

Advanced gravitational lensing techniques for precision cosmology

THÈSE N° 8338 (2018)

PRÉSENTÉE LE 26 JANVIER 2018
À LA FACULTÉ DES SCIENCES DE BASE
LABORATOIRE D'ASTROPHYSIQUE
PROGRAMME DOCTORAL EN PHYSIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Markus REXROTH

acceptée sur proposition du jury:

Prof. H. M. Rønnow, président du jury
Prof. J.-P. R. Kneib, directeur de thèse
Prof. A. Refregier, rapporteur
Dr E. Jullo, rapporteur
Prof. A. Bay, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2018

The eternal mystery of the world is its comprehensibility
— Albert Einstein

Acknowledgements

Completing this PhD thesis has been a wonderful and exciting experience. I have so many people to thank. First of all of course my PhD advisor, Jean-Paul Kneib, for all his support and the possibility to work on such a large range of projects - and to even go observing in the Chilean desert. I thank the members of the jury committee for generously taking their spare time, Alexandre Refregier, Eric Jullo, Henrik Rønnow, and Aurelio Bay. I gratefully acknowledge my wonderful and supportive friends and colleagues at EPFL and the Geneva observatory, Thibault, Olga, Christoph, Rémy, Vivien, Tim, Loïc, Marc, Frédéric, Pierre, Baptiste, Arthur, William, Claudio, et al. All research would come to a halt without the great technical and administrative staff, Yves, Olivier, Gilles, Sophie, Carol, and Claire. I am also very grateful to my friends in Switzerland, Germany, and the rest of the world for their friendship and support, Mark, Francis, Szymon, Constantin, Alex, Manuel, Markus, Philip, Christoph, Johann, Adrian, Austen, Vishal, Justin, Crystal, Ana, and all the others who I cannot mention here by name due to space limits. A big thanks to the wonderful scientists I had the privilege to collaborate with - Johan Richard, Mathilde Jauzac, Priya Natarajan, Gilles Fourestey, Hakim Atek, the LENSTOOL team, David Goldberg, Dominique Eckert, et al. If I should have forgotten to mention someone in this acknowledgment, please be assured that this was accidental and that I am very grateful for your support and friendship. Last, but most importantly, I am infinitely indebted to my parents, Horst and Ute, my brother Jonas, my grandparents, Inge and Gisela, aunts and uncles, Lothar, Hildegard, Beate, Ernst, Sonja, Mathias, and Waltraud, and my cousins Lea and David. It is a true blessing to be a part of this family.

Lausanne, 30 October 2017

M.R.

Abstract

Gravitational lensing - the deflection of light by gravity - has greatly developed since its famous first observation in 1919, which validated Einstein's General Theory of Relativity. The strength of this effect does not depend on the nature of the mass which produces the gravitational field and thus it is a great tool to weigh both the visible and the invisible parts of the universe. Consequently gravitational lensing has become a pillar of observational cosmology over the last decades, and it is used to study the nature of Dark Matter and Dark Energy, the two mysterious quantities which dominate our universe, but are not yet understood by physics theory. The success of this endeavor rests on a thorough understanding of lensing theory and observations, including their systematics, the availability of a sufficient amount of precise data, and the development of efficient software to precisely measure these lensing effects in digital astronomical images.

This thesis presents advanced techniques which can improve several of these areas. In the first part, we develop a new theoretical method to break the mass-sheet degeneracy, which prevents accurate mass determinations from lensing observations. The second part focuses on spectroscopic data from MUSE, a second-generation Integral-Field Spectrograph installed on one of the largest ground-based telescopes on earth. We present a pipeline which permits the efficient determination of the redshift of a source observed by MUSE. The redshift indicates the distance of the source from us and depends on the expansion of the universe. In addition, we use MUSE observations of a galaxy cluster to improve the determination of its total weight, including the dominant Dark Matter component. In the third part of this thesis, we investigate how we can accelerate the computation of these mass maps. In the era of big data and large surveys, computing efficiency is key to obtaining new scientific insights. We use High Performance Computing techniques like graphics card acceleration to improve the code performance and we develop a method which harnesses extra performance from using single precision without losing the required accuracy. In the last section, we present first results from measuring flexion, a higher order lensing effect which could substantially increase the resolution of lensing mass maps and thus lead to a sharper view of structure in the universe.

Key words: astrophysics, cosmology, cosmological parameter, dark matter, flexion, gravitational lensing, graphics card acceleration, high performance computing, lens modeling, mass-sheet degeneracy, redshift, spectroscopy, software, theory, wavelets

Résumé

L'effet de lentille gravitationnelle - la déflexion de la lumière par la gravité - s'est beaucoup développé depuis la première observation en 1919, la première validation expérimentale de la théorie de la Relativité Générale d'Einstein. L'importance de l'effet ne dépend pas de la nature de l'objet qui agit comme lentille, ce qui en fait un outil puissant pour peser à la fois l'univers visible et sombre. L'effet de lentille gravitationnel est donc devenu un des piliers de la cosmologie observationnelle des dernières décennies. Il est utilisé pour étudier la nature de la matière sombre et de l'énergie sombre, deux quantités mystérieuses qui dominent notre univers, mais qui ne sont toujours pas comprises. Leur compréhension passe par une connaissance approfondie de la théorie de l'effet lentille et des observations, y compris de leur systématiques, l'accès à un volume suffisant de données, et le développement d'outils d'analyses efficaces qui mesurent précisément ces effets dans les images astrophysiques digitales.

Cette thèse présente des techniques avancées qui permettent des avancées dans plusieurs des domaines cités au-dessus. Dans la première partie, nous développons une nouvelle méthode théorique pour lever la dégénérescence de masse du déflecteur, qui empêche une mesure exacte de la masse dans des observations lentillées. La deuxième partie se penche sur des données spectroscopiques provenant de MUSE, un instrument de deuxième génération qui permet la spectrographie intégrale de champ, installé sur l'un des plus grands télescopes au sol. Nous présentons une méthode qui détermine efficacement le décalage vers le rouge d'une source observée par MUSE. Le décalage vers le rouge indique la distance nous séparant de la source, ce qui dépend de l'expansion de l'univers. De plus, nous utilisons des observations MUSE d'un amas de galaxies pour améliorer la détermination de sa masse totale, y compris la partie sombre dominante. Dans la troisième partie, nous investiguons comment nous pouvons accélérer la modélisation numérique de ces cartes de masse. Nous utilisons des principes de calculs à haute performance, comme l'accélération graphique, pour améliorer les performances et développons une méthode qui exploite toute la performance des calculs à précision simple sans perdre en précision. Dans la dernière section, nous présentons les premiers résultats de mesure de flexion, un effet d'ordre élevé, qui pourrait augmenter sensiblement la résolution des cartes de masses et donc conduire à une meilleure compréhension de la structure de l'univers.

Mots clefs : astrophysique, cosmologie, paramètre cosmologique, matière sombre, flexion, effet de lentille gravitationnel, accélération graphique, calculs à haute performance, modélisation de lentilles, dégénérescence de masse du déflecteur, décalage vers le rouge, spectroscopie, logiciel, théorie, décomposition en ondelettes

Zusammenfassung

Der Gravitationslinseneffekt - die Ablenkung des Lichts durch die Schwerkraft - hat sich seit seiner ersten Beobachtung im Jahre 1919, die die Einstein'sche Allgemeine Relativitätstheorie bestätigte, enorm entwickelt. Die Stärke dieses Effektes hängt nicht von der Natur der Masse ab, die das Gravitationsfeld erzeugt, und aus diesem Grund ist dieser ein grossartiges Werkzeug um sowohl die sichtbaren als auch die unsichtbaren Teile des Universums zu wiegen. Daher wurde der Gravitationslinseneffekt in den letzten Jahrzehnten zu einem Stützpfiler der beobachtenden Kosmologie und er wird benutzt um die Beschaffenheit der dunklen Materie und der dunklen Energie zu untersuchen. Diese zwei mysteriösen Messgrössen dominieren unser Universum, aber sie können von der physikalischen Theorie noch nicht hinreichend erklärt werden. Der Erfolg dieses Forschungsunterfangens beruht auf einem genauen Verständnis von Theorie und Beobachtung des Linseneffektes, einschliesslich ihrer Systematiken, auf der Verfügbarkeit einer ausreichend grossen Menge von präzisen Daten, und auf der Entwicklung von effizienter Software um den Linseneffekt in digitalen astronomischen Bildern präzise zu messen.

Diese Arbeit präsentiert fortgeschrittene Techniken, die mehrere dieser Bereiche verbessern können. Im ersten Teil entwickeln wir eine neue theoretische Methode um die Massenflächenentartung aufzuheben, die akkurate Massenbestimmungen durch Beobachtungen des Linseneffektes verhindert. Der zweite Teil konzentriert sich auf spektroskopische Daten von MUSE, einem Spektrographen der zweiten Generation für das gesamte Sichtfeld, der an einem der grössten terrestrischen Teleskope der Welt angebracht ist. Wir präsentieren eine Pipeline, die die effiziente Bestimmung der Rotverschiebung einer mit MUSE beobachteten Lichtquelle erlaubt. Die Rotverschiebung gibt die Entfernung der Lichtquelle von uns an und hängt von der Expansion des Universums ab. Zusätzlich benutzen wir Beobachtungen eines Galaxienhaufens mit MUSE, um die Messung von dessen Gesamtgewicht einschliesslich des dominanten dunkle Materie Anteils zu verbessern. Im dritten Teil dieser Arbeit untersuchen wir, wie wir die Berechnung dieser Massenlandkarten beschleunigen können. In der Ära grosser Datenmengen und umfangreicher Vermessungen ist die Berechnungseffizienz ein Schlüsselement um neue wissenschaftliche Erkenntnisse zu gewinnen. Wir benutzen Techniken des Hochleistungsrechnens wie Grafikkartenbeschleunigung um die Codeleistung zu verbessern und wir entwickeln eine Methode, die zusätzliche Leistung gewinnt, indem sie einfache Genauigkeit benutzt ohne die notwendige Exaktheit zu verlieren. Im letzten Teil präsentieren wir erste Ergebnisse von einer Messung der Beugung. Diese ist ein Linseneffekt höherer Ordnung, der die Auflösung der Massenlandkarten, die durch den Linseneffekt gewonnen werden, beträchtlich erhöhen und

Zusammenfassung

damit zu einer schärferen Sicht auf die Strukturen im Universum führen könnte.

Stichwörter: Astrophysik, Kosmologie, kosmologischer Parameter, dunkle Materie, Beugung, Gravitationslinseneffekt, Grafikkartenbeschleunigung, Hochleistungsrechnen, Linsenmodellierung, Massenflächenentartung, Rotverschiebung, Spektroskopie, Software, Theorie, Wavelets

Contents

Acknowledgments	i
Abstract (English, Français, & Deutsch)	iii
List of figures	xi
List of tables	xiii
1 Introduction	1
2 Modern cosmology	3
2.1 The cosmological standard model in a nutshell	3
2.2 Are we there yet? - The road to the standard model	9
2.3 Open questions	31
3 Cosmology from gravitational lensing	35
3.1 A concise history of gravitational lensing	35
3.2 Gravitational lensing theory	37
3.3 Cosmological constraints	46
3.4 The road ahead: Precision cosmology from lensing and its challenges	50
4 The mass-sheet degeneracy: A limiting factor?	53
4.1 Theoretical foundation	53
4.2 Impact on lens models and proposed techniques to lift the degeneracy	56
4.3 A new method to break the degeneracy	58
5 Improving lens models with spectroscopic redshifts	81
5.1 Photometric vs spectroscopic redshifts	81
5.2 MUSE: A spectroscopic redshift machine	82
5.2.1 The MUSE instrument	82
5.2.2 MUSE observations of cluster lenses	85
5.3 IFS-RedEx: A redshift extraction pipeline with wavelet filtering	86
5.4 An improved lens model for the galaxy cluster MACSJ1931	94
5.4.1 Introduction	94
5.4.2 MUSE data and reduction	95

Contents

5.4.3	Spectrum and redshift extraction	95
5.4.4	Redshifts and line emission maps	96
5.4.5	Lens model	98
6	High Performance Computing for cluster lensing	113
6.1	Wanted: A power boost for LENSTOOL	113
6.1.1	LENSTOOL, a successful lens modeling software	113
6.1.2	The strong lensing algorithm	115
6.2	Single vs Double Precision and GPUs vs CPUs	119
6.2.1	CPU and GPU implementations	119
6.2.2	Single and double precision	120
6.2.3	Computing finite precision errors for strong lensing	122
6.2.4	Fixing the missing accuracy close to critical lines	123
6.2.5	Early performance measurements and upcoming work	124
6.3	Appendix: Finite machine precision error in $\nabla\varphi_\epsilon$ computation for SIE	125
6.3.1	Error propagation rules	126
6.3.2	Error computation	127
6.3.3	Upper error bounds for cluster- and galaxy-scale SIE lenses	133
7	Flexion: A promising frontier	137
7.1	Increasing the lens model resolution with flexion	137
7.2	How to measure the curvature of a few pixels?	138
7.3	Flexion measurements from an automated pipeline	138
8	Outlook and conclusion	143
	Bibliography	170
	Acronyms and abbreviations	171
	Symbols and mathematical notation	175
	Photographic credits	177
	List of publications	181
	Curriculum Vitae	183

List of Figures

2.1	Timeline of the universe	4
2.2	Standard model of particle physics	5
2.3	Illustration of a baryon	5
2.4	Cosmic Microwave Background Anisotropies	7
2.5	Large Scale Structure of the universe zoom-in	8
2.6	Energy density content of the universe pie chart	9
2.7	Spacetime curvature in General Relativity	11
2.8	Possible spatial geometries in FLRW universe models	16
2.9	Evolution of CMB observations	31
3.1	Sketch of a gravitational lens system	38
3.2	Illustration of gravitational lensing shear and flexion	42
3.3	Strong lensing galaxy cluster Abell 370	45
5.1	The MUSE instrument on the VLT	83
5.2	MUSE datacube	84
5.3	MACSJ1931 with detected cluster members and background sources	97
5.4	Spectra and source locations for six high redshift sources	105
5.5	H α map of the BCG	108
5.6	OII emission map for a galaxy at $z = 0.8$	109
5.7	OII emission maps for “jellyfish” galaxies	110
5.8	Critical lines, caustics, and multiple images for MACSJ1931	111
5.9	Amplification map for the MACSJ1931 lens model	112
6.1	LENSTOOL strong lensing algorithm	116
6.2	Early performance benchmark results	126
7.1	Simulation of galaxies with strong flexion	139



List of Tables

5.1	List of multiple image systems	100
5.2	Best-fit parameters for the MACSJ1931 lens model	102
5.3	Spectroscopic redshifts for the sources detected with MUSE	103
6.1	Theoretical maximum computing performance of used CPU and GPUs	122

1 Introduction

The development of Albert Einstein's Theory of General Relativity a bit over a hundred years ago was a milestone for physics. It describes gravity in a beautiful, geometric framework and it has led to several wonderful applications - not only in astronomy, but also in tools which we use on a daily basis like the GPS navigation system.

One of these applications, namely gravitational lensing, has truly prospered over the last decades. It is almost a miracle of technological and scientific progress that we can measure tiny distortions in the shapes of galaxies which existed billions of years in the past - and we can even use it to study the universe! Einstein himself did not see this coming - it was the astronomer Fritz Zwicky who predicted that gravitational lensing must necessarily be observed and will be very useful to study Dark Matter and cosmology.

Indeed, he was absolutely right. Today, astronomers are using the strong lensing effect, which greatly distorts observed images and even produces multiple images of the same source, to study the distribution of Dark Matter in galaxies and clusters of galaxies - and subsequently they use the magnification effect of these clusters as cosmic telescopes to look even deeper into the universe. Other scientists use the weak lensing effect, tiny distortions of the galactic isophotes, to study the large scale distribution of matter in the universe and infer knowledge about Dark Energy, the mysterious energy field which accelerates the expansion of the universe. It is even feasible to observe colliding galaxy clusters and to compute upper bounds on the Dark Matter self-interaction by using the gravitational lensing effect.

None of these applications, however, would be possible without the advancement of technology in recent decades. Ever increasing telescope sizes, CCD cameras, and adaptive optics deliver fantastic observational data which is the fuel of lensing science. Similarly, the exponential development of computer power facilitated the use of more and more sophisticated algorithms and image processing techniques, which have in turn dramatically improved the scientific yield from lensing.

Chapter 1. Introduction

The last decade ushered in the age of precision cosmology, and when the statistical error bars shrink to the percent-level, systematic effects start to become relevant. The current decade leads us on a straight path towards big data cosmology and with the upcoming telescope generations this will soon become a reality - at the very latest once the SKA radio telescope is operational, which will produce more data than the whole current Internet.

For these reasons, we have investigated several areas of gravitational lensing - theory, observation, modeling, and shape measurement - and we have developed advanced techniques to ameliorate some of their difficulties.

