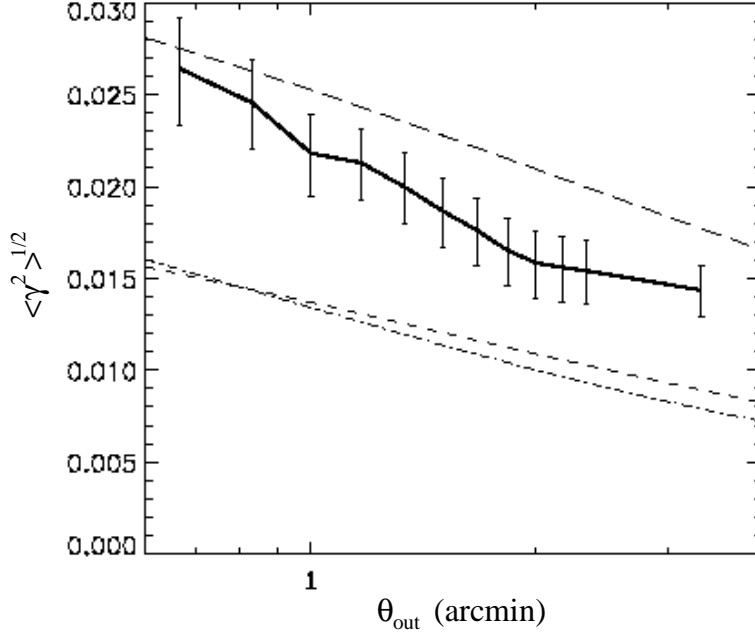


Figure 58: Measured shear variance signal (Figure 1 from [71]): Shown is the square-root of the measured shear variance as a function of the radius of the top-hat window (solid line). The maximum angular scale, 3.5 arc-minutes radius, is fixed by the maximum angular scale defined by individual CCDs (7'). Error bars are computed over 1000 random realisations of the galaxy catalogue. The other lines are theoretical predictions of the same quantity for different cosmological models in the non-linear regime: the long-dashed line corresponds to  $(\Omega_0 = 1, \Omega_\Lambda = 0, \sigma_8 = 0.6)$ , the dashed line to  $(\Omega_0 = 0.3, \Omega_\Lambda = 0, \sigma_8 = 0.6)$ , and the dot-dashed line to  $(\Omega_0 = 0.3, \Omega_\Lambda = 0.7, \sigma_8 = 0.6)$ .

anisotropy. For this we have binned our galaxy catalog according to stellar ellipticity. For each bin we then measure ellipticity averages one time with the full PSF correction applied  $\langle g_\alpha \rangle$  and a second time without the PSF correction  $\langle g_\alpha^{\text{ani}} \rangle$  (without applying  $P^{\text{sm}} \mathbf{q}^*$ ). We have pasted galaxies from all over the survey, so any local alignment of galaxies originating from possible mass concentrations should vanish and a possible non vanishing value for  $\langle g_1 \rangle$  or  $\langle g_2 \rangle$  could be attributed to uncorrected PSF effects. Figure 70 shows  $\langle g_1 \rangle$  and  $\langle g_2 \rangle$  (dashed lines) and  $\langle g_1^{\text{ani}} \rangle$  and  $\langle g_2^{\text{ani}} \rangle$  (solid lines) versus, respectively,  $\langle q_1^* \rangle$  and  $\langle q_2^* \rangle$ . The solid lines exhibit a direct correlation between the galaxy and the star ellipticities, showing that the PSF anisotropy does indeed induce a strong spurious anisotropy in the galaxy shapes of a few percents. However, the dashed lines show that the corrected galaxy ellipticities are no longer correlated with the star ellipticity, the average  $\langle g_1 \rangle$  fluctuates around  $-1\%$ , while  $\langle g_2 \rangle$  is consistent with zero. This figure confirms the remarkable accuracy of the PSF correction method that we found in our simulations (see section 6). Error bars in these plots are calculated assuming Gaussian errors for the galaxies in a given bin. The significant offset of  $\langle g_1 \rangle$  of  $1\%$  might be interpreted as a systematic induced by the CCD, as we will see in the next paragraph, and can be easily corrected for. Figure 71 shows that this systematic is nearly galaxy-size independent, and affects all galaxies in the same way. This also argues for the CCD-induced systematic, since we expect that a PSF-induced systematic (which is a convolution) would depend on the galaxy size (see section 6).

- Using the same method as in the previous paragraph, we can also investigate whether there is a systematic alignment of galaxies with CCD lines/columns. Here, instead of sorting the

Simulation #	$\Gamma$	$\Omega_0$	$\Omega_\Lambda$	$\sigma_8$
(1) $\Lambda$ CDM	0.21	0.3	0	0.85
(2) $\tau$ CDM	0.21	1	0	0.6
(3) $\tau$ CDM	0.21	1	0	1

Table 6: List of the ray tracing simulations we used (see [29] for details). The redshift of the sources is 1.

galaxies according to the star ellipticity, they are sorted according to their  $X$  or  $Y$  location on each CCD frame. By averaging the galaxy ellipticities  $\langle g_1 \rangle$  and  $\langle g_2 \rangle$  in either  $X$  or  $Y$  bins, we suppress the cosmic shear signal and keep only the possible systematics associated with the CCD frame. Figure 72 shows  $\langle g_1 \rangle$  and  $\langle g_2 \rangle$  versus  $X$  (dashed lines) and versus  $Y$  (solid lines). The plots from the top-left to bottom-right correspond respectively to  $\langle g_1 \rangle$  versus  $X$ ,  $\langle g_2 \rangle$  versus  $X$ ,  $\langle g_1 \rangle$  versus  $Y$ , and  $\langle g_2 \rangle$  versus  $Y$ . We see that  $\langle g_1 \rangle$  is systematically negative by  $\sim -1\%$  for both  $X$  and  $Y$  binnings, while  $\langle g_2 \rangle$  does not show any significant deviation from zero. This result is fully consistent with the dashed lines in Fig. 70 which demonstrate that the  $-1\%$  systematic is probably a constant systematic which affects all the galaxies in the same way, and which is not related to the PSF anisotropy correction. The origin of this constant shift is still not clear, it might have been produced during the readout process, since a negative  $\langle g_1 \rangle$  corresponds to an anisotropy along columns of the CCDs. We corrected for the effect by adding a  $1\%$   $g_1$  shear to all the galaxies.

After we have performed all these tests, we are confident that our signal is of cosmological origin. In the final section we interpret our signal with respect to cosmological constraints.

## 8.5 Cosmological constraints

As described above, the error in the measurement of  $\langle \gamma^2 \rangle$  has two parts. First, the intrinsic ellipticity distribution. The error of this on the final measurement is already fairly small with the current data so that a significant detection (up to about  $5\sigma$ ) of the signal could be obtained (see Fig. 59). To be able to reject cosmological models we have to estimate the second source of uncertainty, the cosmic variance, for our survey. This cannot be done analytically but has to be estimated from cosmological ray tracing simulations. We have used data from Jain (see [29]) for this purpose. Table 6 lists three cosmological models for which the cosmic variance was modelled. We note that the simulations were done with a constant redshift for the source galaxies at  $z_s = 1$ . Although this is not appropriate for our data we believe that it is a very good estimate for the mean redshift of our galaxies. Figure 73 shows the scale dependence for the variances of the three models compared to our signal. We can marginally reject models (1) and (3) and our measurement is in very good agreement with the cluster normalised model (2). So, our analysis is consistent with the currently favoured cosmological models, although we cannot yet reject other scenarios with high significance. Also, because of the unknown redshift distribution of our sources, the interpretation might still be subject to modifications.

We are very confident that we will put strong constraints on the cosmology when our survey will be completed. It will consist of 16 square degrees in four broad-band colours.



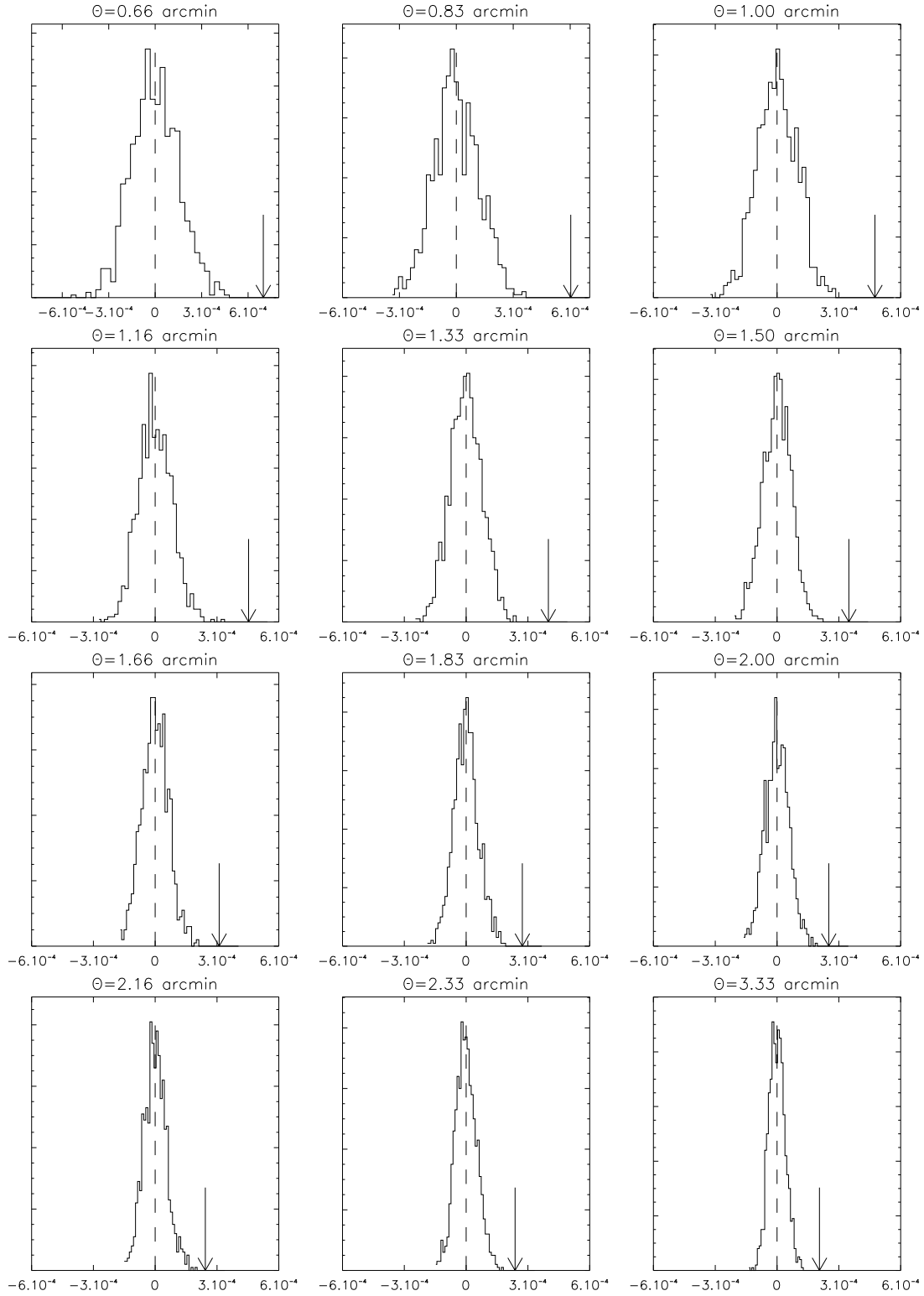


Figure 59: Significance of the shear variance (Figure 2 from [71]):  $\theta = \theta_{\text{out}}$  in this figure. For all different scales for which the shear variance was measured, the value of the signal (given by the arrow) compared to the signal measured in the randomised catalogues (histograms) is shown. This Figure illustrates how significant the signal is with respect to a pure random orientation of the galaxies.

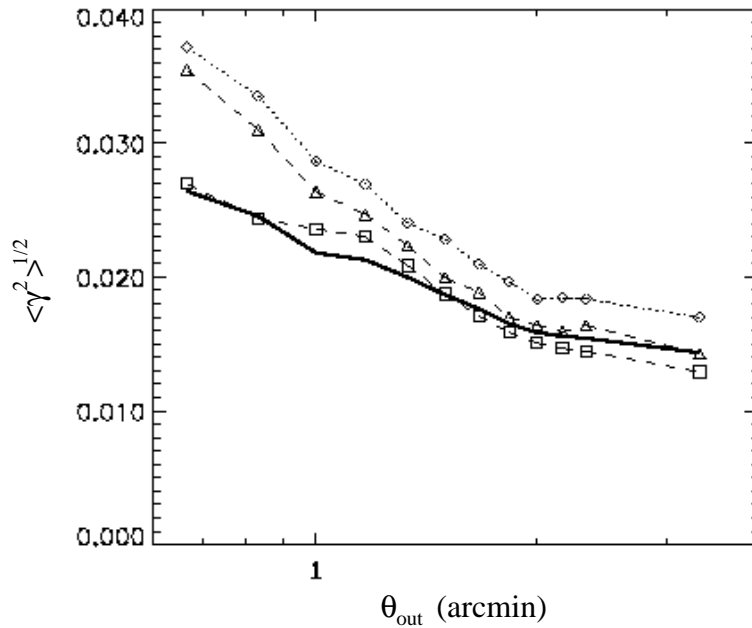


Figure 60: Bias due to overlapping object isophotes (Figure 3 from [71]): The thick solid line shows the signal as plotted on Fig. 58. It was obtained with a catalog of galaxies where galaxies having a neighbour closer than 10 pixels were rejected. The three other curves show the same signal measured with different rejection criteria: the diamond-dotted line is for no rejection at all, the triangle-dashed line for galaxies closer than 5 pixels rejected and the square-dashed line for galaxies closer than 20 pixels rejected. This figure illustrates that the overlapping isophotes of close galaxies tends to overestimate the shear variance.

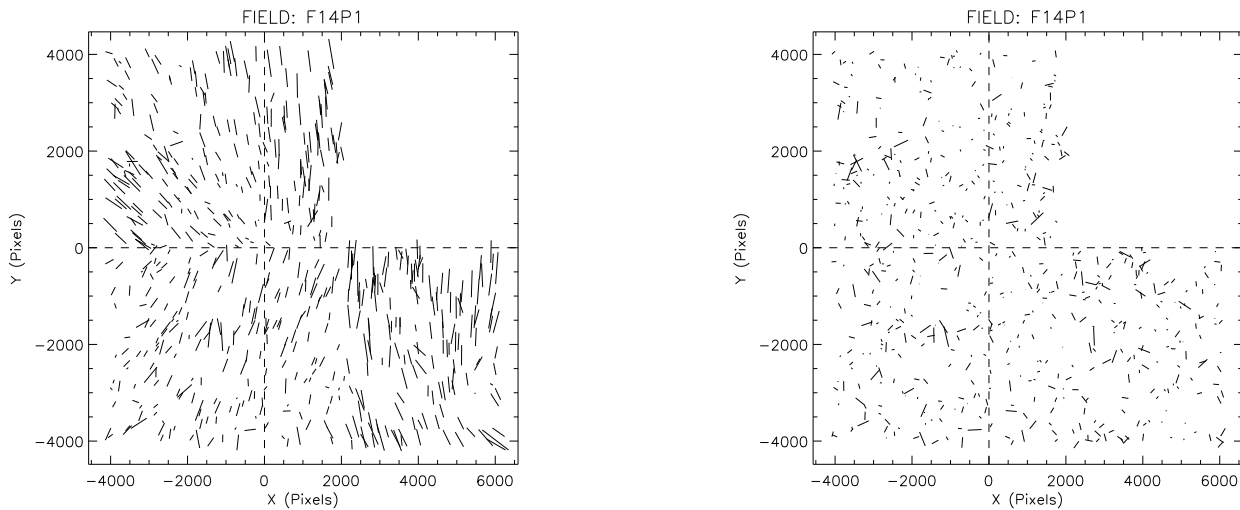


Figure 61: Uncorrected (left) and corrected (right) star ellipticities for FIELD F14P1 (Figure 12 from [71]): The dashed cross shows the location of the optical centre. Frames are graduated in pixels.

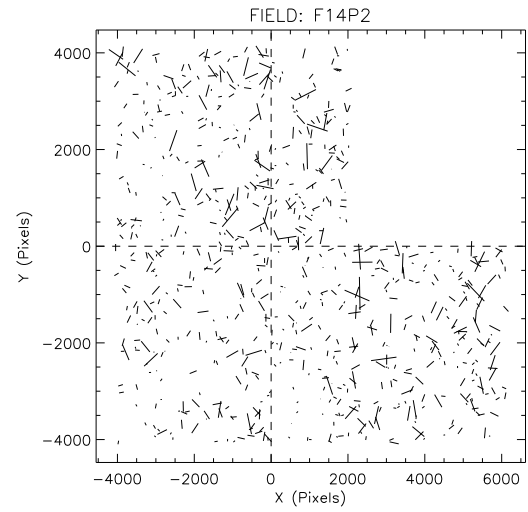
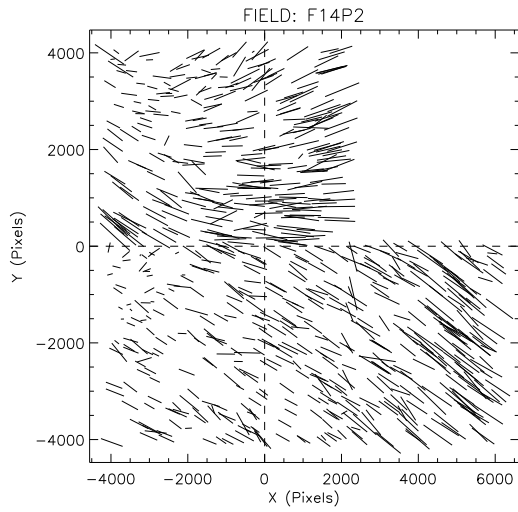


Figure 62: Same as Fig. 61 for FIELD F14P2 (Figure 13 from [71]).