The outline for this thesis is as follows. In the next chapter, we give an introduction to cosmology and the history of the universe. We introduce the mathematical formalism which is required in the remainder of this thesis. In the third chapter, we introduce gravitational lensing and its mathematical formalism. After that, we present a new theoretical method to break the mass-sheet degeneracy of gravitational lensing. In chapter 5, we present a software pipeline for the efficient extraction of redshifts from astronomical data and we show that we can use these redshifts to construct an improved lens model of a galaxy cluster. In chapter six, we study the potential of High Performance Computing techniques to accelerate a widely used lens modeling software. In chapter seven, we present first results for a measurement software for flexion, a higher order lensing effect with great potential which is difficult to measure. We conclude in chapter eight.

2 Modern cosmology

2.1 The cosmological standard model in a nutshell

In this section we give a concise overview of the history of the universe according to the standard model of modern cosmology. In the next section, we will show how we have found out what we know and we introduce the necessary mathematical formalism along the way. The history and development of our universe is summarized in figure 2.1. The universe began 13.8 billion years ago with a process that is called the “Big Bang”. The newborn universe is a truly special place. In fact, its physical properties are so extreme that the currently known laws of physics cannot describe it - we need a unification of general relativity and quantum mechanics for that purpose, and this “Theory of Everything” has yet to be found. Right after its birth, the universe expands exponentially during the so-called “inflationary phase”. Within $\approx 10^{-30}$ seconds, it reaches a size that is many orders of magnitudes greater than the total universe observable today (Linde, 2008). As a result, any inhomogeneities and curvature of the universe are exponentially diluted. The cosmos which we observe today is flat and homogeneous on the largest scales (Linde, 2008). The value of the inflation field at a given position can deviate slightly from the average value in the cosmos due to quantum fluctuations. Therefore the duration of inflation can slightly differ at two separate locations in the universe. When inflation ends, the density of the universe starts to decrease with time t as $\propto t^{-2}$. Consequently, some parts of the universe will have lost less density and some will have lost more density at a given time and thus they have a slightly higher or lower density than the average (Linde, 2008). During inflation, all energy is concentrated in the inflation field. Soon after the end of the inflationary phase, the field begins to oscillate near the minimum of its effective potential and it produces many elementary particles, which in turn interact with each other and reach a state of thermal equilibrium (Linde, 2008).

Once the inflationary phase is completed, the universe continues to expand, but at a much slower rate. Consequently, the universe cools with time. Once the temperature has dropped so much that no new baryon-anti-baryon pairs are created, essentially all anti-baryons react with baryons and disappear, but a small rest of baryons survives due to the baryon-anti-baryon

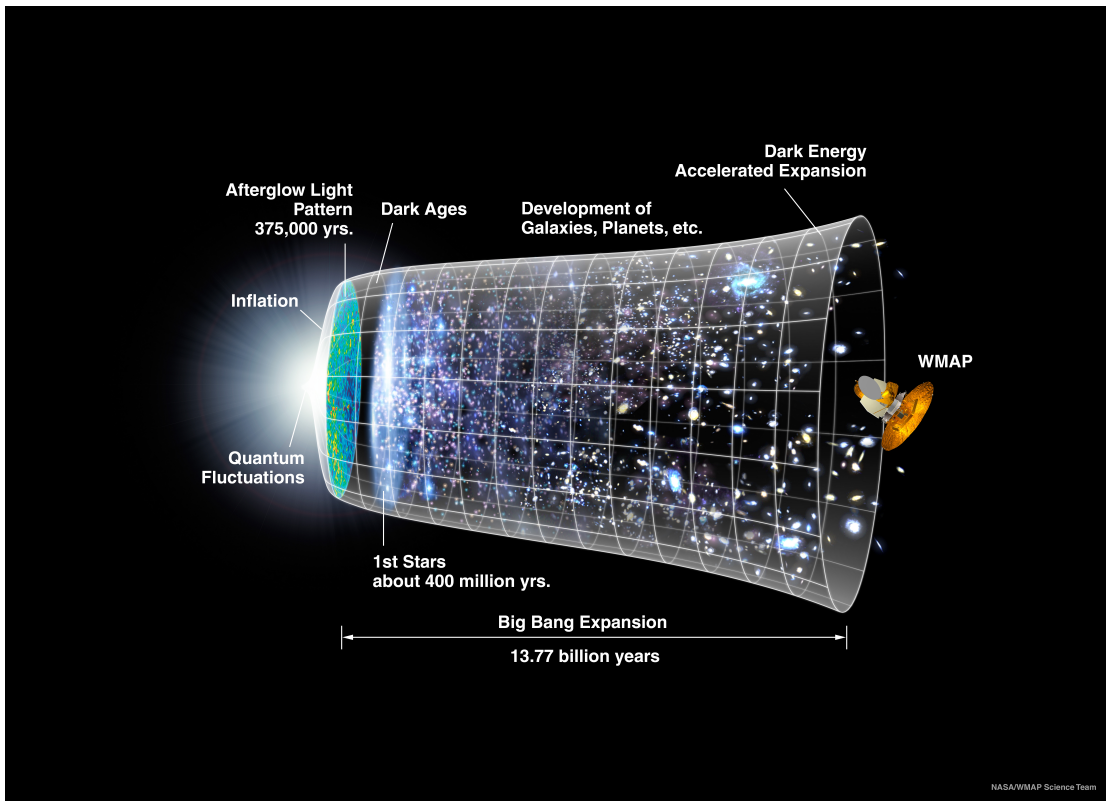


Figure 2.1 – Timeline of the universe in the cosmological standard model (Image credit: NASA / WMAP Science Team)

asymmetry (Riotta and Trodden, 1999) (see figures 2.2 and 2.3 for an overview of the standard model of particle physics and baryons). At these energy scales, the unified electroweak force is already split into the electromagnetic and the weak force (Riotta and Trodden, 1999). Now we have a plasma consisting of protons, neutrons, electrons, positrons, photons, neutrinos, and anti-neutrinos in thermodynamic equilibrium (Iocco et al., 2009). The energy density of the universe is dominated by radiation and relativistic particles (Iocco et al., 2009). The cosmos continues to expand and cool. When the temperature drops below 2 – 3 MeV (approximately 23 – 34 billion K), the interaction rates of the three flavors of neutrinos with the photons and the electron-positron pairs become lower than the expansion rate of the universe and the neutrinos decouple from them (Steigman, 2007). The electron-neutrinos and anti-neutrinos continue to interact with the baryons until the universe is about 1 second old and the temperature sinks below 0.8 MeV (≈ 9 billion K) (Steigman, 2007). Now all the neutrinos have completely decoupled and are free. Neutrons and protons cease to be in equilibrium and the neutron-to-proton density ratio freezes out at 1/6 (Steigman, 2007). The ratio lowers with time due to the decay of free neutrons until the beginning of nucleosynthesis (Iocco et al., 2009). While the temperature is already below the deuterium binding energy of 2.2 MeV at this stage, the large photon-nucleon density ratio of about 10^9 delays deuterium synthesis until the photo-dissociation becomes ineffective, the so-called deuterium bottleneck (Iocco et al., 2009).

2.1. The cosmological standard model in a nutshell

Once the temperature decreases below the electron rest mass of 511 keV (≈ 6 billion K), the electron-positron pairs annihilate and due to the lepton-anti-lepton asymmetry, essentially all positron particles disappear while just enough electrons survive to be equal in number to the protons and thus to ensure charge neutrality (Steigman, 2007).

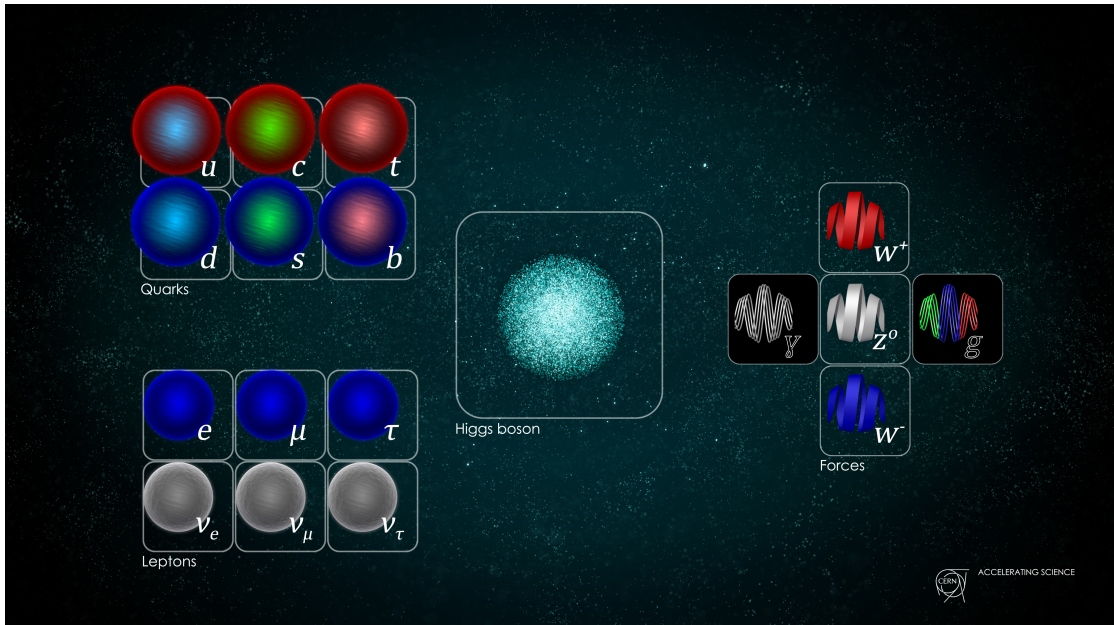


Figure 2.2 – The standard model of particle physics. The elementary particles are grouped into quarks and leptons. Each group consists of three pairs, for example the electron e and electron-neutrino ν_e are the first pair of the lepton group. The fundamental forces are the electromagnetic force, carried by photons γ , the weak force, carried by W and Z bosons, and the strong force, carried by gluons g . The fourth fundamental force, gravity, could not yet be united with the standard model. Finally, the Higgs boson is responsible for the mass of some fundamental particles. (Image credit: CERN / D. Dominguez)

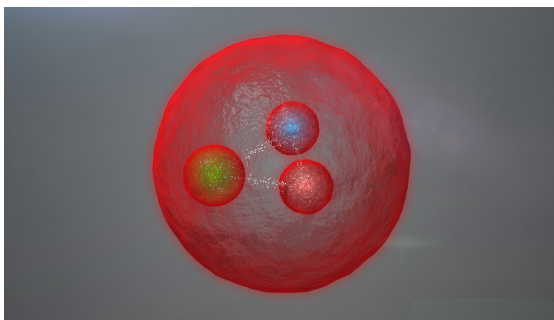


Figure 2.3 – Illustration of a baryon. Baryons are composite particles which are made out of three quarks. Common baryons are the proton and the neutron, which make up the nuclei of atoms. (Image credit: CERN / D. Dominguez)

Approximately 200 seconds after the birth of the universe, the temperature has dropped below 80 keV (≈ 930 million K) and nucleosynthesis begins (Steigman, 2007). Neutrons and protons are rapidly combined to form the hydrogen isotope nuclides deuterium (D) and tritium (^3H) and the helium isotope nuclides ^3He and ^4He (Steigman, 2007). There are no stable mass-5 nuclides and thus a new bottleneck appears at ^4He . As a result, nuclear reactions quickly

Chapter 2. Modern cosmology

incorporate all available neutrons into ${}^4\text{He}$ (Steigman, 2007). To overcome the mass-5 gap, Coulomb-suppressed reactions of ${}^4\text{He}$ with D, ${}^3\text{H}$, or ${}^3\text{He}$ are required and consequently the abundances of heavier nuclides than ${}^4\text{He}$ is strongly suppressed (Steigman, 2007). This results in a relatively low abundance of mainly lithium (${}^7\text{Li}$) and beryllium (${}^7\text{Be}$) nuclides, and the ${}^7\text{Be}$ nuclides decay into ${}^7\text{Li}$ by electron capture once the cosmos has further cooled (Steigman, 2007). Finally, there is another gap at mass-8 and thus no heavier nuclides are created during nucleosynthesis in astrophysically interesting abundances (Steigman, 2007). When the cosmos is about 20 minutes old and the temperature falls below 30 keV (≈ 350 million K), all nuclear reactions stop due to the Coulomb barriers and the absence of free neutrons (Steigman, 2007). The primordial nucleosynthesis is completed.

The universe continues to expand and the plasma consisting of nuclides, electrons, and photons continues to cool. The free electrons act as glue between the photons and the baryons through Thomson and Coulomb scattering and thus the plasma is a tightly coupled photon-baryon fluid (Hu and Dodelson, 2002). As a result, the baryonic matter cannot simply fall into the overdense regions originally created by inflation and acoustic oscillations occur (Hu and Dodelson, 2002). Once the universe has expanded further, its energy density starts to be dominated by matter and no longer by radiation and relativistic particles (Planck Collaboration et al., 2016). Cold Dark Matter (CDM), an unknown form of matter not yet included in the standard model of particle physics, becomes important for the amplitude of the acoustic peaks (Hu and Dodelson, 2002). Once the universe is about 380,000 years old and the temperature sinks to ≈ 3000 K, electrons and protons can combine to form neutral hydrogen. This event is called recombination (Hu and Dodelson, 2002). Now the universe becomes transparent and the photons can travel freely (Hu and Dodelson, 2002). We can see these photons today in the Cosmic Microwave Background (CMB), see figure 2.4. During their voyage to us, the photons are highly redshifted due to the expansion of the universe. The CMB is an almost perfect black-body radiation with temperature $T = 2.73$ K and tiny temperature fluctuations (Hu and Dodelson, 2002). These fluctuations correspond to the slightly over- or underdense regions in the early universe. After recombination, the baryons are no longer affected by the pressure of the photons and they can simply fall into the overdense regions with the CDM (Hu and Dodelson, 2002). These overdense regions evolve with time and give rise to the Large Scale Structure of the universe observed today, see figure 2.5. The composition of the energy density of the universe at recombination is shown in figure 2.6.

After recombination, the so-called “Cosmic Dark Ages” begin, the era before the first stars are born, see figure 2.1. The structure in the universe forms hierarchically through mergers of smaller halos with CDM and baryons into larger ones (Bromm and Yoshida, 2011). The first stars are Population III stars, which are chemically pristine and thus essentially metal-free (Stark, 2016). When we use the word metal in an astrophysical context, we mean every element heavier than helium. Population III stars are born in minihalos, which are the progenitors of the first galaxies (Bromm and Yoshida, 2011). These minihalos consist of CDM with total mass $\approx 10^6 M_{\odot}$ and if we could observe them today, they would have a redshift $z \approx 20 - 30$,

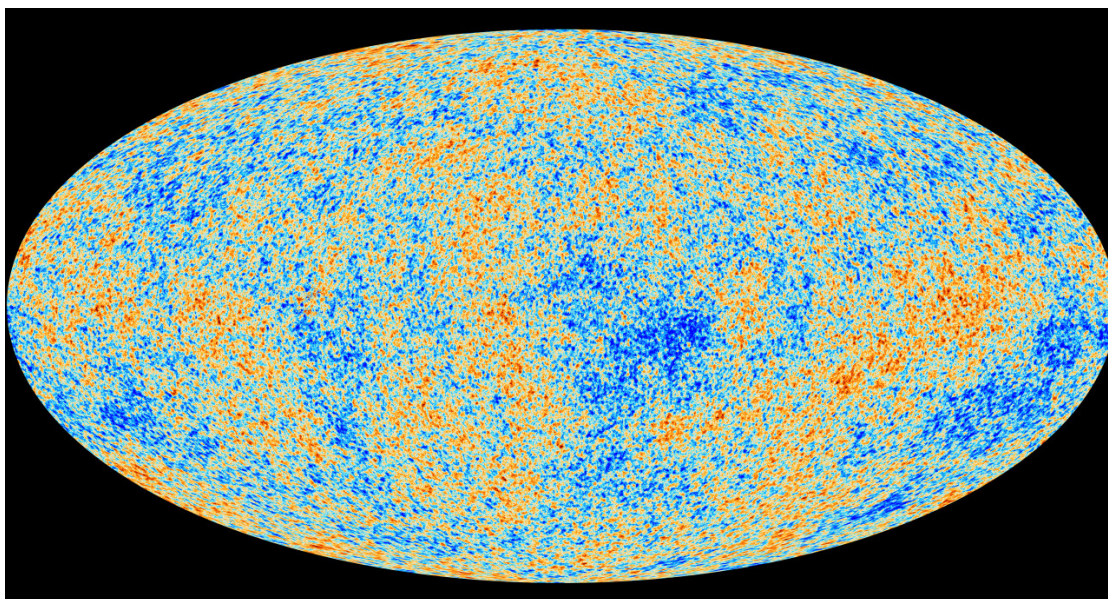


Figure 2.4 – The anisotropies of the Cosmic Microwave Background as observed by the Planck satellite. The different colors represent tiny temperature fluctuations. (Image credit: ESA / Planck Collaboration)

which corresponds to an age of the cosmos between 100 and 180 million years (Bromm and Yoshida, 2011). The nuclear reactions inside the Population III stars produce metals and thus these stars initiate the gradual metal enrichment of the universe (Stark, 2016). The first galaxies appear and they might already be metal-enriched by supernovae triggered by the first Population III stars (Bromm and Yoshida, 2011). In this case it is possible that they already hosted predominantly Population II instead of Population III stars (Bromm and Yoshida, 2011). Population II stars are metal-poor, but not metal-free objects. Over time, galactic scale outflows caused by radiation pressure from hot stars and supernovae enrich the Intergalactic Medium (IGM) with metals (Bromm and Yoshida, 2011). The first exemplars of a Supermassive Black Hole (SMBH) with mass $\geq 10^{10} M_{\odot}$ appear before the universe is one billion years old. They have likely grown from smaller seed black holes (Bromm and Yoshida, 2011). By the time the universe is 1 billion years old at redshift $z \approx 6$, radiation has reionized the neutral hydrogen in the IGM (Stark, 2016). The dominant sources for reionization were not the most luminous galaxies, but most likely a large amount of faint galaxies. Significant contributions from other sources like mini-quasars and massive Population III stars cannot yet be ruled out (Bromm and Yoshida, 2011; Stark, 2016). At last, clusters of galaxies begin to form. They correspond to large overdensities in the primordial density fluctuations and it takes several billion years for the cluster to reach equilibrium (Kravtsov and Borgani, 2012). The star-formation rate density in the universe peaks 3.5 billion years after the Big Bang, at $z \approx 2$. At this time, about $0.1 M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}$ of stellar mass are produced, ten times more than today (Madau and Dickinson, 2014). About 7 billion years after the Big Bang ($z \approx 0.73$), the expansion of the cosmos is no longer decelerating, but it starts to accelerate and this

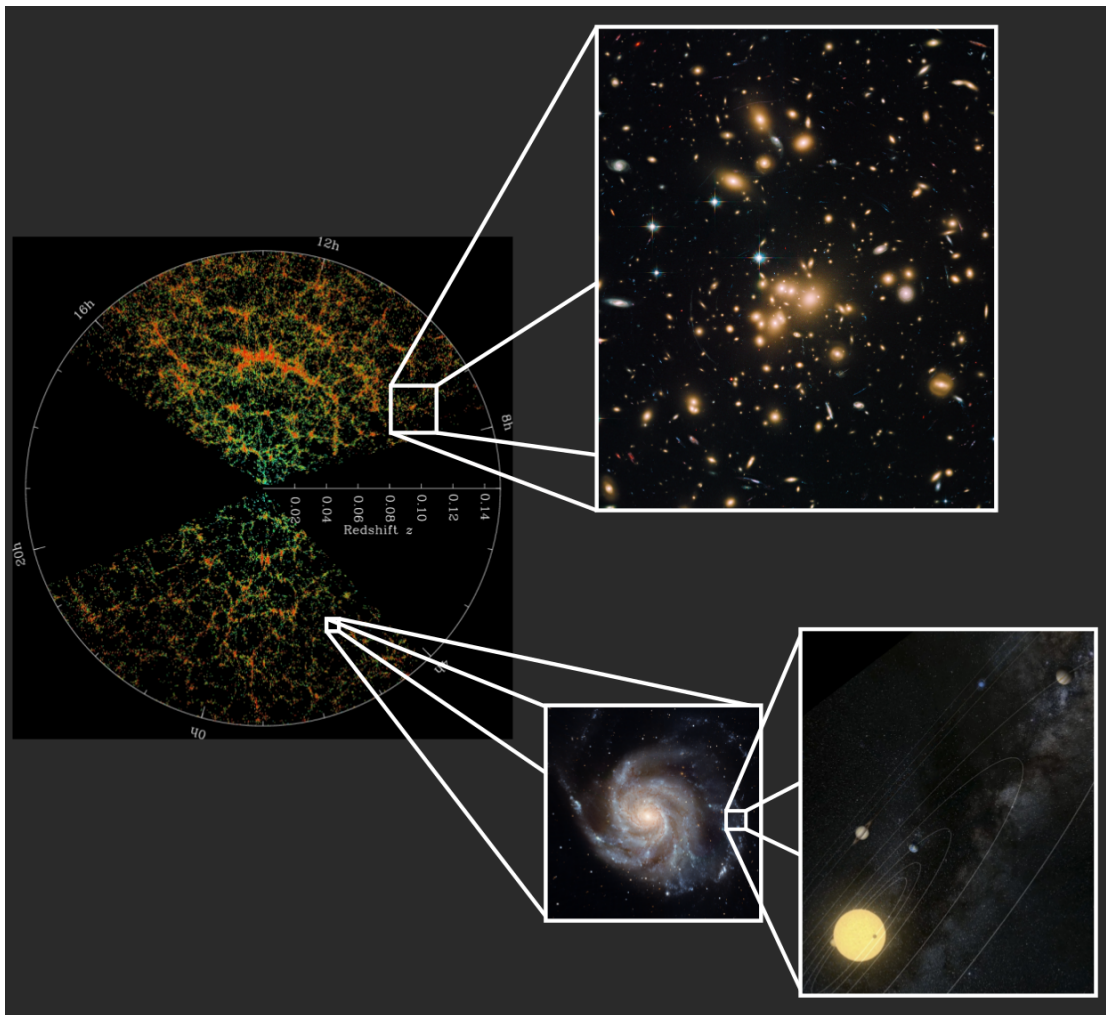


Figure 2.5 – A zoom-in on the Large Scale Structure of the universe. Each dot in the Sloan Digital Sky Survey map (leftmost image) represents a galaxy. The galaxies are the building blocks of the cosmic web, which consists of filaments of galaxies and voids, which are essentially empty. The nodes of the filaments are galaxy clusters, which consist of hundreds or even thousands of galaxies (image top right). The bottom center image shows a spiral galaxy in a filament, where the galaxy density is much lower than in a cluster. Spiral galaxies typically consist of billions of stars. The bottom right picture shows a zoom-in on one of the galaxies' stars which hosts a planetary system. (Image credit: Own work / SDSS / NASA / ESA / STScI, see photographic credits)

acceleration is still going on today (Riess et al., 1998; Perlmutter et al., 1999). This effect is caused by an unknown energy source called Dark Energy (DE). DE can be described by a constant term in Einstein's General Theory of Relativity (GR), but its physical nature is still a mystery. When the cosmos is approximately 10 billion years old at $z \approx 0.4$, DE begins to dominate the energy density of the universe and its importance increases to this day, see figure 2.6 (Perlmutter et al., 1999). At about the same time, our sun and the solar system

2.2. Are we there yet? - The road to the standard model

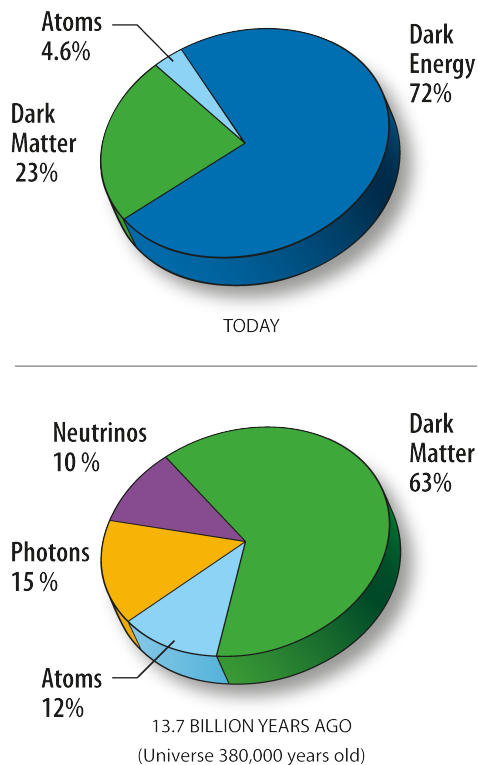


Figure 2.6 – The energy density content of the universe changes with time. The pie chart from WMAP 5-year data shows the energy density content at recombination, when the cosmos was about 380,000 years old, and today. (Image credit: NASA / WMAP Science Team)

begin to form in our home galaxy, the Milky Way, which has been assembled over time through numerous merger and accretion events (Bromm and Yoshida, 2011). In the following 4 billion years, the sun and earth evolve to their present form and earth begins to harbor life.

2.2 Are we there yet? - The road to the standard model

The road that led to today's standard model of cosmology has been a bumpy one. A GPS navigation system for science has yet to be invented, and so the only option scientists have when they face a scientific cross-road is to choose a direction and explore the territory it leads to. While this may occasionally lead to dead-ends, more often than not it will result in truly exciting discoveries. The history of cosmology is no exception to this rule. In this section, we will present this history and use it to introduce the mathematical formalism of cosmology.

Around the year 1900, the scientific view of the universe was drastically different from today. For most scientists, the cosmos exists essentially out of our home galaxy, the Milky Way, and the density of stars drastically decreases beyond its bounds (O'Raifeartaigh et al., 2017; Kragh, 2014). There have been observations of faint "nebulae", but many astronomers think that they are solar systems in early stages of evolution inside our Milky Way instead of extragalactic objects (O'Raifeartaigh, 2013). However, already in the 18th century Thomas Wright and Immanuel Kant had also raised another possibility, namely that these nebulae

are very distant galaxies similar to our own Milky Way (O’Raifeartaigh et al., 2017). This theory was supported in the 19th century by observations of spiral structure in some nebulae and the fact they appeared to contain stars. But there was also evidence which seemed to support the contrary point of view. First, it was found that these nebulae were clustered near the poles of the Milky Way and second, an extremely bright nova was observed in the Andromeda nebula (O’Raifeartaigh et al., 2017). In the early 20th century, Slipher (1913) uses a spectrograph to obtain the first radial velocity measurement of a nebula, namely of Andromeda. He measures the color shift of the nebula’s spectrum due to its motion, the so-called Doppler effect. If an object approaches us, its spectrum will be shifted towards the blue, and if it recedes from us, the spectrum will be shifted towards the red. With this method Slipher finds that the Andromeda nebula is approaching us with a velocity of 300 km/s, the greatest velocity observed to this date and much larger than stellar velocities (Slipher, 1913). Over the next years, he expands his work to include 25 nebulae and he finds that they all have velocities of similar magnitude (Slipher, 1917). In addition, Slipher (1917) observes that there are more nebulae moving away from us than nebulae approaching us. The measured velocities are so large that some astronomers already speculate that the nebulae cannot be gravitationally bound by the Milky Way and must reside outside of its bounds. This debate was only settled a decade later, when Edwin Hubble succeeded in measuring the distances to the nebulae (O’Raifeartaigh, 2013).

Around the same time, Albert Einstein presented the final version of his General Theory of Relativity (GR), which describes the gravitational force (Einstein, 1916). He explains gravity by a curvature of the four-dimensional spacetime, see figure 2.7. Larger masses cause a stronger deformation of spacetime than smaller ones and thus their gravitational effect is stronger. This explains for example the motion of the moon around the earth in a beautiful geometric manner: The moon, which corresponds to the small red ball in figure 2.7, is attracted by the strong curvature of the earth, which corresponds to the medium-sized orange ball. If the moon was at rest, it would be accelerated towards the earth by the curvature in spacetime, and finally crash into it. Luckily, the moon moves with a certain velocity v_{moon} , which is just fast enough to counteract the gravitational pull. As a result, the moon moves in a circle with constant radius around the earth. In addition, there is an even larger mass than the earth, namely the sun, which is represented by the large yellow ball in the figure. The sun causes an even stronger deformation of spacetime than the earth and consequently both the earth and the moon circle around it.

Now we want to quantitatively describe this effect. For this purpose, we introduce the contravariant coordinate x^μ to describe a location in spacetime. The value at $\mu = 0$ is the coordinate on the timelike axis, which is a time t multiplied with the speed of light c , and the values for $\mu = 1, 2, 3$ are the coordinates on the spacelike axes. We also introduce the covariant metric tensor $g_{\mu\nu}$, which must be symmetric (Einstein, 1916),

$$g_{\mu\nu} = g_{\nu\mu}. \tag{2.1}$$

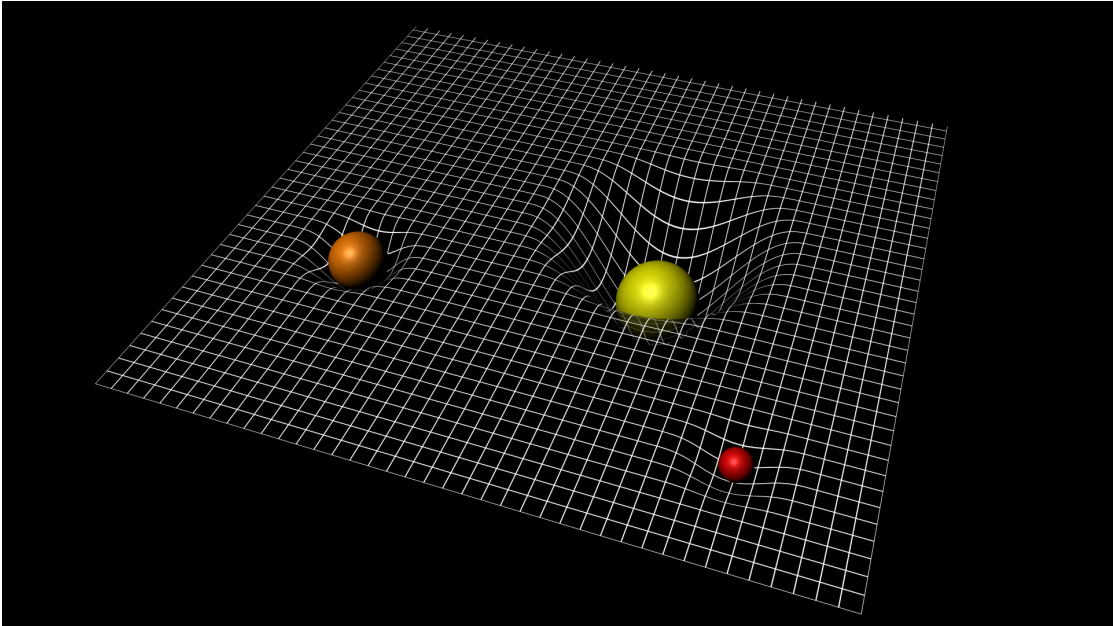


Figure 2.7 – Einstein’s GR explains gravity by the curvature of spacetime: A large mass (yellow ball) causes a larger curvature than smaller ones (orange and red balls) and thus the gravitational pull towards this object is the strongest. (Image credit: ESA / C. Carreau)

The metric tensor $g_{\mu\nu}$ contains geometric information about spacetime, for example about its curvature, and thus it describes the gravitational field (Einstein, 1916). The values of $g_{\mu\nu}$ are a function of the spacetime location x^μ . We can also use the metric tensor to compute the square of an infinitesimal line element ds (Einstein, 1916),

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \quad (2.2)$$

where we introduced the Einstein summation convention,

$$a_i b^i = \sum_i a_i b^i. \quad (2.3)$$

This quantity is important, because ds^2 is invariant under coordinate system transformations (Einstein, 1916). Its value determines an important property: When $ds^2 > 0$, then the line element is spacelike, and when $ds^2 < 0$, it is timelike (Einstein, 1916). For example, if a massive particle at coordinate x^μ is moving at half the speed of light c along one spacelike coordinate in the absence of gravity, we have $ds^2 < 0$ and it is timelike. For an external observer, the time $dt = dx^0/c$ has passed, but for the particle the *eigentime* $d\tau = \sqrt{-ds^2/c^2}$ has passed and the two times are not the same (see e.g. Einstein, 1916, for the relation between *eigentime* and ds^2). For photons, we have $ds^2 = 0$ (Einstein, 1916). Note that we use a different sign convention than Einstein (1916) for the metric tensor. As a result we have in the absence of

gravity:

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2.4)$$

which is exactly the metric tensor for the Special Theory of Relativity (Einstein, 1916). In general, $g_{\mu\nu}$ will not have the simple form shown in equation 2.4, and it has to be computed by using the Einstein field equations (Einstein, 1916, shown in the form used in e.g. Friedmann (1922)),

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \quad (2.5)$$

where $R_{\mu\nu}$ is the Ricci curvature tensor, R is the Ricci scalar, G is the gravitational constant, and $T_{\mu\nu}$ is the stress-energy-momentum tensor. A detailed mathematical discussion or even derivation of these equations would require the introduction of the mathematical theory of differential geometry and it would take a substantial amount of space. Therefore we refer the reader to the original 54-page paper by Einstein (1916) or one of the several excellent books written on this topic for further information about their mathematical structure and we present instead a qualitative discussion. The left hand side of equation 2.5 describes the curvature of spacetime. Note that both $R_{\mu\nu}$ and R depend on the metric $g_{\mu\nu}$. The curvature described on the left hand side is determined by the terms on the right hand side of the equation, namely the stress-energy-momentum tensor $T_{\mu\nu}$. As the name implies, it encapsulates information about the energy density, momentum density, pressure, and shear stress at spacetime location x^μ . These deform the spacetime and thus cause gravity. Once they are determined for a given x^μ , we can compute $g_{\mu\nu}$.

Let us now assume that we have an empty space with a central mass M and that we have computed $g_{\mu\nu}$ for every coordinate x^μ , so that we know exactly how the spacetime curvature looks like. Next, we would naturally like to know how a particle moves in this gravitational field. We make the simplifying assumptions that the only non-zero force is gravity and that this test particle is mass-less and does not cause any gravitation. The reason for this is as follows: If the particle had a non-negligible mass, then spacetime itself would change when the position of the particle changes, and thus the complexity of this example would drastically increase. Now we can use the Principle of Stationary Action: The path along which the particle will travel, the so-called geodesic, must be a path which satisfies (Einstein, 1916):

$$\delta \left(\int_{P_1}^{P_2} ds \right) = 0, \quad (2.6)$$

where P_1 and P_2 are the start and end points of the particle in spacetime. Thus the invariant line element ds plays again a crucial role: The particle will travel along a path which lets the

2.2. Are we there yet? - The road to the standard model

variation of the integral of ds vanish.