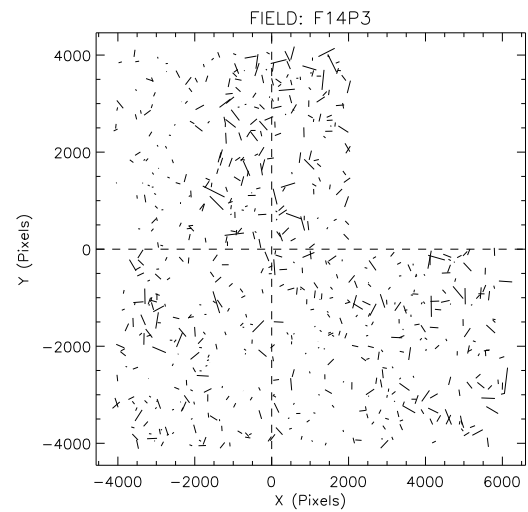
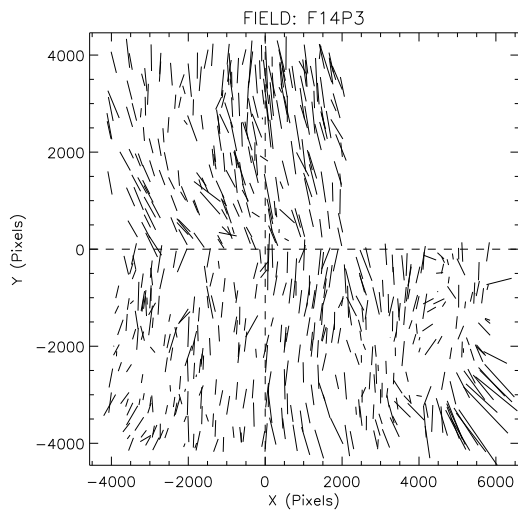


Figure 63: Same as Fig. 61 for FIELD F14P3 (Figure 14 from [71]).

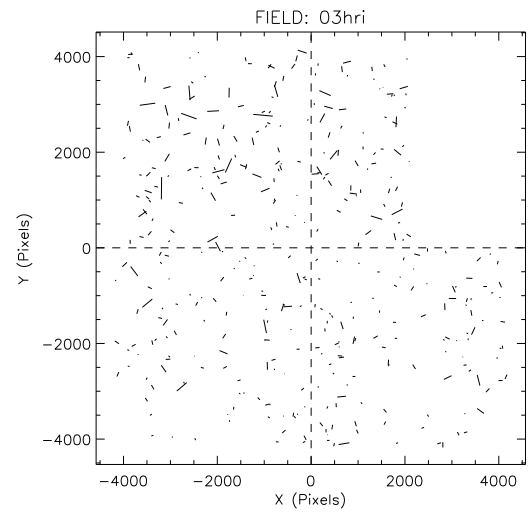
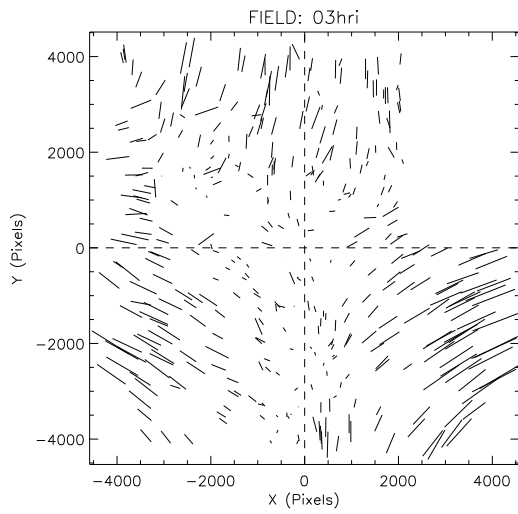


Figure 64: Same as Fig. 61 for FIELD CFDF-03 (Figure 15 from [71]).

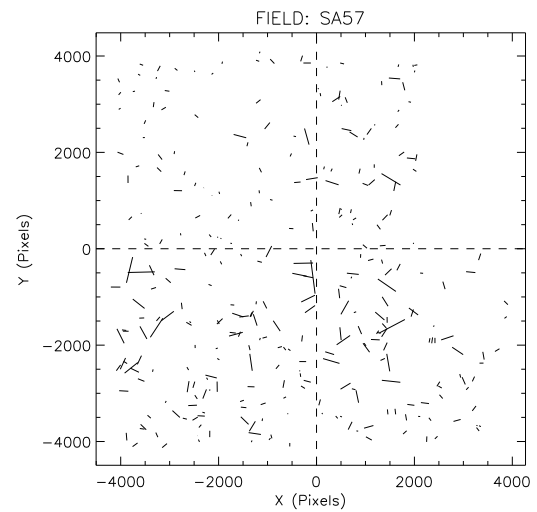
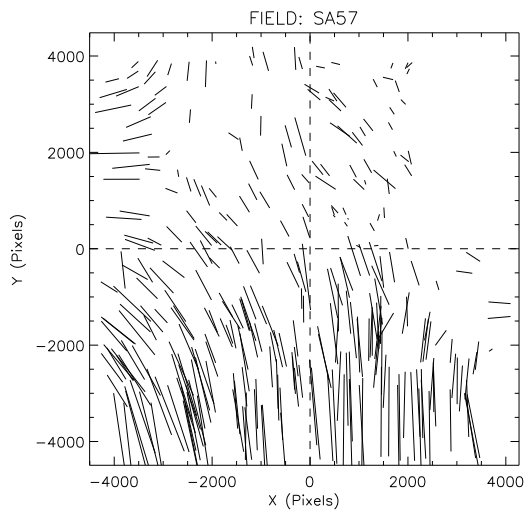


Figure 65: Same as Fig. 61 for FIELD SA57 (Figure 16 from [71]).

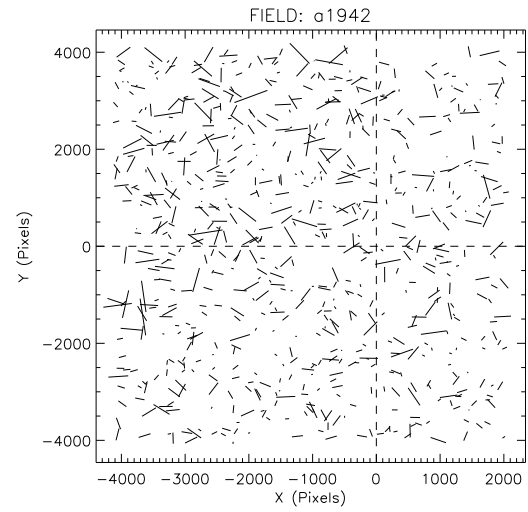
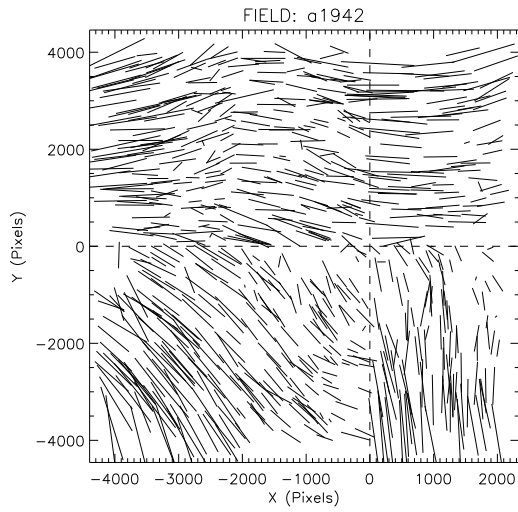


Figure 66: Same as Fig. 61 for FIELD a1942 (Figure 17 from [71]).

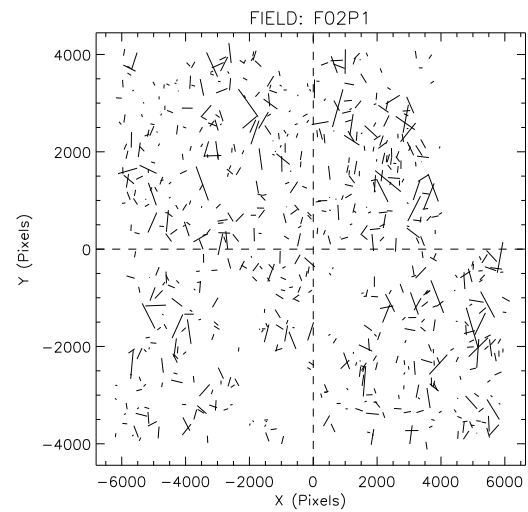
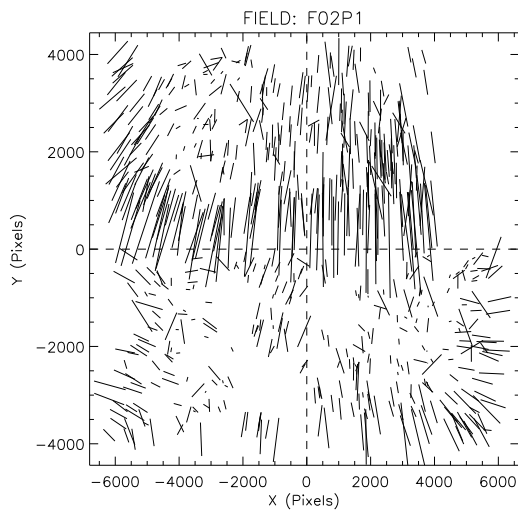


Figure 67: Same as Fig. 61 for FIELD F02P1 (Figure 18 from [71]).