Any theory must pass the test of reality to be accepted by the scientific community, and GR is no exception to this rule. For example, it correctly explains the observed perihelion shift of the planet Mercury, which is not possible with Newton's theory of gravity (Einstein, 1916). However, the probably most spectacular confirmation of GR comes only a few years later in 1919, when Dyson, Eddington, and Davidson use a solar eclipse to measure the gravitational lensing of star light by the sun and find that the deflection is approximately as predicted by GR and thus two times larger than predicted by the Newtonian theory (Dyson et al., 1920). This makes Einstein an international celebrity.

Already in the year 1917, Einstein applies his theory to the universe as a whole, which was at that time for many scientists still only our Milky Way (Einstein, 1917). However, he does not attempt to find general cosmological solutions of the GR equations, but instead he develops a particular model of the universe. He is guided by the somewhat philosophical assumption that the universe is static and eternal. However, this assumption seems very reasonable at this time, and no astronomical observations to the contrary are known to Einstein (O'Rai feartaigh et al., 2017). In addition, he assumes that matter is uniformly distributed over the largest scales of the universe. In other words, he implicitly assumes that the universe is homogeneous and isotropic on the largest scales, an assumption which is later much more explicitly formulated (see e.g. Milne, 1933, but note that he develops a cosmology model which is not based on GR in this paper) and called the "cosmological principle". Even though the cosmological principle was at odds with the astronomical data around 1917, today's measurements show that it actually holds true and it is one of the pillars of modern cosmology (O'Rai feartaigh et al., 2017). As a result, Einstein obtains a spatially closed universe, but the GR equations do not allow it to be static (Einstein, 1917). He realizes that it is possible to extend the field equations of GR by adding the term $\Lambda g_{\mu\nu}$ and that this leads to a static solution. The Einstein field equations are now:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} - \Lambda g_{\mu\nu}. \quad (2.7)$$

However, the static matter distribution in the universe model comes at the price of introducing an unknown "cosmological constant" Λ . Einstein himself seems to be not fully at ease with this solution, as he describes the constant in Einstein (1917) as "*eine neue, durch unser tatsächliches Wissen von der Gravitation nicht gerechtfertigte Erweiterung*" (English translation: "*a new extension which is not justified by our actual knowledge of gravitation*").

Shortly thereafter, the astronomer Willem de Sitter proposes his cosmological model. In de Sitter (1917), he develops a universe which he claims is static and spatially closed, but it is empty of matter. It could be interpreted as an approximation to a very low density universe (Kragh, 2014). The model predicts an interesting effect, namely that matter in this universe would recede from an observer and be redshifted (de Sitter, 1917). However, it is later

shown that the de Sitter universe is actually not static, but evolving with time ([Lemaître, 1926](#)).

A major step forward is taken when [Friedmann \(1922, 1924\)](#) analyzes the GR equations and finds the cosmological solutions for a homogeneous and isotropic universe, including the Einstein and de Sitter universe models. Many of these solutions describe dynamic universes which are evolving with time. This does not comply with the prevailing opinion at the time that the cosmos must be static, and Einstein believes that Friedmann's solutions are simply a mathematical exercise without any physical relevance ([Kragh, 2014](#)). As a result, the work is little noticed in the scientific community ([Kragh, 2014](#)). However, the Friedmann solutions are later rediscovered and [Robertson \(1935\)](#) and [Walker \(1935\)](#) give a mathematical proof that the used metric tensor is the only metric tensor which exists on a homogeneous and isotropic spacetime. Independently of Friedmann, [Lemaître \(1927\)](#) discovers cosmological solutions of GR which lead to an expanding universe model ([Kragh and Smith, 2003](#)). In their honor, the metric tensor which we use in cosmology to this day is called the Friedmann-Lemaître-Robertson-Walker metric (FLRW metric).

We now derive the FLRW metric and the Friedmann equations, which determine the dynamics of spacetime. The derivation follows [Bartelmann and Schneider \(2001\)](#), but note that they use a different sign convention for the metric tensor $g_{\mu\nu}$ and a different definition of x^0 than we did when we discussed GR in this chapter. Since these conventions do not change the physics and are merely a matter of taste, we will follow [Bartelmann and Schneider \(2001\)](#) and change the sign of the metric tensor, so that the $g_{\mu\nu}$ shown in equation 2.4 would pick up a factor of (-1) , and define $x^0 = t$ instead of $x^0 = ct$. The speed of light factor will be absorbed by the metric tensor, which we will from now on simply call "metric".

We assume that the cosmological principle holds, i.e. that hypersurfaces of constant time are homogeneous and isotropic on the largest scales. They can be flat or curved. We formulate the cosmological principle using two postulates ([Bartelmann and Schneider, 2001](#)):

1. When averaged over sufficiently large scales, there exists a mean motion of radiation and matter in the universe with respect to which all averaged observable properties are isotropic.
2. All fundamental observers, i.e. imagined observers which follow this mean motion, experience the same history of the universe, i.e. the same averaged observable properties, provided they set their clocks suitably. Such a universe is called observer-homogeneous.

These two postulates constrain the allowed form of the metric substantially. In the following, Roman indices like i refer to the spatial indices 1,2,3, while Greek indices like μ refer to all four indices. We call spatial coordinates which are constant for fundamental observers

2.2. Are we there yet? - The road to the standard model

comoving coordinates. In these coordinates, $dx^i = 0$ and thus

$$ds^2 = g_{00} dt^2. \quad (2.8)$$

We require that the *eigentime* of fundamental observers equals the cosmic time, and thus we have (Bartelmann and Schneider, 2001)

$$g_{00} = c^2. \quad (2.9)$$

Due to the isotropy of the cosmos, clocks can be synchronized such that the space-time components of $g_{\mu\nu}$ vanish (Bartelmann and Schneider, 2001),

$$g_{0i} = 0. \quad (2.10)$$

Consequently, we have for the metric

$$ds^2 = c^2 dt^2 + g_{ij} dx^i dx^j, \quad (2.11)$$

where g_{ij} is the metric of the spatial hypersurfaces (Bartelmann and Schneider, 2001). As a consequence of isotropy, the spatial metric can only isotropically contract or expand with a scale function $a(t)$. The scale function is a function of time only, because otherwise the expansion would differ at different locations, and this would violate the assumption of homogeneity (Bartelmann and Schneider, 2001). Thus we have

$$ds^2 = c^2 dt^2 - a^2(t) dl^2, \quad (2.12)$$

where dl is the line element of the homogeneous and isotropic three-space. Due to homogeneity, all quantities describing matter and radiation in the universe, e.g. the density ρ and the pressure p , can only be functions of time and we can set an arbitrary point as origin of the coordinate system (Bartelmann and Schneider, 2001). We introduce the angles θ and ϕ to define the locations on the unit sphere around the origin and we use a radial coordinate w . As a result, we obtain the squared line element for the FLRW metric (Bartelmann and Schneider, 2001),

$$ds^2 = c^2 dt^2 - a^2(t) [dw^2 + f_k^2(w) (d\phi^2 + \sin^2(\theta) d\theta^2)] \quad (2.13)$$

$$= c^2 dt^2 - a^2(t) [dw^2 + f_k^2(w) d\omega^2], \quad (2.14)$$

where

$$f_k(w) = \begin{cases} k^{-\frac{1}{2}} \sin\left(k^{\frac{1}{2}} w\right) & \text{if } k > 0 \\ w & \text{if } k = 0 \\ (-k)^{-\frac{1}{2}} \sinh\left((-k)^{\frac{1}{2}} w\right) & \text{if } k < 0. \end{cases} \quad (2.15)$$

Due to homogeneity, $f_k(w)$ must be either a trigonometric, a linear, or a hyperbolic function

of w , depending on the curvature k , which can be positive, zero, or negative, and this leads to the presented form of $f_k(w)$ (Bartelmann and Schneider, 2001). A positive curvature corresponds to a spatially closed universe, zero curvature to a flat Euclidean universe, and negative curvature to a hyperbolic universe, see figure 2.8.

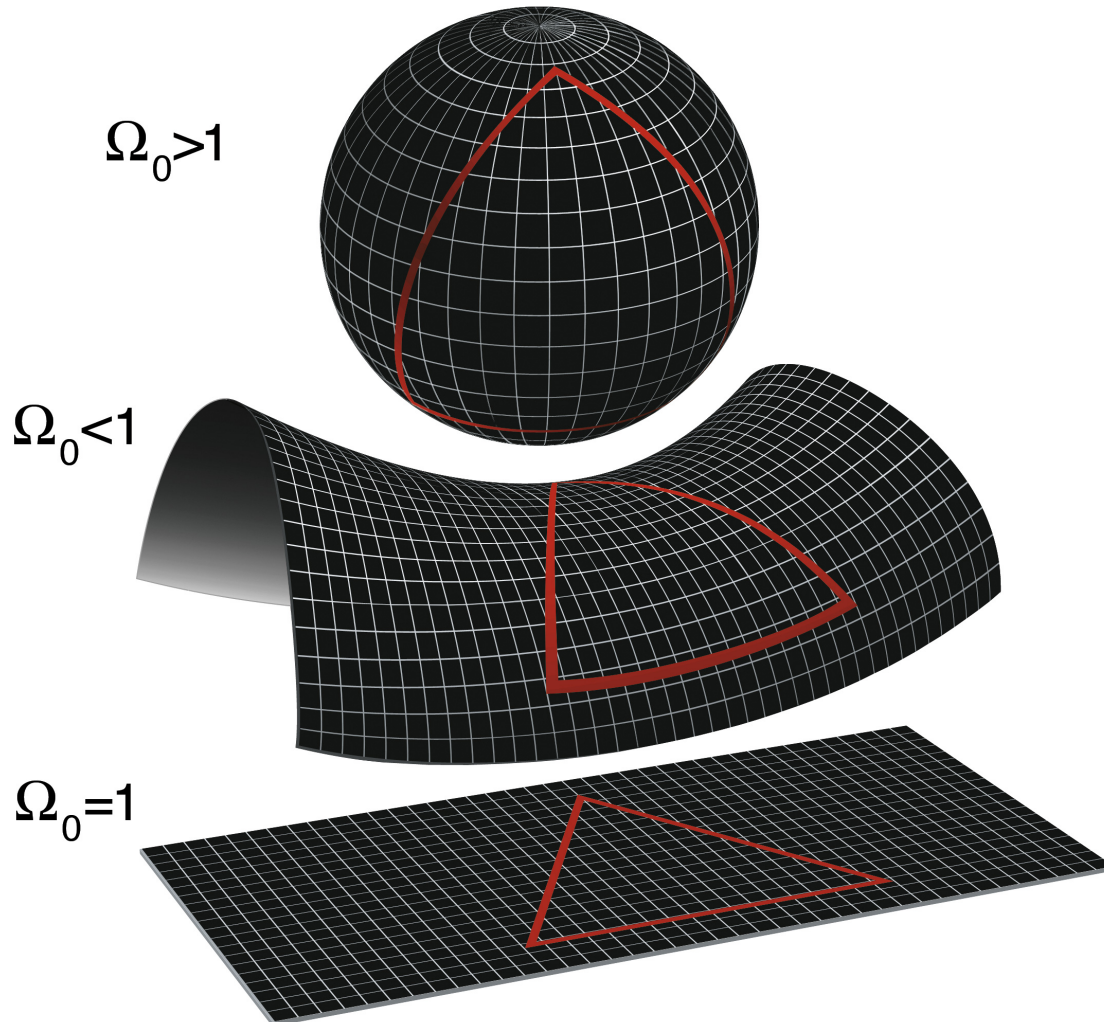


Figure 2.8 – The FLRW metric permits three different spatial geometries. If the total density parameter Ω_{tot} (denoted with Ω_0 in this image) is larger than unity, the space is closed and not Euclidean (top). The sum of the angles of a triangle is larger than 180 degrees. If it equals one, the space is flat and Euclidean (bottom). The sum of the triangle angles is exactly 180 degrees. If $\Omega_{\text{tot}} < 1$, the space is hyperbolic, not Euclidean, and the angle sum is < 180 degrees (center). (Image credit: NASA / WMAP Science Team)

In the next step, we want to derive the dynamics of spacetime, so we need to investigate the dependence of the scale factor a on time and the dependence of the curvature k on the matter and radiation in spacetime. For this purpose, we use the Einstein field equations in formula 2.7, but note that the sign of the Λ term is inverted and the $T_{\mu\nu}$ term picks up an extra factor c^2 when using the Bartelmann and Schneider (2001) conventions. The

2.2. Are we there yet? - The road to the standard model

highly symmetric FLRW metric in equation 2.14 implies that $T_{\mu\nu}$ must take the form of a stress-energy-momentum tensor of a homogeneous perfect fluid characterized by its density $\rho(t)$ and pressure $p(t)$ (Bartelmann and Schneider, 2001). Note we have density and pressure due to non-relativistic matter, ρ_m and p_m , and density and pressure due to relativistic matter and radiation, ρ_R and p_R . The total density and pressure are

$$\rho = \rho_m + \rho_R, \quad (2.16)$$

$$p = p_m + p_R. \quad (2.17)$$

Due to the homogeneity, both density and pressure can only depend on time (Bartelmann and Schneider, 2001). As a result, the Einstein field equations dramatically simplify and they result in the two independent Friedmann equations (Bartelmann and Schneider, 2001):

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3}, \quad (2.18)$$

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}, \quad (2.19)$$

where \dot{a} and \ddot{a} are the first and second derivatives of the scale factor with respect to time. We define $a = 1$ at the present epoch t_0 and now the scale factor $a(t)$ is fully determined (Bartelmann and Schneider, 2001). Finally, we combine equations 2.18 and 2.19 to obtain the adiabatic equation (Bartelmann and Schneider, 2001):

$$\frac{d}{dt} \left[a^3(t)\rho(t)c^2 \right] + p(t) \frac{da^3(t)}{dt} = 0. \quad (2.20)$$

It has an intuitive interpretation: The first term $a^3(t)\rho(t)$ is proportional to the energy in a fixed comoving volume and the change in this “internal” energy equals the pressure times the change in proper volume. Thus equation 2.20 is the first law of thermodynamics in a cosmological context (Bartelmann and Schneider, 2001).

In a third step, we would like to reformulate the first Friedmann equation in terms of the density parameters of the universe and use it to study the spatial curvature of the cosmos. We define the Hubble¹ parameter H as the relative expansion rate of the universe,

$$H = \frac{\dot{a}}{a}, \quad (2.21)$$

and its value at the present epoch t_0 is the Hubble constant H_0 (Bartelmann and Schneider, 2001),

$$H_0 = H(t_0). \quad (2.22)$$

The value of the Hubble constant has been very uncertain until recently, so it is often expressed

¹The reason for naming this parameter after Edwin Hubble will become clear later in this chapter

Chapter 2. Modern cosmology

as

$$H_0 = 100 h \frac{\text{km}}{\text{sMpc}}. \quad (2.23)$$

Current observations show that it is approximately (e.g., [Riess et al., 2016](#); [Bonvin et al., 2017](#))

$$H_0 \approx 70 \frac{\text{km}}{\text{sMpc}}. \quad (2.24)$$

The timescale for the expansion of the universe is roughly the inverse of H_0 ([Bartelmann and Schneider, 2001](#)). Following [Bartelmann and Schneider \(2001\)](#), we define the critical density of the universe,

$$\rho_{\text{crit}} = \frac{3H_0^2}{8\pi G} \approx 1.9 \times 10^{-29} h^2 \frac{\text{g}}{\text{cm}^3}, \quad (2.25)$$

and further the density parameters at the current epoch for matter, Λ , and radiation and relativistic particles, which are respectively denoted with Ω_m , Ω_Λ , and Ω_R :

$$\Omega_m = \frac{\rho_m(t_0)}{\rho_{\text{crit}}}, \quad (2.26)$$

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2}, \quad (2.27)$$

$$\Omega_R = \frac{\rho_R(t_0)}{\rho_{\text{crit}}}. \quad (2.28)$$

The total density parameter at the current epoch Ω_{tot} is given by

$$\Omega_{\text{tot}} = \Omega_m + \Omega_\Lambda + \Omega_R \quad (2.29)$$

and the spatial curvature density parameter Ω_k by

$$\Omega_k = -\frac{kc^2}{H_0^2}. \quad (2.30)$$

The equation of state $p = p(\rho)$ relates the pressure to the energy density. For non-relativistic matter, we have $p \ll \rho c^2$ and for radiation and relativistic particles we have $p = \rho c^2/3$ ([Bartelmann and Schneider, 2001](#)). We insert these formulas into the adiabatic equation 2.20 and we find

$$\rho_m(t) = \frac{\rho_m(t_0)}{a^3(t)}, \quad (2.31)$$

$$\rho_R(t) = \frac{\rho_R(t_0)}{a^4(t)}, \quad (2.32)$$

so the energy density of relativistic matter drops quicker than the energy density of non-

2.2. Are we there yet? - The road to the standard model

relativistic particles (Bartelmann and Schneider, 2001). Now we can reformulate the first Friedmann equation 2.18 as

$$\frac{H^2(t)}{H_0^2} = \frac{\Omega_R}{a^4(t)} + \frac{\Omega_m}{a^3(t)} + \frac{\Omega_k}{a^2(t)} + \Omega_\Lambda. \quad (2.33)$$

Let us use this equation to draw a first conclusion about the universe. For the current epoch, the equation becomes

$$1 = \Omega_{\text{tot}} + \Omega_k \quad (2.34)$$

and we see that the spatial geometry of the cosmos is determined by its total density parameter. By combining equations 2.30 and 2.34, we find the following three cases, see figure 2.8:

$$\begin{aligned} \Omega_{\text{tot}} < 1 &\rightarrow k < 0 \rightarrow \text{Hyperbolic geometry} \\ \Omega_{\text{tot}} = 1 &\rightarrow k = 0 \rightarrow \text{Flat Euclidean geometry} \\ \Omega_{\text{tot}} > 1 &\rightarrow k > 0 \rightarrow \text{Closed geometry.} \end{aligned} \quad (2.35)$$

In a nutshell, the Friedmann-Lemaître-Robertson-Walker universe model is characterized by the reformulated first Friedmann equation 2.33 and four parameters: The present expansion rate H_0 and the density parameters Ω_m , Ω_R , and Ω_Λ (Bartelmann and Schneider, 2001). So the “cosmological problem” is now to determine these quantities - but it is not an easy task to weigh and measure the cosmos, in particular since many scientists in the early 20th century still do not even know that it is larger than our Milky Way!

This debate is settled by Hubble (1925, 1926) who uses Cepheid variable stars in the nebulae to determine their distance from us. The Cepheids have a strong relationship between their luminosity and their pulsation periods and thus they are excellent standard candles for cosmic distance measurements (Leavitt, 1908; Leavitt and Pickering, 1912; Hertzsprung, 1913; Shapley, 1918). It is now firmly established that the faint nebulae are galaxies similar to our Milky Way and that the universe is much larger than previously thought!

Three years later, Hubble (1929) combines his distance measurements with Slipher’s velocity data. This leads to Hubble’s linear distance-velocity relation: The further away from us a galaxy is, the faster it moves away from us. Consequently the universe is not static, but expanding! However, this assumes that the measured redshifts can really be interpreted as Doppler shifts due to velocity. At the time, this is indeed not undisputed. Shortly after Hubble’s discovery, Zwicky (1929) and Belopolsky (1929) are the first to propose a different interpretation: The redshifts are not caused by velocities, but they are due to “tired light”. The term refers to a hypothetical energy loss mechanism of photons which would result in a redshift as they travel large distances in the universe (Kragh, 2017). This idea fell slowly out of favor over the following decades due to its ad hoc nature and the lack of supporting observational data, but it was a serious alternative interpretation at the time (Kragh, 2017).

Chapter 2. Modern cosmology

Despite these alternatives, Hubble's distance-velocity relation leads over the following years to the abandoning of static models of the universe and to the rediscovery of Friedmann's and Lemaître's work (Kragh, 2014). The slope of the linear distance-velocity relation is the constant H_0 , which is named in Hubble's honor. According to the data presented in Hubble (1929), it is approximately $H_0 \approx 500 \text{ km}/(\text{s Mpc})$, a value which is an order of magnitude too high by today's measurements. In his 1929 paper, Hubble does unfortunately not cite Slipher, even though he uses his velocity data. Thus Slipher's name is not included in the name of the constant (O'RaiFeartaigh, 2013). Interestingly, Hubble himself does not claim that the universe is actually expanding for the remainder of his life - he is satisfied with having established an empirical law between distance and apparent velocity and he leaves the interpretation to others (Kragh, 2014).

The priest and astronomer Lemaître on the other hand firmly believes in the expansion of the universe. Two years before Hubble publishes the distance-redshift relation, Lemaître (1927) derives the cosmological GR equations for an expanding universe without knowing about the Friedmann papers (Kragh, 2014). He knows about Slipher's redshift measurements and he already has a few distance measurements from Hubble at hand (O'RaiFeartaigh et al., 2017). He uses these to compute a first estimate for the value of H_0 for his theory of the expanding universe - two years before Hubble publishes his famous relation (Lemaître, 1927; Kragh and Smith, 2003). However, he publishes his results in French and in a little known journal. In addition, Einstein is opposed to dynamic universes. As a result, his work is little noticed until interest in expanding universe theories surges after Hubble's 1929 paper. The new observational evidence leads to a discussion between Eddington and Lemaître and thus to the rediscovery of his work (O'RaiFeartaigh, 2013).

The distance-redshift relation arises naturally in expanding universes with FLRW metric and we will briefly discuss it here. Due to the expansion of the cosmos, photons are redshifted while they propagate from the emitting galaxy to the observer. Following Bartelmann and Schneider (2001), we consider a comoving source which emits light at time t_{emit} . The photons reach the comoving observer at $w = 0$ at time t_{receive} . For photons, $ds = 0$ and using the FLRW metric in equation 2.14 for a backwards-directed radial light ray, we have (Bartelmann and Schneider, 2001)

$$|c dt| = a dw. \quad (2.36)$$

The comoving separation between source and observer $w_{\text{source,obs}}$ is by definition constant,

$$w_{\text{source,obs}} = \int_{\text{obs}}^{\text{source}} dw = \int_{t_{\text{emit}}}^{t_{\text{receive}}(t_{\text{emit}})} \frac{c dt}{a} = \text{constant}, \quad (2.37)$$

and thus

$$\frac{dw_{\text{source,obs}}}{dt_{\text{emit}}} = 0. \quad (2.38)$$

2.2. Are we there yet? - The road to the standard model

As a result, we have (Bartelmann and Schneider, 2001)

$$\frac{dt_{\text{receive}}}{dt_{\text{emit}}} = \frac{a(t_{\text{receive}})}{a(t_{\text{emit}})} \quad (2.39)$$

and we can identify $(dt)^{-1}$ with the light frequency ν to obtain

$$\frac{dt_{\text{receive}}}{dt_{\text{emit}}} = \frac{\nu_{\text{emit}}}{\nu_{\text{receive}}} = \frac{\lambda_{\text{receive}}}{\lambda_{\text{emit}}}. \quad (2.40)$$

We define the redshift z as the relative change in wavelength (Bartelmann and Schneider, 2001),

$$1 + z = \frac{\lambda_{\text{receive}}}{\lambda_{\text{emit}}} \quad (2.41)$$

and finally we obtain the relation between redshift and the expansion of the universe,

$$1 + z = \frac{a(t_{\text{receive}})}{a(t_{\text{emit}})}. \quad (2.42)$$

Let us now assume that an observer simultaneously receives light from a source at a small physical distance and from a source at a large physical distance. The light from the distant source was consequently emitted at a much earlier time t_{emit} than the light emitted from the close-by source. As a result, the scale factor a will be much smaller for the distant source than for the close-by source and the redshift z will be much greater.

In the year 1931, Einstein finally adopts the point of view that the universe is expanding and he no longer assumes that it has to be static (O’Raifeartaigh, 2013). Consequently, he removes the Λ term from the GR field equations, as it was only introduced in Einstein (1917) to allow a static model of the cosmos. Allegedly, he later called the introduction of Λ “my biggest blunder” (O’Raifeartaigh et al., 2017). Together with de Sitter, he publishes a simple model of an expanding universe in Einstein and de Sitter (1932), but it is scarcely noticed by the scientific community in the 1930s and neither author takes it too seriously (Kragh, 2014). However, it is nevertheless the first well-known model of a flat and infinite universe (Kragh, 2014; O’Raifeartaigh et al., 2015). Like other models of expanding universes, the Einstein-de Sitter model has necessarily a singularity at $t = 0$, but this is not even mentioned in the paper. For Einstein and many of his colleagues, the singularity at $t = 0$ is merely a mathematical feature and not something with a physical representation (Kragh, 2014). Lemaître, however, thinks that this singularity must have a physical meaning. In Lemaître (1931), he proposes that the universe consisted of a “primeval atom” at $t = 0$ and he develops a front-runner for today’s Big Bang Theory (BBT). However, there is no compelling physical evidence for a Big Bang and the theory of the expanding universe will rest on very shaky grounds until the discovery of the CMB in 1965.

In the meantime, astrophysics and nuclear physics are flourishing and they will have an

Chapter 2. Modern cosmology

enormous effect on cosmological research in the next decades. In a seemingly unrelated paper, the astronomer Fritz Zwicky investigates the gravitational potential of the Coma cluster and he finds that its luminous mass is much too low to account for it (Zwicky, 1933). Thus Zwicky has found the first evidence for the existence of large amounts of non-luminous Dark Matter (DM) in the 1930s. In addition, the old astrophysical question of energy generation in the sun is finally solved. The advance of nuclear physics allows Weizsäcker (1937, 1938) and Bethe (1939) to find the proton-proton and the CNO cycles by which the sun converts hydrogen into helium. These fusion reactions do not only explain why the sun shines and makes earth habitable, but they also show how light elements are formed in the universe (Kragh, 1996). However, this mechanism cannot explain the existence of heavier elements (Kragh, 1996). Already in the year 1938, Weizsäcker suggests to use the cosmic element abundances to investigate the earlier state of the universe in which they must have formed. He speculates that they must have been produced in a great matter aggregation under extreme conditions (Weizsäcker, 1938). At roughly the same time, the first relative cosmic abundances of isotopes based on terrestrial, meteoritic, and astrochemical measurements is published by Goldschmidt (Kragh, 1996).

These developments are noticed by Chandrasekhar and Henrich (1942), who compute today's element abundances by assuming that the nuclear species in the early universe were in dynamic equilibrium. However, their results only match the data for light elements, not for heavy ones. They suggest that the different elements might be formed at different epochs of the expanding universe and thus in different physical conditions (Kragh, 1996). The next step is taken by Gamow (1946), who combines the GR description of the expanding universe with nuclear physics and develops the cold BBT (Kragh, 1996). It describes an expanding universe filled with a cold, thick soup of neutrons. As a result of the expansion, we have larger and larger neutronic complexes which by emission of β -particles turn into the known elements (Kragh, 1996).

However, Gamow and his collaborators soon realize that the early cosmos must have been a very hot place, and in a series of papers, Alpher et al. (1948a); Alpher (1948); Gamow (1948) develop the hot BBT. It reformulates the cold, matter-filled model into a hot, radiation-filled model of the early cosmos (Kragh, 1996). The basic mechanism is a decay of neutrons into protons and electrons followed by subsequent neutron capture and β decay, which are needed to form heavier elements (Kragh, 1996). As a result, the formation of the nuclei of the elements can be explained and a very rough agreement with the empirical abundance curve is reached. However, this fit is only reached under the artificial assumption that this process does not begin at $t = 0$, but at a later stage (Kragh, 2014, 1996). As a result, this model can successfully predict the abundance of light elements, but it has great difficulties to account for the heavier elements. Indeed, Gamow, Alpher, and Herman cannot find nuclear reaction mechanisms for the primordial formation of heavier elements (Kragh, 2014).

In his 1948 paper, Gamow already realizes the importance of the decoupling time, which is

2.2. Are we there yet? - The road to the standard model

the epoch when the radiation and matter density of the universe are equal, and [Alpher and Herman \(1948\)](#) use the fact that the radiation density decreases much quicker than the matter density to predict a CMB with a temperature of 5K ([Kragh, 1996](#)). However, this prediction was not very explicitly formulated in the paper and it did not attract much interest among scientists ([Kragh, 2014](#)).

The hot BBT is further improved when [Alpher et al. \(1948b\)](#) realize that thermonuclear processes following the initial neutron capture process have to be taken into account and their calculations predict that helium makes up between 25% and 35% of the total mass in the universe. Unfortunately the observational data at this time is not very accurate, so the comparison with the data is of limited value ([Kragh, 1996](#)). Two years later, [Hayashi \(1950\)](#) shows that protons and neutrons in the early universe must have been in thermal equilibrium due to the weak force. Consequently it is no longer necessary to arbitrarily assume that the early universe was full of neutrons. Instead, the abundances of neutrons and protons can now be exactly calculated ([Longair, 2004](#)). Over the following years, the hot BBT is gradually improved and a comprehensive theory is presented in [Alpher et al. \(1953\)](#).

However, the interest in the hot BBT soon disappears when [Burbidge et al. \(1957\)](#) show in their famous paper that the heavy elements are formed in stars and not in the early universe. Their theory can explain the abundances of virtually all elements, so the creation of metals in the early universe and thus the hot Big Bang are no longer required ([Kragh, 2014](#)). Ironically, the [Burbidge et al. \(1957\)](#) theory cannot explain the abundances of helium and deuterium, which the [Alpher et al. \(1953\)](#) theory can provide. But this theory relies on the still very hypothetical model of the expanding universe and it was already shown that it cannot predict the heavy element abundances, which it was originally designed to do. As a result, the interest of the scientific community vanishes, but not for long ([Kragh, 2014](#)).

Indeed, the FLRW cosmology model still rests only on Hubble's distance-velocity relation, and it has a major problem: Cosmological models of an expanding universe, at least those without cosmological constant Λ , predict an age of the universe which is much lower than the age of the earth determined from radioactivity measurements and the age of the stars from astrophysical data ([O'Raiheartaigh et al., 2015](#)). Some relief to this tension comes when Hubble's measurement of H_0 is revised to a substantially lower value ([Sandage, 1958](#); [Baade, 1956](#); [Kragh, 2014](#)). Nevertheless, it seems very promising to search for plausible alternatives to the "Big Bang" theories, a name which is coined by the astronomer Fred Hoyle during a BBC interview on 28 March 1949 ([Kragh, 2013](#)). The [Hoyle \(1948\)](#) and [Bondi and Gold \(1948\)](#) papers developed a new type of "steady state" models of the cosmos, which could explain Hubble's relation just as well ([Kragh, 2014](#)). Their theory postulated the creation of matter from vacuum and this allowed their universe model to be at the same time not only dynamic and expanding, but also unchanging and eternal. It is evident that cosmology is still a very uncertain "science" in the 1950s and due to its speculative character and lack of data, it is not taken too seriously. The astrophysicist Hermann Bondi writes in his autobiography

Chapter 2. Modern cosmology

that “*I always detest being referred to as a cosmologist*” (Kragh, 2014).

However, this changes dramatically over the following decades, when the advent of ever more and ever better data revolutionizes the field of cosmology and turns it into a highly respected science. First, Osterbrock and Rogerson (1961) estimate a cosmic helium content of 32% and thus give the first solid evidence for the hot BBT (Kragh, 2014). Only four years later, Penzias and Wilson (1965) are puzzled when their radio antenna measures a completely isotropic background radiation with a temperature of roughly 3 K and Dicke et al. (1965) correctly interpret this signal as the CMB (Kragh, 2014). This firmly establishes the BBT as the consensus cosmological model and rules out the “steady state” theory.

During the 1960s, Schmidt (1963) discovers the first Quasi-stellar radio source (Quasar) and these objects subsequently permit the study of the universe at higher redshifts. The observations of Bahcall et al. (1967) give reliable evidence that the fine-structure constant does not vary significantly with time and thus that it is reasonable to assume that the laws of physics do not change, which is always an implicit assumption in cosmological models. In addition, the papers of Peebles (1966) and Wagoner et al. (1967) derive the correct cosmic helium, deuterium, and lithium abundances from the BBT and compare it to data, which strengthens the theories’ credibility.

In the following decade, Zwicky’s DM celebrates its big comeback onto the stage of astrophysics. Over many years, Vera Rubin and Kent Ford study the rotation curves of galaxies and find that they are flat, even though the distribution of their luminous mass predicts that they should decline (Rubin and Ford, 1970; Rubin et al., 1978, 1980). Consequently, the galaxies must have much more DM than luminous matter. In combination with the findings of Zwicky (1933), this indicates that both clusters of galaxies and galaxies themselves must have additional halos of DM. So the universe is once again much more mysterious than scientists had imagined. A truly imaginative idea is published by Tryon (1973), who proposes that the universe is the result of a quantum fluctuation. He proposes that the cosmos has a net energy of zero, because the negative gravitational potential energy offsets the energy contained in all of the mass in the universe, and thus the law of energy conservation is not violated. Using less speculative physics and the BBT, Steigman et al. (1977) can use cosmological and astrophysical arguments to show that at most 5 heavy lepton types exist.

While expanding cosmology models based on the FLRW metric are now the de-facto standard due to their ability to account for the distance-velocity relation, the CMB, and the abundances of the light elements, they are not free of problems. Two of these problems are already implicitly contained in the cosmological principle: The universe is homogeneous and isotropic on the largest scales, but why? When we observe distant objects, we look back in time, because the speed of light c is constant. When we look far enough back, we will find objects which cannot have been in causal contact with each other, because their separation is so large that light would have taken more time than the age of the universe at this epoch to cross this distance.