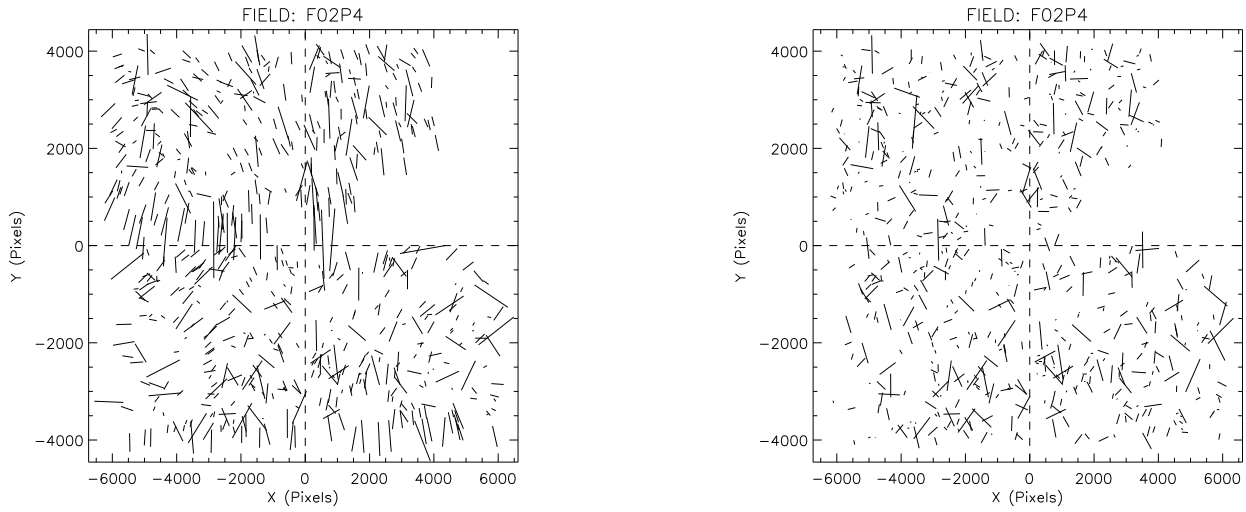


Figure 68: Same as Fig. 61 for FIELD F02P4 (Figure 19 from [71]).

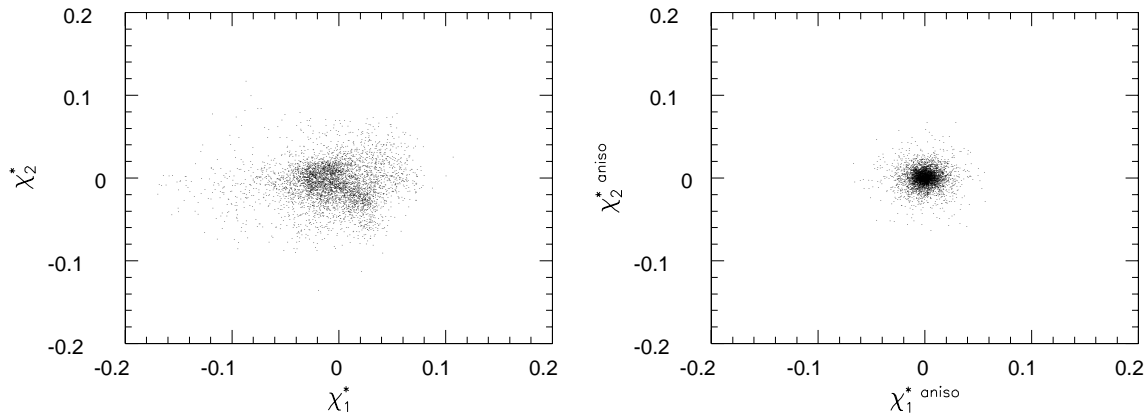


Figure 69: Star ellipticities of all the survey before (left panel) and after (right panel) the correction (Figure 11 from [71]): After correction, the star ellipticity is randomly distributed around zero, as expected.

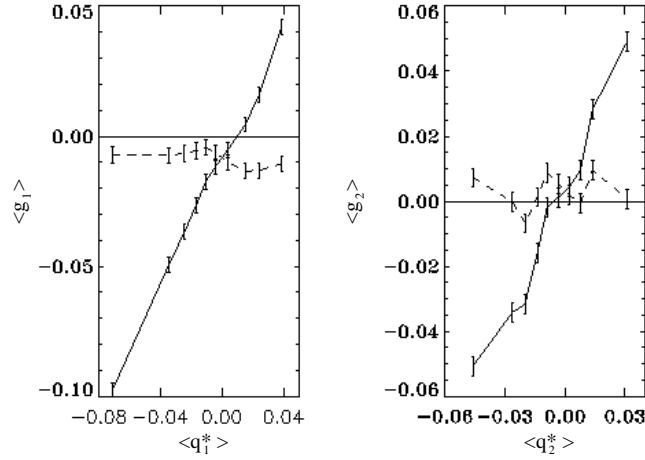


Figure 70: Accuracy of the PSF anisotropy correction (Figure 4 from [71]): Average galaxy ellipticity  $\langle g_\alpha \rangle$  versus the average star ellipticity  $\langle q_\alpha^* \rangle$  (obtained from a polynomial fit) for both components  $\alpha = 1, 2$ . The dashed lines are obtained from the fully corrected galaxy ellipticities, as given by eq. (147). The solid lines are obtained from the galaxy ellipticities corrected for the seeing, but without the anisotropy correction term  $P^{*sm}\mathbf{q}^*$ . Each ellipticity bin contains about  $N = 16000$  galaxies, and the error bars are calculated assuming Gaussian errors. Except for a constant small bias along the  $g_1$  direction, the corrected galaxies are uncorrelated with the stellar ellipticity, which demonstrates that the PSF correction method works well.

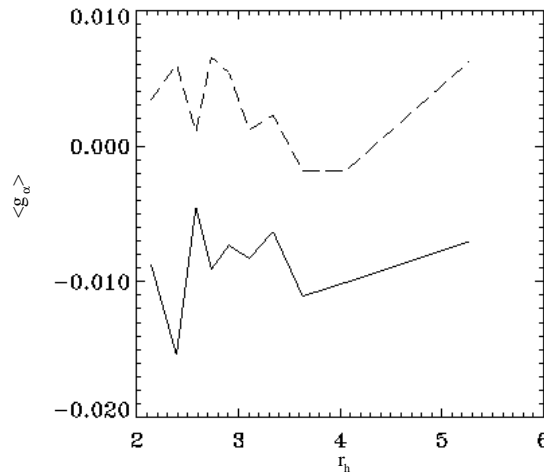


Figure 71: Size dependence of the systematic  $-1\%$  bias in  $g_1$  (Figure 5 from [71]): Average galaxy ellipticity  $\langle g_1 \rangle$  (solid line) and  $\langle g_2 \rangle$  (dashed line) as a function of the half light radius  $r_h$ . It is shown that the systematic bias of  $-1\%$  along the  $g_1$  component is fairly size independent.

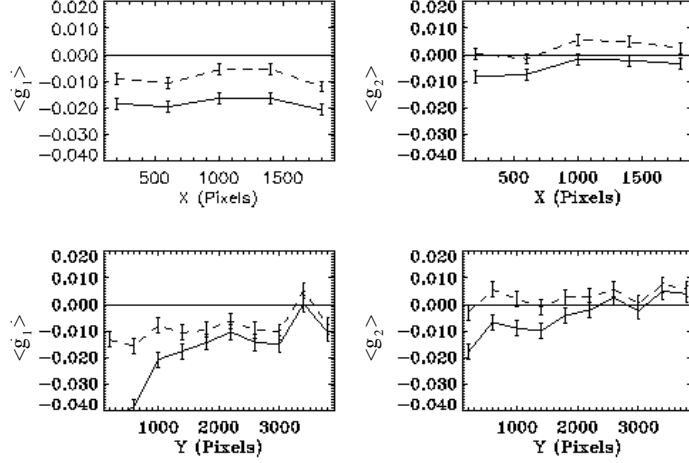


Figure 72: Average galaxy ellipticity  $\langle g_\alpha \rangle$  versus the  $X$  and  $Y$  location on the CCDs (Figure 7 from [71]): As for Fig. 70, the dashed lines are obtained from the fully corrected galaxy ellipticities, as given by eq. (147), and the solid lines are obtained from the corrected galaxy ellipticities where the anisotropy correction term  $P^{*sm} \mathbf{q}^*$  has not been applied. The systematic negative mean value of  $\langle g_1 \rangle$  along lines or columns of the CCD (the two left panels) show that the galaxies are preferentially aligned with the columns of the CCD in the whole survey. A positive systematic value for  $\langle g_2 \rangle$  (the two right panels) is also visible, although much less significant.