2.2. Are we there yet? - The road to the standard model

Nevertheless, we observe homogeneity and isotropy. This problem is often called the horizon problem (see e.g. [Misner, 1969](#)). Another problem is the creation of structure: If the universe is homogeneous and isotropic, how can cosmic structures form by gravitational interaction? These problems are elegantly solved by a theory called “inflation”. A first semi-realistic model for inflation was created by Starobinsky ([Starobinskiĭ, 1979](#); [Starobinsky, 1980](#)), more realistic ones by [Guth \(1981\)](#), [Linde \(1982\)](#), and [Albrecht and Steinhardt \(1982\)](#), and finally [Linde \(1983\)](#) introduced his “chaotic inflation” scenario ([Linde, 2008](#)). In a nutshell, the theory postulates that a hypothetical inflation field existed in the very early universe. Within a tiny fraction of a second, this field expands the cosmos exponentially and thus exponentially dilutes any inhomogeneities and curvature. Due to quantum fluctuations of this field, the duration of the inflationary phase can differ slightly at different spots in the universe, and this gives rise to small over- and underdensities in a largely homogeneous and isotropic universe. These are the seeds for the formation of cosmic structure. While the required inflation field is still completely hypothetical, the theory can successfully explain several cosmological problems and naturally produces quantum fluctuations, which were already proposed as seeds for cosmic structure formation by [Mukhanov and Chibisov \(1981\)](#).

Over the last decades, data has completely transformed the field of cosmology, and now there is yet another revolution taking place thanks to the rapid progress of computing technology. Several groups are incorporating the recently discovered DM into numerical simulations of the cosmos. They implement it once as so-called Hot Dark Matter (HDM), which means that the individual DM particle is quite light and thus travels with large velocities, and as CDM, which means that the DM particle is quite heavy and travels at lower speeds. They find that simulations with HDM do not match the observed universe, while those with CDM do (e.g., [Davis et al., 1985](#); [White et al., 1987](#)). However, the intermediate case called Warm Dark Matter (WDM) cannot be ruled out. Big surveys like the Center for Astrophysics 2 Redshift Survey map the Large-Scale Structure (LSS) of the universe and discover structures like the “Great Wall” ([Geller and Huchra, 1989](#)).

Up to now, we have taken a detailed mathematical look at the universe itself and how it develops with time. One may now wonder how the structure in the universe, like galaxies and clusters, form and how this can be incorporated into numerical simulations of the universe². For this purpose, we investigate the growth of overdensities over cosmic time. We will illustrate this by using a simple cosmology with $\Omega_m \approx \Omega_{\text{tot}} = 1$ and $p = 0$, which is approximately an Einstein-de Sitter model, and we follow once more the presentation given in [Bartelmann and Schneider \(2001\)](#). The first Friedmann equation 2.33 reduces in the Einstein-de Sitter limit to ([Bartelmann and Schneider, 2001](#))

$$H(t) = H_0 a^{-\frac{3}{2}} \tag{2.43}$$

²Note that this generally requires very powerful computers and only became feasible in recent times

and the age of the universe is

$$t_{\text{universe}} = \frac{2}{3H_0}. \quad (2.44)$$

We assume that inflation resulted in small density fluctuations which are uncorrelated and have a Gaussian distribution of their amplitudes (Bartelmann and Schneider, 2001). These are the seeds for the cosmic structure. The amplitudes of the relative density fluctuations grow by gravitational instability. As long as the relative density contrast is $\ll 1$, we can treat them as small perturbations of the homogeneous and isotropic background density and use linear perturbation theory (Bartelmann and Schneider, 2001). We further assume that the universe density is dominated by CDM at late epochs $a \gg a_{\text{eq}}$, where a_{eq} is the scale factor at matter-radiation equality, when the energy density of non-relativistic matter is equal to the energy density of relativistic matter and radiation (Bartelmann and Schneider, 2001). We define the density contrast

$$\delta(\vec{x}, a) = \frac{\rho(\vec{x}, a) - \bar{\rho}(a)}{\bar{\rho}(a)}, \quad (2.45)$$

where \vec{x} is the comoving position and

$$\bar{\rho} = \frac{\rho(t_0)}{a^3} \quad (2.46)$$

is the average cosmic density. The horizon size is the size of causally connected regions in the universe. It is given by the distance which light can cross in time t since the Big Bang (Bartelmann and Schneider, 2001),

$$d_{H,\text{phys}} = \frac{c}{H(a)}, \quad (2.47)$$

and thus the comoving horizon size is

$$d_{H,\text{comov}} = \frac{c}{aH(a)}. \quad (2.48)$$

The Hubble radius is the length (Bartelmann and Schneider, 2001)

$$\frac{c}{H_0} = \frac{3}{h} \text{ Gpc}. \quad (2.49)$$

Relativistic and non-relativistic perturbation theory show that perturbations with $\delta \ll 1$ grow in the Einstein-de Sitter limit like

$$\delta(a) \propto a^{n-2} = \begin{cases} a^2 & \text{before } a_{\text{eq}} \\ a & \text{after } a_{\text{eq}}. \end{cases} \quad (2.50)$$

We now decompose δ into Fourier modes. As long as we can use linear perturbation theory, individual Fourier components evolve independently (Bartelmann and Schneider, 2001). A

2.2. Are we there yet? - The road to the standard model

perturbation of comoving wavelength λ enters the horizon when

$$\lambda = d_{H,\text{comov}}(a). \quad (2.51)$$

Until a_{eq} , the expansion timescale of the universe $1/H$ is determined by the radiation density and it is shorter than the collapse timescale of DM. Thus the fast radiation-driven expansion prevents DM perturbations from collapsing, but this effect is limited to scales smaller than the horizon. Larger-scale perturbations are consequently not affected. Therefore the horizon size at matter-radiation equality $d_{H,\text{comov}}(a_{\text{eq}})$ is an important scale for growth of structure (Bartelmann and Schneider, 2001). A perturbation with $\lambda < d_{H,\text{comov}}(a_{\text{eq}})$ will be suppressed by the factor (Bartelmann and Schneider, 2001)

$$f_{\text{suppression}} = \left(\frac{a_{\text{enter}}}{a_{\text{eq}}} \right)^2. \quad (2.52)$$

The density perturbation enters the horizon at the scale factor a_{enter} which satisfies the condition (Bartelmann and Schneider, 2001)

$$\lambda = d_{H,\text{comov}} = \frac{c}{a_{\text{enter}} H(a_{\text{enter}})}, \quad (2.53)$$

and in the Einstein-de Sitter limit, we have

$$\lambda \propto \begin{cases} a_{\text{enter}} & \text{if } a_{\text{enter}} \ll a_{\text{eq}} \\ a_{\text{enter}}^{\frac{1}{2}} & \text{if } a_{\text{eq}} \ll a_{\text{enter}} \ll 1. \end{cases} \quad (2.54)$$

We define now the wave number of the perturbation (Bartelmann and Schneider, 2001)

$$k = \frac{1}{\lambda} \quad (2.55)$$

and

$$k_0 = \frac{1}{d_{H,\text{comov}}(a_{\text{eq}})} \quad (2.56)$$

to reformulate the suppression factor as

$$f_{\text{suppression}} = \left(\frac{k_0}{k} \right)^2. \quad (2.57)$$

We can completely characterize the Gaussian density fluctuations $\delta(\vec{x})$ by their power spectrum $P_\delta(k)$, which we define by (Bartelmann and Schneider, 2001)

$$\langle \hat{\delta}(\vec{k}) \hat{\delta}^*(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} - \vec{k}') P_\delta(k), \quad (2.58)$$

where $\hat{\delta}(\vec{k})$ is the Fourier transform of δ , δ_D is the Dirac δ -distribution, and $*$ denotes complex conjugation. Strictly speaking, the Fourier decomposition is only valid in flat space, but at

Chapter 2. Modern cosmology

early enough times space is flat for any cosmological model and at late times the interesting k^{-1} are much smaller than the curvature radius of the universe, so it is acceptable to use it (Bartelmann and Schneider, 2001). Let us now investigate the primordial power spectrum at very early time (Bartelmann and Schneider, 2001),

$$P_i(k) = |\delta_i^2(k)|. \quad (2.59)$$

The density contrast grows according to equation 2.50 and thus (Bartelmann and Schneider, 2001)

$$P_\delta(k) \propto a^{2(n-2)} \quad (2.60)$$

and at a_{enter} , the spectrum has evolved to

$$P_{\text{enter}}(k) \propto a_{\text{enter}}^{2(n-2)} P_i(k) \propto \frac{1}{k^4} P_i(k), \quad (2.61)$$

where we used equation 2.54 for $k \gg 1/d_{H,\text{comov}}(a_{\text{eq}})$. The total power of the density fluctuations at a_{enter} is typically assumed to be scale-invariant, and thus we have

$$P_{\text{enter}}(k) k^3 = \text{constant} \quad (2.62)$$

and the primordial spectrum scales with k ,

$$P_i(k) \propto k \quad (2.63)$$

and it is called the Harrison-Zel'dovich spectrum (Bartelmann and Schneider, 2001). Finally, we combine this with the suppression of small-scale modes in equation 2.57 to obtain the power spectrum

$$P_\delta(k) \propto \begin{cases} k & \text{for } k \ll k_0 \\ k^{-3} & \text{for } k \gg k_0. \end{cases} \quad (2.64)$$

We see now that our cosmology model predicts little power for structures on the largest scales (small k), a peak on intermediate scales, and again a drop off and little power on the smallest scales (large k). We have assumed that the energy density is dominated at late epochs by CDM, and thus we have now a description for the power distribution for CDM over different scales. Baryonic matter falls into the gravitational wells produced by the dominant CDM and starts to evolve on cosmic timescales to form stars, galaxies, and clusters of galaxies. We can simulate this numerically by assuming observationally motivated prescriptions for cooling efficiency, SMBH feedback, etc. (see e.g. Rees and Ostriker, 1977, for a classic paper on gas cooling and galaxy formation). We still have to normalize the power spectrum and we can do this using one of the three following methods (Bartelmann and Schneider, 2001):

2.2. Are we there yet? - The road to the standard model

1. Normalization by CMB temperature anisotropies, which can be translated into an amplitude for $P_\delta(k)$.
2. Normalization by the local variance of counts of galaxies, which are biased tracers of the underlying CDM fluctuations, within a certain volume. Conventionally, the variance of galaxy counts is measured within spheres of radius $8/h$ Mpc and the result is approximately $\sigma_{8,\text{galaxies}} \approx 1$.
3. Normalization by the local abundance of galaxy clusters, which form by gravitational instability from CDM density perturbations. Their spatial number density reflects the amplitude of appropriate CDM fluctuations in a very sensitive manner. Typical scales for CDM fluctuations which collapse to galaxy clusters are $\approx 10/h$ Mpc.

Last, but not least, we have to note that we did not account for non-linear growth of density fluctuations, which becomes important at late stages of the evolution and on small scales. As a result, fluctuations of different sizes interact and the power spectrum must be evaluated numerically ([Bartelmann and Schneider, 2001](#)).

Around 1990, cosmology has progressed dramatically, but the values for the important parameters H_0 , Ω_m , Ω_{tot} , and Ω_Λ are still very uncertain. There is no consensus about a standard cosmology model at all. This starts to change now, and the revolutionary data quality from new instruments is playing a major role in this story. At the beginning of this exciting decade, the Cosmic Background Explorer (COBE) observes the CMB with unprecedented accuracy and precision, see figure 2.9. The data shows that the CMB is indeed an almost perfect black body with temperature $T \approx 2.73$ K ([Mather et al., 1990, 1994](#)) and furthermore, it permits the detection of tiny CMB anisotropies, which are the seeds of cosmic structure ([Smoot et al., 1992](#)). These are well described by scale-invariant fluctuations with a Gaussian distribution and the computed primordial fluctuation power-law spectrum is consistent with the Harrison-Zel'dovich spectrum, exactly as predicted by inflation ([Smoot et al., 1992](#)). Only a few years later, the cosmological constant Λ , Einstein's "biggest blunder", makes a spectacular comeback: [Riess et al. \(1998\)](#) and [Perlmutter et al. \(1999\)](#) use supernovae of type Ia as standard candles to measure distances in the local universe and combine their results with the respective redshifts. They find that the expansion of the universe is not slowing down, but it is accelerating! This acceleration is driven by the Ω_Λ , which is not only non-zero, but dominating the total energy density of the cosmos at the current epoch. At the same time, advanced probes of the CMB like the Balloon Observations Of Millimetric Extragalactic Radiation AND Geophysics (BOOMERanG) and Millimeter Anisotropy eXperiment IMaging Array (MAXIMA) experiments detect the first acoustic peak in the CMB anisotropy power spectrum. The data strongly suggests that the cosmos is approximately spatially flat, $\Omega_{\text{tot}} \approx 1$, and that the baryonic matter is only responsible for a few percent of Ω_{tot} (e.g., [Jaffe et al., 2001](#)).

In the 2000s, the evidence for a standard cosmology model continues to mount up. A

large amount of high quality data in combination with ever-increasing computing power ushers in a new era of “precision cosmology”. Parameters which were only loosely constrained about a decade ago can now be precisely determined. The Two degree Field Galaxy Redshift Survey (2dF) measures Ω_m and it finds that matter makes up about 30% of the energy density in the universe (Peacock et al., 2001). In combination with the CMB data, this confirms the findings of the supernova experiments that Λ dominates Ω_{tot} . The study of the CMB is further improved with the launch of the next generation of space-based CMB experiments, namely the Wilkinson Microwave Anisotropy Probe (WMAP), see figure 2.9. It finds that the measured anisotropies are highly consistent with the standard model of cosmology and inflation (e.g., Spergel et al., 2003). Two years later, the 2dF and the Sloan Digital Sky Survey (SDSS) detect Baryon Acoustic Oscillations (BAO) in the galaxy distribution, which are periodic fluctuations in the density of the visible baryonic matter. Their origin lies in the fluctuations of the CMB and their length scale can serve as a cosmic standard ruler. The BAO results also confirm the cosmological model (Eisenstein et al., 2005; Cole et al., 2005).

Finally, we have arrived at a consensus, standard cosmological model. Its energy density is dominated by DE at the current epoch, and its mass density is dominated by CDM, see figure 2.6. Therefore the model is called Λ Cold Dark Matter (Λ CDM). The current decade further lowers the error bars on the most important cosmological parameters. The Planck space observatory again pushes the limits for the study of the CMB and provides a precise determination of the Ω parameters (Planck Collaboration et al., 2016). Supernova measurements supported by the Hubble Space Telescope (HST) and time-delay measurements from gravitational lensing reduce the uncertainty for the Hubble constant H_0 (Riess et al., 2016; Bonvin et al., 2017). BAO measurements made with the Baryon Oscillation Spectroscopic Survey (BOSS) determine the value of H for higher redshift (Delubac et al., 2015). The Hubble Frontier Fields (HFF) initiative uses gravitational lensing clusters as cosmic telescopes to study the early universe and reionization (Lotz et al., 2017; Atek et al., 2015a). Weak gravitational lensing and galaxy clustering surveys like the Kilo-Degree Survey (KiDS), the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS), and the Dark Energy Survey (DES) constrain the cosmic matter density and σ_8 (Hildebrandt et al., 2017; Joudaki et al., 2017; DES Collaboration et al., 2017). Finally, the Laser Interferometer Gravitational-Wave Observatory (LIGO) detects for the first time gravitational waves and starts a new era of astronomy by enabling multi-messenger observations (Abbott et al., 2016).

Cosmological science is preparing for a bright future: In the next few years, the new space-based observatories James Webb Space Telescope (JWST) and Euclid will be launched. On the ground, huge new telescopes like the European Extremely Large Telescope (E-ELT) and survey telescopes like the Large Synoptic Survey Telescope (LSST) are under construction. Already existing telescopes like the Very Large Telescope (VLT) have been upgraded with fantastic instruments like the Multi Unit Spectroscopic Explorer (MUSE). And in the more distant future, the space-based Wide Field Infrared Survey Telescope (WFIRST) and Laser Interferometer Space Antenna (LISA) will provide new insights on the universe. But now that

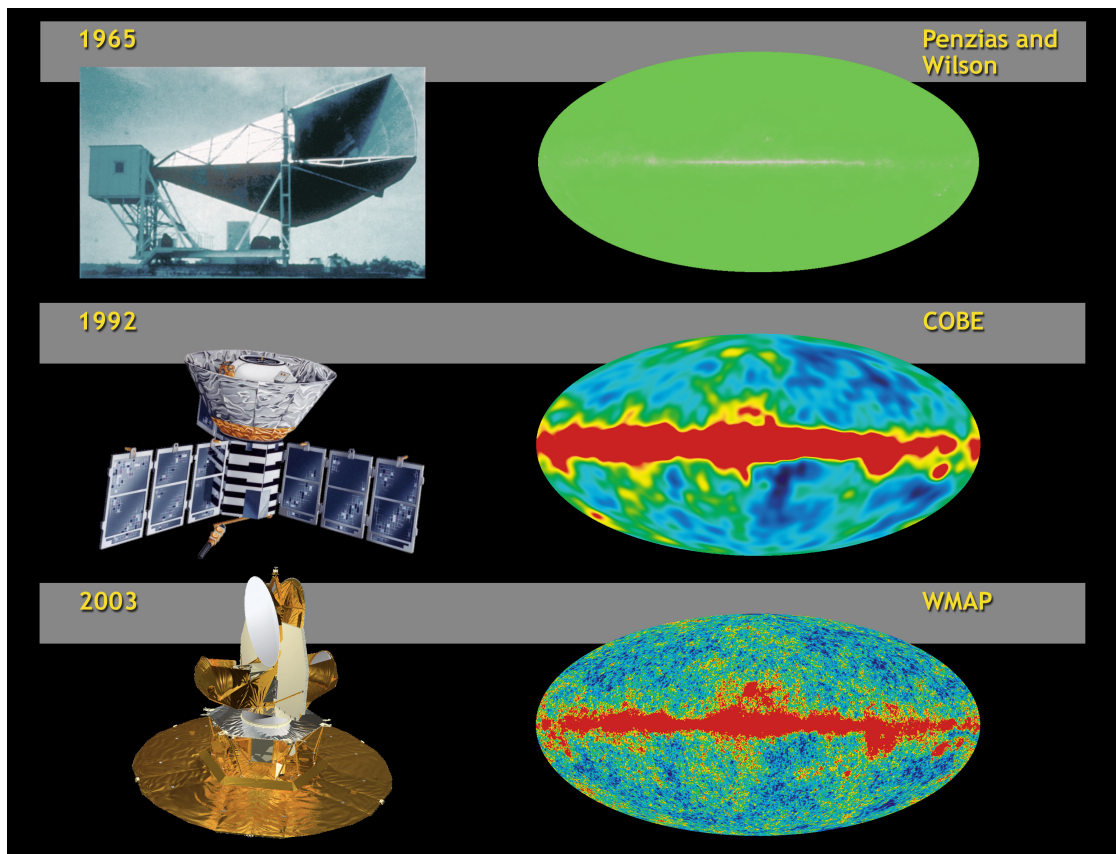


Figure 2.9 – The observational data of the CMB has dramatically improved over the last decades due to the technological progress. The radio antenna of Penzias and Wilson (top left) could only measure an excess temperature due to the CMB. The top right image shows a simulation of the sky viewed with their instrument. The COBE satellite (center left) was a giant leap forward, and the data quality has once again greatly improved with the WMAP space observatory (bottom left) and most recently with the Planck satellite (not shown). The sky shown in the center right image is the four-year map from COBE and the sky on the bottom right is a three-year WMAP map. Note that the shown sky maps are not cleaned, so they show the CMB and unrelated signals, for example the large additional signal due to the Milky Way. These noise signals are removed before working with the data, see further figure 2.4. (Image credit: NASA / WMAP Science Team)

we have a concordance cosmological model which is determined up to a few percent, what is left to discover? As we discuss in the next section, more than enough exciting discoveries still await.

2.3 Open questions

The Λ CDM cosmology model shows that the universe is completely dominated by DE and DM, but what is their physical nature? The very successful standard model of particle physics

cannot account for their existence, and the type of matter which the standard model describes so accurately only makes up about 5% of today's energy density of the cosmos. Cosmological data has already ruled out HDM and thus constrained the possible mass of the hypothetical DM particle. However, is DM really cold or is it actually WDM, the intermediate case between the CDM scenario with very heavy particles and the HDM with very light particles? In addition, it is assumed that the particle does not interact with light, but is this really true or is the interaction just so weak that we cannot yet detect it? Does DM interact with other DM particles? A good DM candidate has to account for all these open questions. It was suggested that a major part of the missing matter density might be accounted for by non-luminous baryonic matter, e.g. in the form of a Massive Astrophysical Compact Halo Object (MACHO), but this was ruled out by using microlensing (Tisserand et al., 2007). It is much more promising to investigate proposed extensions to the standard model of particle physics, for example the Weakly Interacting Massive Particle (WIMP), the Axion, and the sterile neutrino. But since DM is invisible, how can we study it? This is where gravitational lensing comes into play. The lensing effect is sensitive to the total matter, i.e. dark and baryonic matter, and can therefore be used to trace the DM distribution. Thus it is possible to study DM in galaxies and clusters, which are dominated by DM. In addition, we can also use gravitational lensing studies of cluster collisions to obtain upper bounds on the DM self-interaction cross-section (e.g., Randall et al., 2008; Bradač et al., 2008; Harvey et al., 2015; Wittman et al., 2017). Gravitational lensing can be divided into two regimes: Strong Gravitational Lensing (SL) and Weak Gravitational Lensing (WL). We will discuss this further in the next chapter.

The second big unknown is the nature of the cosmological constant Λ . Why does it exist and what is its physical nature? Is it really a constant or does it very slowly vary with time? Is it related to the vacuum energy, which also has negative pressure? A highly precise determination of Ω_Λ at different ages of the universe might yield some insights. It can be probed in different ways, e.g. with supernova measurements (Riess et al., 1998; Perlmutter et al., 1999) or with cluster cosmography from SL (Jullo et al., 2010). Another example is the DES which will constrain Λ by combining measurements of supernovae, BAO, galaxy clusters, and WL (see e.g. DES Collaboration et al., 2017, for first cosmology results).

While the inflationary model is supported by the astronomical data, it is completely unclear what the inflation field is and how it relates to the standard quantum field theory. Is it related to a hypothetical Grand Unification of the forces? Similarly, the baryon-anti-baryon asymmetry is strongly suggested by astronomical data - and by the simple fact that we exist. Its origin cannot yet fully be explained by particle physics, but it is known that the standard models of particle physics and cosmology fulfill the three Sakharov (1967) conditions, namely B violating interactions, C and CP violating interactions, and departure from thermal equilibrium. However, the amount of CP violation and the strength of the electroweak phase transition are not enough to account for the magnitude of the asymmetry (Iocco et al., 2009). Similarly, an explanation for the lepton-anti-lepton asymmetry is needed. However, the answers to these questions will have to come from particle physics theory and experiment rather than from

cosmology. The sum of the neutrino masses, on the other hand, can be constrained by cosmological observations, and e.g. [Planck Collaboration et al. \(2016\)](#) find that it is less than 0.23 eV.

Another burning question is when the epoch of reionization of the universe took place and how long it lasted. Observations of Lyman- α emitters suggest that this occurred roughly between redshifts $z \approx 6$ and $z \approx 8$, while the Thomson scattering optical depth from CMB probes seems to indicate a higher redshift ([Stark, 2016](#)). The most dominant sources of reionization might have been faint galaxies, but further data is needed to validate early results. The use of strong lensing galaxy clusters as “cosmic telescopes” can help to find faint sources at high redshifts and thus to contribute to an answer (e.g., [Atek et al., 2015a](#); [Ishigaki et al., 2015](#)).

Another exciting development is the discrepancy between the measured H_0 values from the Planck satellite and local measurements using supernovae and time delays from gravitational lensing ([Planck Collaboration et al., 2016](#); [Riess et al., 2016](#); [Bonvin et al., 2017](#)). Does this indicate interesting new physics or can the values be reconciled, e.g. by accounting for systematic error effects? Now that the statistical errors have shrunk to small values in the era of “precision cosmology”, exciting discoveries might be ahead. However, this requires large data sets to beat down statistical noise, a thorough understanding and accounting for systematic effects in cosmology theory and observation, and the necessary software and hardware power to make all this possible. In short, advanced techniques are required. In this thesis we will present such techniques for one of the most important tools to probe cosmology, namely gravitational lensing.

3 Cosmology from gravitational lensing

3.1 A concise history of gravitational lensing

Before we discuss the theory behind gravitational lensing and derive the lens equation, we will briefly summarize its historical development. As mentioned in section 2.2, the first GR “lensing community” includes the theorist Einstein, who calculated the lensing deflection of a star by the sun in [Einstein \(1916\)](#), and the observers Dyson, Eddington, and Davidson who published their lensing measurements during the 1919 solar eclipse in [Dyson et al. \(1920\)](#). In the same year, [Lodge \(1919\)](#) coins the term “lens” in the context of light deflection and shortly thereafter [Chwolson \(1924\)](#) suggests the existence of so-called “Einstein rings” if source and lens are perfectly aligned. After being approached by the Czech engineer Rudi Mandl, [Einstein \(1936\)](#) discusses the lensing effect by a star and he investigates the position and magnification of the lensed images ([Schneider et al., 2006](#)). He concludes that there is no great chance to observe this effect in practice.

Fritz Zwicky, however, considers lensing by galaxies instead of stars. In [Zwicky \(1937a\)](#), he finds that this effect should indeed be observable. In addition, he suggests to use the lensing galaxies as “cosmic telescopes” to be able to observe very distant galaxies and to use the lensing effect to study the masses of the lensing galaxies. In a second paper, [Zwicky \(1937b\)](#) investigates the likelihood that a background source will be strongly enough lensed to produce multiple images and he concludes that the probability is so large that such lenses must inevitably be discovered.

In the 1960s, [Klimov \(1963\)](#) investigates galaxy-galaxy lensing, [Liebes \(1964\)](#) studies if stars in the Milky Way can lens other stars, and [Refsdal \(1964a,b\)](#) focuses on time delays of different lensed images and derives the basic equations of gravitational lens theory ([Schneider et al., 2006](#); [Kneib and Natarajan, 2011](#)). Furthermore, he suggests to use the time delays to measure H_0 and this is a very fruitful and active field of research today (see e.g. [Bonvin et al., 2017](#)).

Chapter 3. Cosmology from gravitational lensing

Until 1979, however, the gravitational lensing research is purely theoretical since no lenses have been observed¹. In this year, [Walsh et al. \(1979\)](#) discover a lensed Quasar and thus launch the field of observational lensing research. In the same year, the first CCD detectors replace photographic plates, and the astronomical community greatly benefits from this technological upgrade ([Schneider et al., 2006](#)). In 1986, [Lynds and Petrosian \(1986\)](#) and [Soucail et al. \(1987\)](#) discover giant luminous arcs and [Paczynski \(1987\)](#) explains them as background sources which are strongly lensed by galaxy clusters. A theoretical treatment of cluster lensing is e.g. given by [Narayan et al. \(1984\)](#). In the year 1988, [Hewitt et al. \(1988\)](#) discover the first Einstein ring and in the following year [Irwin et al. \(1989\)](#) detect microlensing in one of the four lensed images of a Quasar as predicted by [Chang and Refsdal \(1979\)](#).

In 1990, [Tyson et al. \(1990\)](#) are the first to detect WL. The detection and measurement of WL is made possible by advances in optical imaging cameras, such as large mosaic CCD cameras which cover a nearly degree-sized field-of-view, and specific image analysis tools ([Schneider et al., 2006](#)). In 1993, the first microlensing events towards the Large Magellanic Cloud are reported ([Alcock et al., 1993](#); [Aubourg et al., 1993](#)) as predicted by [Paczynski \(1986\)](#). This effect can be used to search for baryonic matter of low luminosity in our Milky Way, such as Jupiter-sized planets and neutron stars. These objects might contribute to the missing matter in galaxies required to explain the flat rotation curves, but we know today that they cannot replace DM (e.g., [Tisserand et al., 2007](#)). Even weaker WL effects caused by galaxy ensembles and the large-scale distribution of matter are detected ([Brainerd et al., 1996](#); [Van Waerbeke et al., 2000](#); [Bacon et al., 2000](#); [Kaiser et al., 2000](#); [Wittman et al., 2000](#)).

Gravitational lensing has continued to mature and today's lensing studies, in particular WL measurements, benefit greatly from the development of modern telescopes and computing software and hardware. Lensing requires superb data and a well-behaved Point Spread Function (PSF), and this is provided by some ground-based telescopes like the Subaru Telescope, and ideally by space-based telescopes like the HST. Large programs like the Cluster Lensing And Supernova survey with Hubble (CLASH), CFHTLenS, KiDS, DES and the HFF provide a huge amount of high quality data. Efficient algorithms, pipelines, and software have been developed to extract the lensing information (e.g., [Kaiser et al., 1995](#); [Refregier and Bacon, 2003](#); [Okura et al., 2007](#); [Refregier and Amara, 2014](#)), model lensing mass distributions ([Jullo et al., 2007](#); [Jullo and Kneib, 2009](#); [Oguri, 2010](#); [Zitrin et al., 2009](#)), and infer cosmological constraints (see e.g. [Jullo et al., 2010](#); [Joudaki et al., 2017](#); [DES Collaboration et al., 2017](#), for the pipelines they use). The scientific return is exciting and includes constraints on the reionization of the universe (e.g., [Atek et al., 2015a](#); [Ishigaki et al., 2015](#)), the cosmological constant Λ and the cosmic matter density (e.g., [Jullo et al., 2010](#); [Caminha et al., 2016](#)), the distribution of DM substructure (e.g., [Natarajan et al., 2017](#)), the cosmic matter density and σ_8 ([Joudaki et al., 2017](#); [Hildebrandt et al., 2017](#); [DES Collaboration et al., 2017](#)), and the Hubble constant ([Bonvin et al., 2017](#); [Vuissoz et al., 2008](#)). In the near future, space-based missions like JWST and Euclid and ground-based telescopes like the LSST will provide fantastic data sets for

¹Except the lensing by the sun in 1919

lensing.

3.2 Gravitational lensing theory

Galaxies and galaxy clusters are so massive that they locally deform spacetime, so in general gravitational lensing has to be described by the GR equations. However, in this thesis we focus on gravitational lensing by galaxies and clusters, and this allows us to make a few assumptions which dramatically reduce the complexity of the lensing equations. We derive this gravitational lens theory in this section and we follow largely the excellent presentation given in [Bartelmann and Schneider \(2001\)](#).

A sketch of the gravitational lensing scenario is shown in figure 3.1. We assume that there is only one lens along the line-of-sight. The astrophysical distances between source, lens, and observer are much larger than the extent of the lens along the line-of-sight, and this permits to describe the path of the photons by two straight light rays with a kink near the deflector instead of using a smooth, slowly varying curve ([Bartelmann and Schneider, 2001](#)). The magnitude and direction of this kink are given by the deflection angle $\hat{\alpha}$. The source and the lens mass distribution are respectively located at redshifts z_s and z_l and angular diameter distances D_{os} and D_{ol} . The angular diameter distance is defined in analogy to the relation in Euclidean space between the physical cross-section ΔA of an object at redshift z_2 and the solid angle $\Delta\omega$ which it subtends for an observer at redshift z_1 ([Bartelmann and Schneider, 2001](#)),

$$\Delta\omega D^2 = \Delta A. \quad (3.1)$$

Using our results from the FLRW metric discussion in section 2.2, we have ([Bartelmann and Schneider, 2001](#))

$$\frac{\Delta A}{4\pi a^2(z_2) f_k^2[w(z_1, z_2)]} = \frac{\Delta\omega}{4\pi}, \quad (3.2)$$

where $a(z_2)$ is the scale factor at photon emission time and $f_k[w(z_1, z_2)]$ is the radial coordinate distance between observer and source. Therefore we obtain

$$D(z_1, z_2) = \left(\frac{\Delta A}{\Delta\omega} \right)^{\frac{1}{2}} = a(z_2) f_k[w(z_1, z_2)]. \quad (3.3)$$

Let us now consider a point mass M . If the impact parameter ξ is much larger than the Schwarzschild radius of the lens,

$$\xi \gg \frac{2GM}{c^2}, \quad (3.4)$$

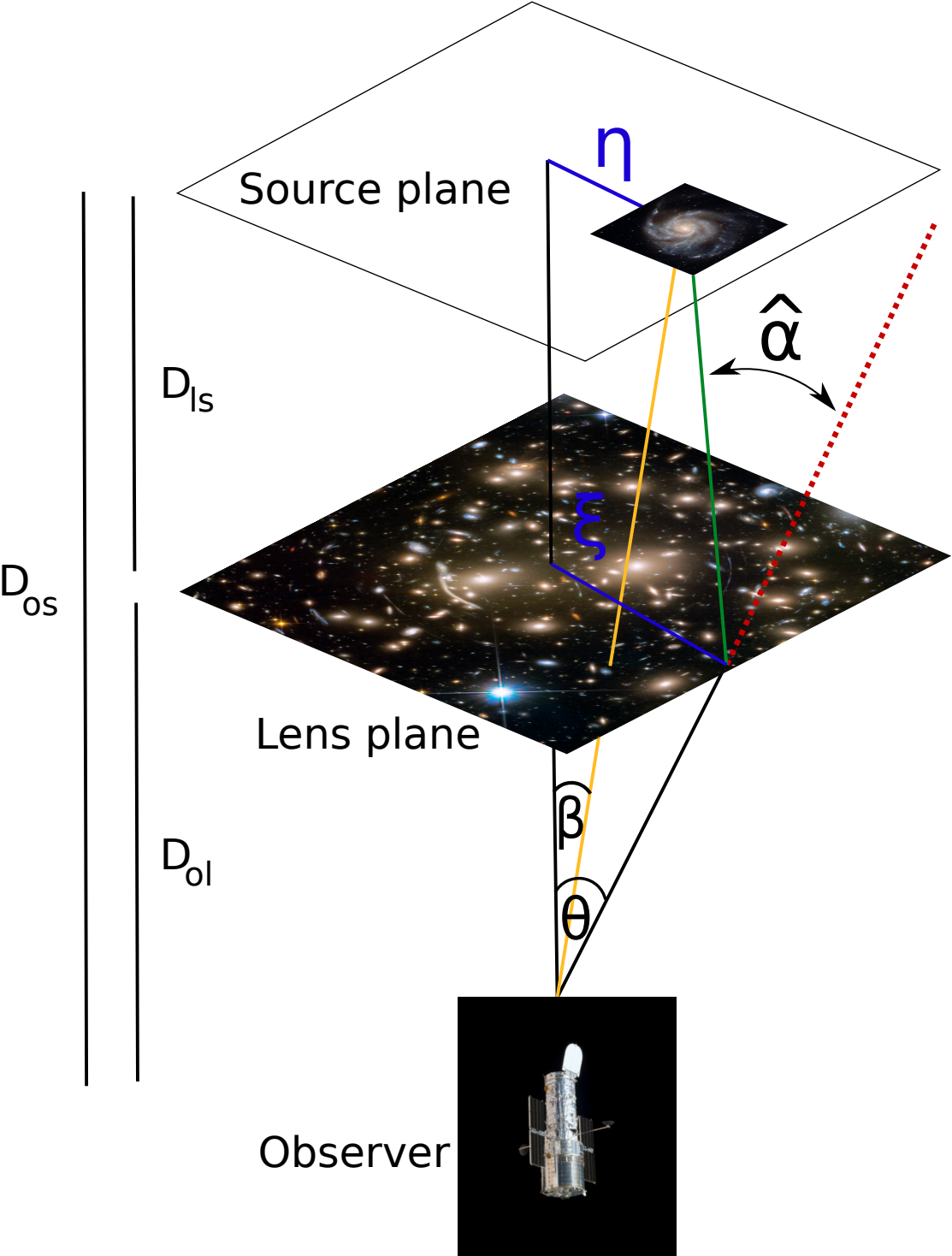


Figure 3.1 – Sketch of a gravitational lens system. (Image credit: Own work / NASA / ESA, see photographic credits)

then the GR prediction for the deflection angle is (Bartelmann and Schneider, 2001)

$$|\hat{\alpha}| = \frac{4GM}{c^2\xi} \quad (3.5)$$

and according to the condition 3.4 the deflection angle is small

$$|\hat{\alpha}| \ll 1. \quad (3.6)$$

In the next step, we observe that we can linearize the field equations of GR in equation 2.7 if the gravitational field is weak. Now we can move to a more general case, because in the weak field limit the deflection angle of an ensemble of point masses is just the vectorial sum of the deflection angles due to individual lenses (Bartelmann and Schneider, 2001).