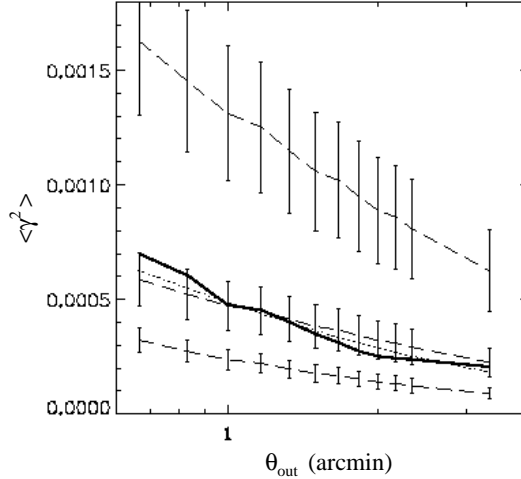


Figure 73: Cosmological constraints from our measured signal (Figure 10 from [71]): Comparison of our signal (thick line) with three cosmological models. The error bars are the cosmic variance measured on five independent realisations of ray-tracing simulations for the cosmologies listed in Table 6. For clarity, the error bars of the signal are not plotted, their amplitude can be read on Fig. 58. From bottom to top, the dashed lines correspond to: model (1), model (2) and model (3) as given in Table 6. The shot-noise error bars of the signal are in fact comparable in amplitude to the cosmic variance error bars of model (2). We show also a cluster-normalised  $\Lambda$  model (dotted line) with  $\Omega_0 = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and a CDM power spectrum with  $\Gamma = 0.5$ . This model was not obtained from a simulation, but computed using the non-linear power spectrum using the fitting formula from Peacock & Dodds (see [51]).



## 9 Outlook

In this thesis, I have presented work on three applications of the weak lensing effect. The first part was purely theoretical; it was investigated how we can optimally constrain a parametrised mass model of a known galaxy cluster. Although its usefulness for quantitative analysis is still in question as we do not know whether we can determine galaxy number counts on small scales with sufficient accuracy, it has already been applied to data by an English collaboration. Some aspects of this work that are worth of further investigation are:

- To also study galaxy clusters at intermediate redshift ( $z \approx 0.3$ ), the redshift distribution of the background galaxy population has to be accounted for. The errors connected with this uncertainty, and to what degree *photometric redshifts* can improve the situation, should be investigated.
- To extend comparisons of cluster mass models. As galaxy clusters are not circularly symmetric objects, generating simulated data with elliptical profiles but analysing the data with a circularly symmetric one is of special interest. Also the inclusion of the more realistic Navarro, Frenk & White mass profile should be done.

The rest of the thesis was dedicated to the exciting new possibilities that wide-field imaging offers. We now have access to several highly sensitive CCD mosaic cameras that cover  $30' \times 30'$  with a single exposure. The associated data flow requires that we can process the data automatically with confidence in the resulting object catalogs. A first step to build up a fully automatic pipeline for weak lensing observations has been done in section 6. We showed with realistic simulations that we can produce catalogs for reliable shear measurements out of raw images without user interaction. Nevertheless, there remains a lot of technical work to be done in this area. Some of the aspects related to lensing are:

- Especially in wide-field imaging, many images suffer from very bright stars causing blooming effects, strong stray light and reflections (see also Fig.22). So far, we usually have marked out affected image regions by hand, but it should be investigated whether there are automatic ways of dealing with it.
- We typically coadd several single exposures to get a final image. These frames usually have slightly different seeing disks. In the past we often simply rejected images that had the worst seeing. We have to investigate what effects the coaddition of images with different PSF properties has on the final shear measurement result.
- As a consequence of the last point, weak-lensing observations have mostly been done with a very compact dither pattern so far (this means taking consecutive exposures of a field with an offset of only a few arcseconds). This has the advantage of allowing a simple coaddition of the images with integer pixel shifts, and the effects of optical image distortions do not need to be taken into account during the image prereduction. First this preserves uncorrelated noise in the image pixels, second it preserves the smoothness of PSF anisotropy on the scale of single chips, and third it brings the practical advantage that we can still deal with single chips instead of very large images. Nevertheless, we think that this approach is not ideal. On the one hand accurate astrometry and photometry is done most easily if information from objects in the overlap between different chips is at hand. The aforementioned advantages may turn into problems in later analysis when the gaps between single chips lead to severe border effects, e.g. when we try to measure cosmic shear on large scales or in the search for dark matter concentrations as in our MOCAM data (see Fig. 43). So we will investigate whether we can overcome the effects of correlated pixel noise caused by remapping for optical distortions and discontinuities of the PSF anisotropy in the final coadded images.

The work of the last two sections marks the start of a new field within weak lensing studies: The determination of the total matter content in the universe. So far, the weak lens effect has

mainly been used to study known objects. The large area of wide-field imaging surveys allows a routine application of cluster-finding algorithms. As lensing is only sensitive to the mass of an object, it is the tool of choice to determine the cosmological mass function and to constrain cosmological models with the number of massive objects we observe in redshift slices (see Fig. 2). Besides creating a complete shear-selected sample of massive objects in the 1.7 square degree of data that we also used for the cosmic shear project, the nature of our discovery has to be studied with further observations. Already approved infrared observations may reveal a galaxy cluster at high redshift. If observing time is allocated, Chandra data will provide a more detailed X-ray map around the dark clump candidate. As a further step we might apply for multi-color observations around our discovery. Thus we can constrain the redshift distribution of the lensed galaxies (photometric redshifts) and hence the redshift of the matter concentration. We note that a lot of theoretical work on the expected number density of dark matter haloes with weak lensing observations has been done in the past years (see e.g. [40] and [55]). According to these studies, about 10 dark matter haloes should be detected with high significance in a moderately deep survey of 1 square degree.

The last section was about the direct measurement of moments of the cosmological density contrast. Also in this field the prospects for the near future are very exciting. Since the work on the first detection of cosmic shear finished, the Descartes project (see <http://terapix.iap.fr/Descart/>), having already provided the first data sets, has accumulated seven square degrees of data that have to be reduced and analysed. Besides lowering the cosmic variance, this data set will also allow the measurement of higher-order moments of the  $\kappa$  field with high significance. Especially the third moment, the skewness, is an interesting complement to the shear variance already obtained. In a simplified model [power-law model (index  $n$ ) for the power spectrum, all sources at the same redshift  $z_s$ , linear evolution of the density contrast, no cosmological constant], the measured shear variance and hence the variance of the  $\kappa$  field  $\sigma_\kappa$  behaves like

$$\sigma_\kappa \propto \sigma_8 \Omega_0^{0.75} z_s^{0.75} \left( \frac{\theta}{1'} \right)^{-\left(\frac{n+2}{n}\right)}, \quad (200)$$

while the behaviour of the skewness is

$$\frac{\langle \kappa^3 \rangle}{\langle \kappa^2 \rangle^2} \propto \Omega_0^{-0.8} z_s^{-1.33}. \quad (201)$$

Hence, a simultaneous measurement of variance and skewness allows an independent determination of  $\Omega_0$  and the normalisation of the power spectrum  $\sigma_8$ .

## References

- [1] Bacon, D., Refregier, A., Clowe D., Ellis, R., 2000, *Numerical Simulations of Weak Lensing Measurements* (MNRAS 318, 625-640)
- [2] Bacon, D., Refregier, A., Ellis, R., 2000, *Detection of Weak Gravitational Lensing by Large-scale Structure* (astro-ph/0003008; submitted to MNRAS)
- [3] Baugh, C.M., Coles, S., Frenk, C.S., 1996, *Faint galaxy counts as a function of morphological type in a hierarchical merger model* (MNRAS 282, L27-L32)
- [4] Bartelmann, M., Schneider, P., 2000, *Weak Gravitational Lensing* (astro-ph/9912508)
- [5] Beckers, J., 1995, (Scientific and Engineering Frontiers for 8-10m Telescopes, eds M. Iye & T. Nishimura, 303)
- [6] Bertin, E., Arnouts, S., 1996, *SExtractor: Software for source extraction* (A&A, 117, 393-404)
- [7] Binggeli, B., Sandage, A., Tarengi, M., 1984, *Studies of the Virgo Cluster. I - Photometry of 109 galaxies near the cluster center to serve as standards* (AJ 89, 64-82)
- [8] Blandford, R. D., Saust, A. B., Brainerd, T. G., Villumsen, J. V., 1991, *The distortion of distant galaxy images by large-scale structure* (MNRAS 251, 600-627)
- [9] Bond, J. R., Efstathiou, G., 1984, *Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter* (ApJ 285, L45-L48)
- [10] Born, M., Wolf, E., 1999, *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light* (Cambridge University Press 1999)
- [11] Bruzual G. A., Charlot S., 1993, *Spectral evolution of stellar populations using isochrone synthesis* (ApJ 405, 538-553)
- [12] Croft, A. C. C., Metzler, C., 2000, *Weak lensing surveys and the intrinsic correlation of galaxy ellipticities* (astro-ph/0005384, submitted to ApJ)
- [13] Dolag, K., 1997, *Korrelationsfunktion von Galaxien und Quasaren in beliebigen Friedmann-Lemaître-Kosmologien unter Berücksichtigung nichtlinearer Effekte* (Diploma thesis, TU München)
- [14] Driver, S.P., Windhorst, R.A., Griffiths, R.E., 1995, *The Contribution of Late-Type/Irrregulars to the Faint Galaxy Counts from Hubble Space Telescope Medium-Deep Survey Images* (ApJ 453, 48-64)
- [15] Erben, T., 1997, *Die Bestimmung von Galaxieneigenschaften durch Galaxy-Galaxy-Lensing* (Diploma thesis, TU München)
- [16] Erben, T., Van Waerbeke, L., Bertin, E., Mellier, Y., Schneider, P., 2000, *How accurately can we measure weak gravitational shear?*, (astro-ph/0007021, submitted to A&A)
- [17] Erben, T., Van Waerbeke, L., Mellier, Y., Schneider, P., Cuillandre, J. C., Castander, F. J., Dantel-Fort, M., 2000, *Mass detection of a matter concentration projected near the cluster Abell 1942: Dark clump or high-redshift cluster?* (A&A 355, 23-36)
- [18] Fliessbach, T., 1989, *Allgemeine Relativitätstheorie* (BI Wissenschaftsverlag)
- [19] Fort, B., Mellier, Y., 1994, *Arc(lets) in clusters of Galaxies* (A&A Review (1994)5:239-292)
- [20] Fort, B.; Mellier, Y.; Dantel-Fort, M. *Distribution of galaxies at large redshift and cosmological parameters from the magnification bias in CL 0024+1654* (A&A 321, 353-362)

- [21] Fried, D.L., 1966, (J. Opt. Soc. Am. 56, 1372)
- [22] Goenner, H., 1994, *Einführung in die Kosmologie* (Spektrum Akademischer Verlag)
- [23] Gray, M., Ellis, R. S., Refregier, A., Bézecourt, J., McMahon, R. G., Beckett, M. G., Mackay, C. D., Hoenig, M. D., 2000, *Infrared Observations of gravitational lensing in Abell 2219 with CIRSI* (astro-ph/0004161)
- [24] Hasan, H., Burrows, C.J., 1995, *Telescope Image Modeling (TIM)* (PASP 107, 289-298)
- [25] Heavens, A., Refregier A., Heymans, H., 2000, *Intrinsic Correlation of Galaxy Shapes: Implications for Weak Lensing Measurements* (astro-ph/0005269)
- [26] Hoekstra, H., Franx, M., Kuijken, K., Squires, G., 1998, *Weak lensing analysis of CL 1358+62 using Hubble Space Telescope observations* (ApJ 504, 636-660)
- [27] Hoekstra, H., Franx, M., Kuijken, K., 2000, *Hubble Space Telescope Weak-Lensing Study of the  $z=0.83$  Cluster MS 1054-03* (ApJ 532, 88-108)
- [28] Hudson, M., Gwyn, S. D. J., Dahle, H., Kaiser, N., 2000, *Galaxy-Galaxy Lensing in the Hubble Deep Field: The Halo Tully-Fisher Relation at Intermediate Redshift* (ApJ 503, 531-542)
- [29] Jain, B., Van Waerbeke, L., 2000, *Statistics of Dark Matter Halos from Gravitational Lensing* (ApJ 503, L1-4)
- [30] de Jong, R.S., Lacey, C., 1999, *The space density of spiral galaxies as a function of their luminosity, surface brightness and scalesize* (The Low Surface Brightness Universe, ASP Conference Series 170, 52)
- [31] Kaiser, N., 1999, *A New Estimator for Weak Lensing Observations* (astro-ph/9904003)
- [32] Kaiser, N., 1992, *Weak Gravitational Lensing of Distant Galaxies* (ApJ 388, 272-286)
- [33] Kaiser, N., Squires, G., 1993, *Mapping the dark matter with weak gravitational lensing* (APJ 404, 441-450)
- [34] Kaiser, N., Squires, G., Broadhurst, T., 1995, *A method for weak lensing observations* (ApJ 449, 460-475)
- [35] Kaiser, N., Squires, G., Fahlmann, G. G., Woods, D., 1994, *Mapping the dark matter in clusters* (Clusters of Galaxies [Proceedings of the XXIXth Rencontre de Moriond; ed. by Durret, F., Mazure, A., Tran Thanh Van, J.]
- [36] Kaiser, N., Wilson, G., Luppino, G., Dahle, H., 1999, *A Photometric Study of the Supercluster MS0302 with the UH8K CCD Camera: Image Processing and Object Catalogs* (astro-ph/9907229; submitted to PASP)
- [37] King, I.R., 1971, *The Profile of a Star Image* (PASP 83, 199-201)
- [38] Kneib, J. P., Mathez, G., Fort, B., Mellier, Y., Soucail, G., Longaretti, P. Y., 1994, *Redshift survey up to  $B_J=27$ : distance of gravitational arclets behind Abell 370* (A&A 286, 701-717)
- [39] Kneib, J. P., Mellier, Y., Fort, B., Mathez, G., 1993, *The Distribution of Dark Matter in Distant Cluster Lenses - Modelling A 370* (A&A 273, 367-376)
- [40] Kruse, G., 2000, *Statistische Untersuchungen zum schwachen kosmologischen Gravitationslinseneffekt* (PhD thesis; LMU München)
- [41] Kuijken, K., 1999, *Weak weak lensing: correcting weak shear measurements accurately for PSF anisotropy* (A&A 352, 355-362)

- [42] Lilly S. J., Tresse L., Hammer F., Crampton D., Lefevre O., 1995, *The Canada-France Redshift Survey. VI. Evolution of the Galaxy Luminosity Function to Z approximately 1* (ApJ 455, 108-124)
- [43] Luppino, G. A., Kaiser, N., 1997, *Detection of Weak Lensing by a Cluster at  $z = 0.83$* . (ApJ 475, 20-28)
- [44] Magliocchetti, M., Maddox, S. J., Lahav, O., Wall, J. V., 1998, *Variance and skewness in the FIRST survey* (MNRAS 300, 257-268)
- [45] Lupton, R., 1993, *Statistics in theory and practice* (Princeton University Press)
- [46] Mellier, Y., 1999, *Probing the Universe with Weak Lensing* (ARA&A 37, 127-189)
- [47] Narayan, R., Bartelmann, M., 1996, *Lectures on Gravitational Lensing* (Proceedings 1995 Jerusalem Winter School)
- [48] Navarro, J. F., Frenk, C. S., White, S. D. M., 1996, *The structure of Cold Dark Matter Halos* (APJ 462, 563)
- [49] Nonino, M, Bertin, E., da Costa, L., Deul, E., Erben, T., Olsen, L., Prandoni, I., Scodreggio, M., Wicenc, A., Wichmann, R., Benoist, C., Freudling, W., Guarnieri, M.D., Hook, I., 1999, *ESO Imaging Survey. I. Description of the survey, data reduction and reliability of the data* (A&AS 137, 51-74)
- [50] Oke, J. B., Sandage, A., 1968, *Energy distributions, K Corrections, and the Stebbins Whitford Effect for giant elliptical galaxies* (APJ 154, 21-32)
- [51] Peacock, J. A., Dodds, S. J., 1996, *Non-linear evolution of cosmological power spectra* (MNRAS 280, L19-L26)
- [52] Peebles, 1993, *Principles of physical cosmology* (Princeton University Press)
- [53] Press, W. H., Saul, A. T., Vetterling, T. V., Flannery, B. P., 1992, *Numerical Recipes in Fortran (second edition)* (Cambridge University Press)
- [54] Racine, R., 1996, *The Telescopic Point-Spread Function* (PASP 108, 699-705)
- [55] Reblinsky, K., 2000, *Projection effects in clusters of galaxies* (PhD-thesis, LMU München)
- [56] Roddier, F., 1981, *The Effects of Atmospheric Turbulence in Optical Astronomy* (Progress in Optics 19, 281)
- [57] Sandage, A. R., Kron, R. G., Longair, M. S., 1993, *The Deep Universe* (Springer Verlag)
- [58] Schechter, P., 1976, *An analytic expression for the luminosity function for galaxies*. (ApJ 203, 297-306)
- [59] Schneider, P., 1996, *Detection of (dark) matter concentrations via weak gravitational lensing* (MNRAS 283, 837-853)
- [60] Schneider, P., Ehlers, Falco, 1992, *Gravitational Lenses* (Springer Verlag)
- [61] Schneider, P., King, L. J., Erben, T., 2000, *Cluster mass profiles from weak lensing: constraints from shear and magnification information* (A&A 353, 41-56)
- [62] Schneider, P., Seitz, C., 1995, *Steps towards nonlinear cluster inversion through gravitational distortions I* (A&A 294, 411-431)
- [63] Schneider, P., Van Waerbeke, L., Jain, B., Kruse, G., 1998 *A new measure for cosmic shear* (MNRAS 296, 873-892)

- [64] Seitz, C., Schneider, P., 1995, *Steps towards nonlinear cluster inversion through gravitational distortions III: Including a redshift distribution of the sources* (A&A 318, 687-699)
- [65] Seitz, C., Kneib, J. P., Schneider, P.; Seitz, S. *The mass distribution of CL0939+4713 obtained from a 'weak' lensing analysis of a WFPC2 image* (A&A 314, 707-720)
- [66] Smith, R. C., 1995, *Observational Astrophysics* (Cambridge University Press)
- [67] Taylor, A. N., Dye, S., Broadhurst, T. J., Benitez, N., van Kampen, E. *Gravitational Lens Magnification and the Mass of Abell 1689* (ApJ, 501, 539-553)
- [68] Umetsu, K., Futamase, T., 2000, *Detection of dark matter concentrations in the field of Cl 1604+4304 from weak lensing analysis* (ApJ Lett., 539, L5-L8)
- [69] Villumsen, J. V., Freudling, W., Da Costa, L. N., 1997, *Clustering of Galaxies in the Hubble Deep Field* (ApJ 481, 578-586)
- [70] Van Waerbeke, L., Bernardeau, F., Mellier, Y., 1999, *Efficiency of weak lensing surveys to probe cosmological models* (A&A 342, 15-33)
- [71] Van Waerbeke, L., Mellier, Y., Erben, T., Cuillandre, J. C., Bernardeau, F., Maoli, R., Bertin, E., Mc Cracken, H. J., Le Fevre, O., Fort, B., Dantel-Fort, M., Jain, B., Schneider, P., 2000, *Detection of correlated galaxy ellipticities on CFHT data: first evidence for gravitational lensing by large-scale structures* (A&A 358, 30)
- [72] Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., Bernstein, G., 2000, *Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales* (Nature 405, 143-148)

## A The Stuff program

In this Appendix a short description of the *Stuff* program is given. It was used to create catalogs with intrinsic galaxy properties for our simulations with the KSB algorithm. The program is freely available at: <ftp://geveor.iap.fr/pub/stuff>.

Given the sky dimensions of the intended simulations, the program distributes galaxies in redshift space that is subdivided into bins. For every bin, representing a volume element according to the specified cosmology (we used  $H_0 = 65\text{km}/(\text{s Mpc})$ ,  $\Omega_0 = 0.3$  and  $\Omega_\Lambda = 0.7$  in all our simulations), the number of galaxies for the Hubble types E, S0, Sab, Sbc, Scd and Sdm/Irr is determined from a Poisson distribution assuming a non-evolving Schechter luminosity function (see [58]). The different galaxy types are simulated by linearly adding exponential  $\mu_d(r)$  (disk component) and de Vaucouleur profiles  $\mu_b(r)$  (bulge components) in different ratios.

$$\mu_b(r) = M_b + 8.3268 \left(\frac{r}{r_b}\right)^{1/4} + 5 \log r_b + 16.6337; \mu_d(r) = M_d + 1.8222 \left(\frac{r}{r_d}\right) + 5 \log r_d + 0.8710, \quad (202)$$

where  $\mu_b$ ,  $\mu_d$  are the surface brightnesses in  $\text{mag}/\text{pc}^2$ , and  $M_b$ ,  $M_d$  are the absolute magnitudes of the bulge and the disk components, respectively. The distributions of scale radii  $r_b$  and  $r_d$  are fixed by an empirical relation given in Binggeli et al. 1994 (see [7]) and a semi-analytical model from de Jong & Lacey 1999 (see [30]) relating these quantities to the absolute magnitudes. The galaxies are assigned a random disk inclination angle and a position angle which define the intrinsic ellipticities of our objects. The output of the program is a catalog of galaxy positions, apparent magnitudes, semi-minor and major axes, and position angles for disks and bulges. These catalogs are sheared (disks and bulges are sheared separately!) before being processed with SkyMaker.

## B Object detection and selection

In this section we describe the full automatic procedure to go from image frames to a final object catalog for shear analysis used in our KSB simulations. We note that our main intention was to obtain a catalog ensuring reliable measurements. We have demonstrated with Gaussian profiles that our analysis with isolated objects in a semi-analytical treatment and with those detected in FITS images gives very comparable results (see Fig. 37). Our procedure contains several conservative rejection criteria for objects and it is not optimised to make the maximal use of data in terms of number density of objects. The procedure consists of the following steps:

1. Objects were detected with the 'hfindpeaks' algorithm from Nick Kaiser: The image is smoothed with Mexican top hat filters  $W_d(|\vec{\theta}|)$

$$W_d(|\vec{\theta}|) = \frac{1}{2r_g^2\pi} \left[ \exp\left(-\frac{|\vec{\theta}|^2}{2r_g^2}\right) - \frac{1}{r_f^2} \exp\left(-\frac{|\vec{\theta}|^2}{2r_f^2 r_g^2}\right) \right] \quad (203)$$

with increasing filter radii  $r_g$ . In every smoothed image, the peaks are detected and linked with the peaks found in previous smoothings. Hereby peaks are assumed to belong to the same object if their positions coincide within  $r_g$ . In this way a peak trajectory is built up for every potential object. For every peak, a signal-to-noise ratio  $\nu$  is calculated where the signal is the peak value in the smoothed image and the noise  $\sigma_d$  is given by

$$\sigma_d = \sigma_{\text{sky}} \sqrt{\int W_d^2(|\vec{\theta}|) d^2\theta} = \sigma_{\text{sky}} \sqrt{\frac{1}{2\sqrt{\pi}r_g} \left(1 + \frac{1}{r_f^2} - \frac{4}{1+r_f^2}\right)}, \quad (204)$$

with  $r_f = 2$ . Fig. 75 shows a comparison between this signal-to-noise estimate and ours from eq. (157). Afterwards, peaks with the highest  $\nu$  value for each trajectory identify objects (peaks with  $\nu < 4$  are immediately rejected). This size  $r_g$  is used in all the following analyses and the pixel centre of the peak position is taken as a starting point for the object centre

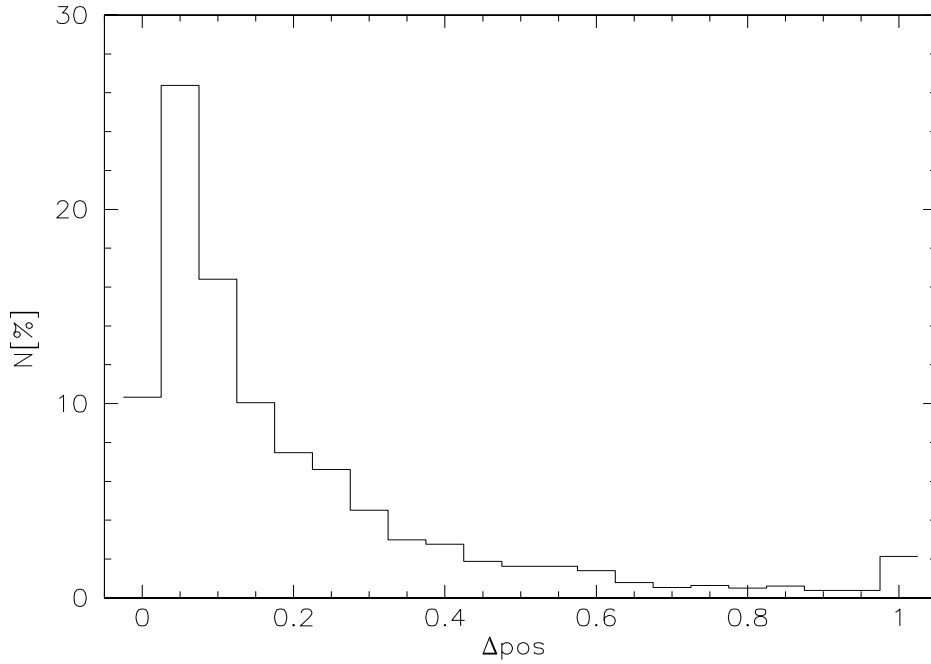


Figure 74: The distribution of the positional differences of the ‘hfindpeaks’ position estimator and ours from eq. (156). We see a clear peak below 0.1 pixel and so we conservatively excluded all objects with a higher difference. See the text for more details.

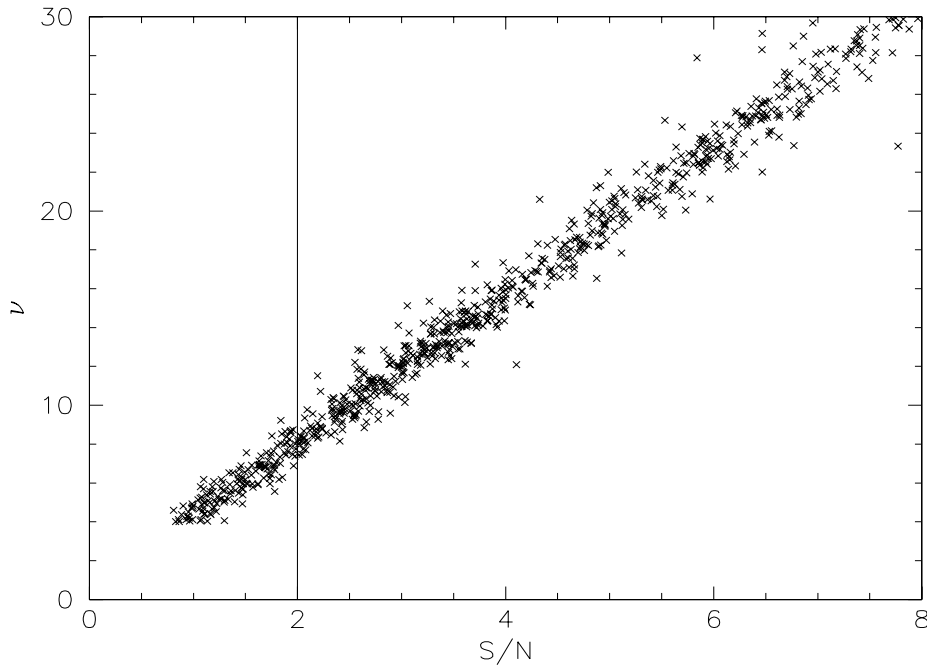


Figure 75: A comparison between the ‘hfindpeaks’ signal-to-noise estimate  $\nu$  and ours from eq. (157). The relation is a straight line and our criterion for a good object ( $S/N > 2$ ) corresponds to  $\nu \approx 7$ .



determination. To get a more accurate object centre, a Newton-Raphson step is performed and the  $x$  and  $y$  pixel-positions are corrected from the pixel centre by  $\delta_x$  and  $\delta_y$

$$\delta_x = -\frac{I_x^s}{I_{xx}^s}; \quad \delta_y = -\frac{I_y^s}{I_{yy}^s} \quad (205)$$

where  $I_x^s, I_y^s$  and  $I_{xx}^s, I_{yy}^s$  are the first and second derivatives of the smoothed light profiles along  $x$  and  $y$ .

2. From the peaks (objects) found in the previous step, all relevant quantities (like  $P^{\text{sh}}, P^{\text{sm}}$ ) are calculated separately with the original, unsmoothed image.
3. From the catalog generated in the previous step, we first removed objects with obvious problems during the analysis process: (1) Objects closer than  $3r_g$  to the border of the image; (2) objects where one of the eigenvalues  $A^2, B^2$  from the  $Q_{ij}$  tensor was negative.

$$\begin{pmatrix} A^2 \\ B^2 \end{pmatrix} = \frac{Q_{11} + Q_{22}}{2} \begin{pmatrix} + \\ - \end{pmatrix} \sqrt{\left(\frac{Q_{11} - Q_{22}}{2}\right)^2 + 4Q_{12}^2}; \quad (206)$$

(3) the object has a total negative flux. (4) Fig. 74 shows the distribution of the difference between the centroid estimates of 'hfindpeaks' and ours from eq. (156). All objects with a difference larger than 0.1 pixel were excluded. We found that this is an efficient way to reject blended objects, since one centroid estimator is based on the smoothed and the other on the unsmoothed image.

4. Objects with a neighbour within  $3r_g$  are rejected.
5. Stars are preselected using the star branch of the  $r_h - m$  diagram and polynomials for the two components of  $\mathbf{q}^*$  are calculated in the following way: a preliminary fit is done for  $q_1^*$  and  $q_2^*$  using a  $\chi^2$  minimisation

$$\chi_j^2 = \sum_{i=1}^{N^*} (q_j^*(x_i, y_i) - p_j(a_k, x_i, y_i))^2, \quad (207)$$

where  $N^*$  is the number of preselected stars,  $j = 1, 2$ ,  $x_i, y_i$  are the positions of the stars and  $p_j$  are two-dimensional, second order polynomials with six unknown parameters  $a_k$ . After determination of  $a_k$  we calculated the expected error  $\sigma_{q_j}$  for every  $q_j^*$  by

$$\sigma_{q_j^*} = \sqrt{\frac{\chi_{j\text{min}}^2}{N^* - 1}}, \quad (208)$$

where  $\chi_{j\text{min}}^2$  is the minimum of  $\chi_j^2$  at the fitted parameters  $a_k$ . Stars which deviate by more than  $1\sigma_{q_j}$  in any of the two components  $q_1$  or  $q_2$  are rejected, and the fit is repeated for the final polynomials.

6. The final sample of stars is reprocessed nine times with filter scales  $r_i = 1..9$ , in order to match all the possible galaxy sizes. After each processing, objects with problems according to step (3) are still rejected. From the remaining stars we calculated the mean of  $\langle \text{tr}[P^{*\text{sh}}] / \text{tr}[P^{*\text{sm}}] \rangle$  and we used this value for the PSF correction.
7. We now have all the quantities in hand to calculate scalar and tensor shear estimates for every objects. Hereby we considered only objects as galaxies whose half light radius was larger than the stellar locus (see Fig. 38).
8. Weights for the tensor and scalar  $\mathbf{g}$  are calculated, which is an important ingredient for the two shear estimators as described in section 6.5. With the whole procedure we end up with a number density of about 30 galaxies/arcmin<sup>2</sup>.

## Acknowledgements

Although this thesis only has one author at the end there are a lot of people without whom I never could have done it.

First and most important I have to thank my mother, Frau Christa Erben, who has supported me in every possible respect during the many years of my education.

My supervisor, Prof. Dr. Peter Schneider, has contributed most to the successful outcome of my work. He always had time when I needed it and did everything to support my scientific career. He intended to get me going on a Galaxy-Galaxy lensing project long before most of the wide-field imaging cameras had become available. For this I spent the first two years of my thesis in the ESO imaging survey team that should perform a unique wide-field survey with a small-field camera at that time. Although the survey failed to provide me with the necessary data material and although I found the work I did there rather useless at the time I did it and the whole project often chaotic, I appreciate it now as an invaluable and extraordinary important experience. It is the basis for the well founded experience on all aspects around astronomical data I now have. I profit a lot from the experience of many people who shared and still share their knowledge with me. They are too numerous to name them all and I mention in particular: Luiz da Costa, who accepted me as a member of his team; Rainer Wichmann, who is a very good friend now and who gave me the pleasure to be witness of his marriage; Mario Nonino, who shared his enormous knowledge about data reduction and script writing; Simone Zaggia and Roeland Rengelink, who taught me how to observe; Erik Deul, Richard Hook and Emmanuel Bertin, from whom I know everything about pipeline processing.

Also in my time at MPA I have made a lot of new friends and had the pleasure to work with very nice colleagues. I want to especially mention here the core members of the lens group: Lindsay King, Matthias Bartelmann, Stella Seitz, Guido Kruse, Katrin Reblinsky, Hannelore Haemmerle, Douglas Clowe, Lara Collodel and Mischa Schirmer.

Outside MPA and ESO my special thanks go to Yannick Mellier and Ludovick van Waerbeke. All major observational projects of this thesis would not have been possible without their help and support. From the very beginning they shared their knowledge and data with me and accepted me as a full member of their team.

Last, I have to thank my friends outside science who have nothing to do with astronomy and who gave me the necessary diversion from my work.

# Lebenslauf

Thomas Erben

## Persönliche Daten

---

<i>geboren am</i>	18. März 1970 in München
<i>Staatsangehörigkeit</i>	deutsch
<i>Familienstand</i>	ledig

## Wissenschaftliche Ausbildung

---

seit Jan. 1999	Promotionsstipendium am Max-Planck-Institut für Astrophysik in Garching bei München
März 1997 - Dez. 1998	Mitglied des ESO Imaging Survey Teams an der Europäischen Südsternwarte in Garching
Feb. 1996 – Feb. 1997	Diplomarbeit am Max-Planck-Institut für Astrophysik über <i>Bestimmung von Galaxieneigenschaften durch Galaxy-Galaxy-Lensing</i> betreut von Prof. Dr. Wolfgang Hillebrandt und Dr. habil. Peter Schneider
Sept. 1992	Vordiplom
Nov. 1990 – Feb. 1997	Studium der allgemeinen Physik an der Technischen Universität in München.

## Schulausbildung

---

1980 – 1989	Pater-Ruppert Mayer Gymnasium in Pullach bei München. Abschluss: Allgemeine Hochschulreife im humanistischen Zweig des Gymnasiums
1976 – 1980	Grundschule in München