We look at a three-dimensional mass distribution with density ρ which we divide into cells of size dV and mass dm (Bartelmann and Schneider, 2001),

$$dm = \rho(\mathbf{r})dV. \quad (3.7)$$

A photon which passes this mass distribution will be deflected, but this lensing effect is so small that we can approximate it as a straight line in the neighborhood of the deflecting lens. This corresponds to the Born approximation in nuclear physics (Bartelmann and Schneider, 2001). As a result, the separation of the photon from a mass element dm is given by

$$\boldsymbol{\xi} - \boldsymbol{\xi}_{dm}, \quad (3.8)$$

where $\boldsymbol{\xi}$ is a two-dimensional vector, and we have for the total deflection angle (Bartelmann and Schneider, 2001)

$$\begin{aligned} \hat{\alpha}(\boldsymbol{\xi}) &= \frac{4G}{c^2} \sum dm \frac{\boldsymbol{\xi} - \boldsymbol{\xi}_{dm}}{|\boldsymbol{\xi} - \boldsymbol{\xi}_{dm}|^2} \\ &= \frac{4G}{c^2} \int d^2\xi_{dm} \int dz \rho(\boldsymbol{\xi}_{dm}, z) \frac{\boldsymbol{\xi} - \boldsymbol{\xi}_{dm}}{|\boldsymbol{\xi} - \boldsymbol{\xi}_{dm}|^2}, \end{aligned} \quad (3.9)$$

and the result is again a two-dimensional vector. Now we define the surface mass density

$$\Sigma(\boldsymbol{\xi}) = \int dz \rho(\boldsymbol{\xi}, z) \quad (3.10)$$

and the previous equation becomes (Bartelmann and Schneider, 2001)

$$\hat{\alpha}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int d^2\xi_{dm} \Sigma(\boldsymbol{\xi}_{dm}) \frac{\boldsymbol{\xi} - \boldsymbol{\xi}_{dm}}{|\boldsymbol{\xi} - \boldsymbol{\xi}_{dm}|^2}. \quad (3.11)$$

Chapter 3. Cosmology from gravitational lensing

This equation holds as long as the deviation of the actual light ray from a straight line within the mass distribution is small compared to the scale on which the mass distribution changes significantly. This is the case for all galaxy-galaxy and cluster lensing scenarios, but not for lensing by LSS (Bartelmann and Schneider, 2001). However, as the two former cases are the main topics of this thesis, we can safely assume that this condition is satisfied.

Now we want to relate the true position of the source to the location observed by a telescope. The source and lens planes are defined perpendicular to the line-of-sight of the observer, see figure 3.1. Let $\boldsymbol{\eta}$ be the position of the source in the two-dimensional source plane and recall that $\boldsymbol{\xi}$ is the position in the lens plane. Then we have (Bartelmann and Schneider, 2001)

$$\boldsymbol{\eta} = \frac{D_{\text{os}}}{D_{\text{ol}}} \boldsymbol{\xi} - D_{\text{ls}} \hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \quad (3.12)$$

and if we introduce angular coordinates $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$,

$$\begin{aligned} \boldsymbol{\eta} &= D_{\text{os}} \boldsymbol{\beta}, \\ \boldsymbol{\xi} &= D_{\text{ol}} \boldsymbol{\theta}, \end{aligned} \quad (3.13)$$

this becomes

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{\text{ls}}}{D_{\text{os}}} \hat{\boldsymbol{\alpha}}(D_{\text{ol}} \boldsymbol{\theta}) = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \quad (3.14)$$

where we have defined the scaled deflection angle $\boldsymbol{\alpha}$,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_{\text{ls}}}{D_{\text{os}}} \hat{\boldsymbol{\alpha}}(D_{\text{ol}} \boldsymbol{\theta}). \quad (3.15)$$

Equation 3.14 is the so-called lens equation and it is the central equation of gravitational lensing. The statement of the lens equation is that a source with true position $\boldsymbol{\beta}$ will be observed at position $\boldsymbol{\theta}$ and it is possible that several values of $\boldsymbol{\theta}$ satisfy equation 3.14, so a single source can be multiply imaged!

Let us now investigate how we can distinguish the SL regime, where multiple images occur, and the WL regime, where the lensed images are only slightly distorted. We define the dimension-less surface mass density κ ,

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\text{crit}}}, \quad (3.16)$$

where

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_{\text{os}}}{D_{\text{ol}} D_{\text{ls}}} \quad (3.17)$$

is the critical surface mass density Σ_{crit} (Bartelmann and Schneider, 2001). Note that Σ_{crit} depends on the redshifts of both source and lens! If a mass distribution has $\kappa \geq 1$ somewhere, then it will produce multiple images for some source positions $\boldsymbol{\beta}$ and hence κ is a good estimator to distinguish the SL and WL regimes. Note that $\kappa \geq 1$ is a sufficient but not a necessary condition for multiple images to occur (Bartelmann and Schneider, 2001).

We will now rewrite the scaled deflection angle (Bartelmann and Schneider, 2001)

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \quad (3.18)$$

and this implies that we can write $\boldsymbol{\alpha}$ by using a deflection potential Ψ

$$\Psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \kappa(\boldsymbol{\theta}') \ln(|\boldsymbol{\theta} - \boldsymbol{\theta}'|) \quad (3.19)$$

as

$$\boldsymbol{\alpha} = \nabla\Psi. \quad (3.20)$$

The deflection potential is the two-dimensional analogue of the Newtonian gravitational potential and we have the Poisson equation (Bartelmann and Schneider, 2001)

$$\nabla^2\Psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}). \quad (3.21)$$

Now that we can compute the locations of the images on the sky, let us investigate their shapes. They will differ from the original, undistorted shapes of the unlensed sources, because light bundles are differentially deflected (Bartelmann and Schneider, 2001). Liouville's theorem and the absence of emission and absorption of photons in gravitational light deflection imply that lensing conserves surface brightness (Bartelmann and Schneider, 2001),

$$I(\boldsymbol{\theta}) = I^s[\boldsymbol{\beta}(\boldsymbol{\theta})], \quad (3.22)$$

where I^s is the surface brightness distribution in the source plane and I the observed surface brightness distribution. If a source is much smaller than the angular scale on which the lensing properties of the mass distribution change, we can locally linearize the lens mapping (Bartelmann and Schneider, 2001). The image distortion is given by the Jacobian matrix (Bartelmann and Schneider, 2001)

$$A_{ij}(\boldsymbol{\theta}) = \frac{\partial\beta_i}{\partial\theta_j} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (3.23)$$

Chapter 3. Cosmology from gravitational lensing

where we have introduced the shear γ , see figure 3.2,

$$\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\phi}, \quad (3.24)$$

$$\begin{aligned} \gamma_1 &= \frac{1}{2}(\Psi_{,11} - \Psi_{,22}), \\ \gamma_2 &= \Psi_{,12}, \end{aligned} \quad (3.25)$$

and we have used the GR convention that a comma denotes a partial derivative, which we will continue to use throughout this thesis.

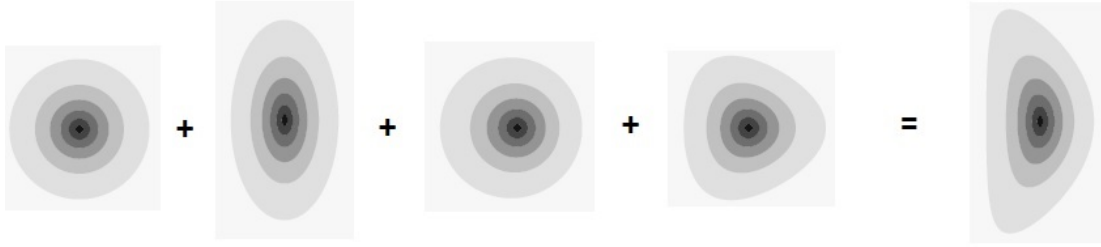


Figure 3.2 – Illustration of gravitational lensing shear γ and first and second flexion \mathcal{F} and \mathcal{G} . First four images, from left to right: An unlensed circular source, shear, \mathcal{F} -Flexion, \mathcal{G} -Flexion. Rightmost image: Shear and flexion together give rise to an arclet. (Image credit: Own work)

We can use the locally linearized lens equation to find for a point $\boldsymbol{\theta}_0$ corresponding to the source point $\boldsymbol{\beta}_0$ (Bartelmann and Schneider, 2001)

$$I(\boldsymbol{\theta}) = I^s[\boldsymbol{\beta}_0 + \mathbf{A}(\boldsymbol{\theta}_0) (\boldsymbol{\theta} - \boldsymbol{\theta}_0)], \quad (3.26)$$

and we see that circular sources are lensed into ellipses. The observed flux and source flux can be computed by integrating over the respective brightness distributions and the ratio of observed flux to source flux is the magnification μ (Bartelmann and Schneider, 2001),

$$\mu(\boldsymbol{\theta}_0) = \frac{1}{\det(\mathbf{A})} = \frac{1}{(1 - \kappa)^2 - |\gamma|^2}. \quad (3.27)$$

Hence the images are deformed in shape and size.

Points where \mathbf{A} is singular, i.e. where

$$\det(\mathbf{A}) = 0, \quad (3.28)$$

form closed curves which are called critical curves. The corresponding curves in the source plane are named caustics (Bartelmann and Schneider, 2001). According to equation 3.27, sources on caustics should be infinitely magnified, but this is not the case in practice, because the geometrical-optics approximation fails near critical curves and a wave-optics description leads to a finite magnification (Bartelmann and Schneider, 2001). Nevertheless, they will be

greatly magnified. Sources which move over a caustic have their number of images changed by ± 2 and the two additional images appear or disappear at the corresponding critical curve in the lens plane. Thus only sources inside a caustic are multiply imaged (Bartelmann and Schneider, 2001).

If γ and κ are not constant over a source image, i.e. the source is not much smaller than the angular scale on which the lens properties change, then we have to expand the locally linearized lens mapping. We obtain (Bacon et al., 2006),

$$\beta_i = A_{ij}\theta_j + \frac{1}{2}D_{ijk}\theta_j\theta_k, \quad (3.29)$$

where

$$D_{ijk} = A_{ij,k}. \quad (3.30)$$

We want to obtain a direct description of the D_{ijk} lensing tensor. Therefore we introduce the first and second flexion \mathcal{F} and \mathcal{G} , which are functions of the derivatives of the two γ components (Bacon et al., 2006),

$$\mathcal{F} = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}, \quad (3.31)$$

$$\mathcal{G} = \begin{pmatrix} \gamma_{1,1} - \gamma_{2,2} \\ \gamma_{2,1} + \gamma_{1,2} \end{pmatrix}. \quad (3.32)$$

Alternatively, we can write the flexions as

$$\mathcal{F} = |\mathcal{F}|e^{i\phi} = \mathcal{F}_1 + i\mathcal{F}_2, \quad (3.33)$$

$$\mathcal{G} = |\mathcal{G}|e^{3i\phi} = \mathcal{G}_1 + i\mathcal{G}_2, \quad (3.34)$$

and we see that \mathcal{F} is a spin-1 field and \mathcal{G} is a spin-3 field, see figure 3.2 (Bacon et al., 2006).

Now we define

$$\mathcal{F}_{ij1} = -\frac{1}{2} \begin{pmatrix} 3\mathcal{F}_1 & \mathcal{F}_2 \\ \mathcal{F}_2 & \mathcal{F}_1 \end{pmatrix},$$

$$\mathcal{F}_{ij2} = -\frac{1}{2} \begin{pmatrix} \mathcal{F}_2 & \mathcal{F}_1 \\ \mathcal{F}_1 & 3\mathcal{F}_2 \end{pmatrix}, \quad (3.35)$$

$$\mathcal{G}_{ij1} = -\frac{1}{2} \begin{pmatrix} \mathcal{G}_1 & \mathcal{G}_2 \\ \mathcal{G}_2 & -\mathcal{G}_1 \end{pmatrix},$$

$$\mathcal{G}_{ij2} = -\frac{1}{2} \begin{pmatrix} \mathcal{G}_2 & -\mathcal{G}_1 \\ -\mathcal{G}_1 & -\mathcal{G}_2 \end{pmatrix}, \quad (3.36)$$

Chapter 3. Cosmology from gravitational lensing

and we can express the D_{ijk} lensing tensor in terms of the flexions as

$$D_{ijk} = \mathcal{F}_{ijk} + \mathcal{G}_{ijk}. \quad (3.37)$$

We see that the flexions are sufficient to fully describe the expansion of the locally linearized lens mapping. The effect of flexion on the lensed images is as follows. In a typical flexion scenario, we have simultaneously non-zero shear γ and non-zero first and second flexions \mathcal{F} and \mathcal{G} . A circular source is thus both sheared and flexed and it is deformed into an arclet instead of an ellipse, see figure 3.2. The flexion is typically negligible far away from the SL area of the lens, but it has to be included if the lensed image is relatively close to it, see figure 3.3. Thus flexion can be regarded as the intermediate case between WL and SL.

We end this section with an illustrative example. We look at a simple lens model, the Singular Isothermal Sphere (SIS). It has a relatively simple mathematical expression and is thus very instructive. The physical motivation for this lens model is as follows: The rotation curves of spiral galaxies are approximately flat due to the DM component. The DM halo necessary to produce these rotation curves has a density profile $\rho \propto 1/r^2$ for large r (Bartelmann and Schneider, 2001). The projected mass density then drops off like $1/\theta$. Such density profiles can be obtained if the velocity dispersion of the DM particles is spatially constant and they are hence called isothermal profiles (Bartelmann and Schneider, 2001). The projected surface mass density is

$$\Sigma(\boldsymbol{\theta}) = \frac{\sigma_v^2}{2GD_{ol}|\boldsymbol{\theta}|}, \quad (3.38)$$

where σ_v is the line-of-sight velocity dispersion of the visible ‘‘particles’’ (e.g. galaxies in a galaxy cluster), which are assumed to be in virial equilibrium (e.g., Bartelmann and Schneider, 2001). Thus we have

$$\kappa(\boldsymbol{\theta}) = \frac{\theta_E}{2|\boldsymbol{\theta}|}, \quad (3.39)$$

$$\theta_E = 4\pi \left(\frac{\sigma_v}{c} \right)^2 \frac{D_{ls}}{D_{os}}, \quad (3.40)$$

where we defined the Einstein deflection angle θ_E . Using equations 3.19 and 3.20, we find that the magnitude of the scaled deflection angle $\boldsymbol{\alpha}$ is constant,

$$|\boldsymbol{\alpha}| = \theta_E. \quad (3.41)$$

We see that the lens equation has infinitely many solutions for $\boldsymbol{\beta} = \mathbf{0}$, namely each point on the circle with radius θ_E . Therefore a background source at this location will be strongly lensed into a perfect Einstein ring. The shear γ is calculated from equation 3.25 and we obtain

$$\gamma(\boldsymbol{\theta}) = -\frac{\theta_E}{2|\boldsymbol{\theta}|} e^{2i\phi}, \quad (3.42)$$

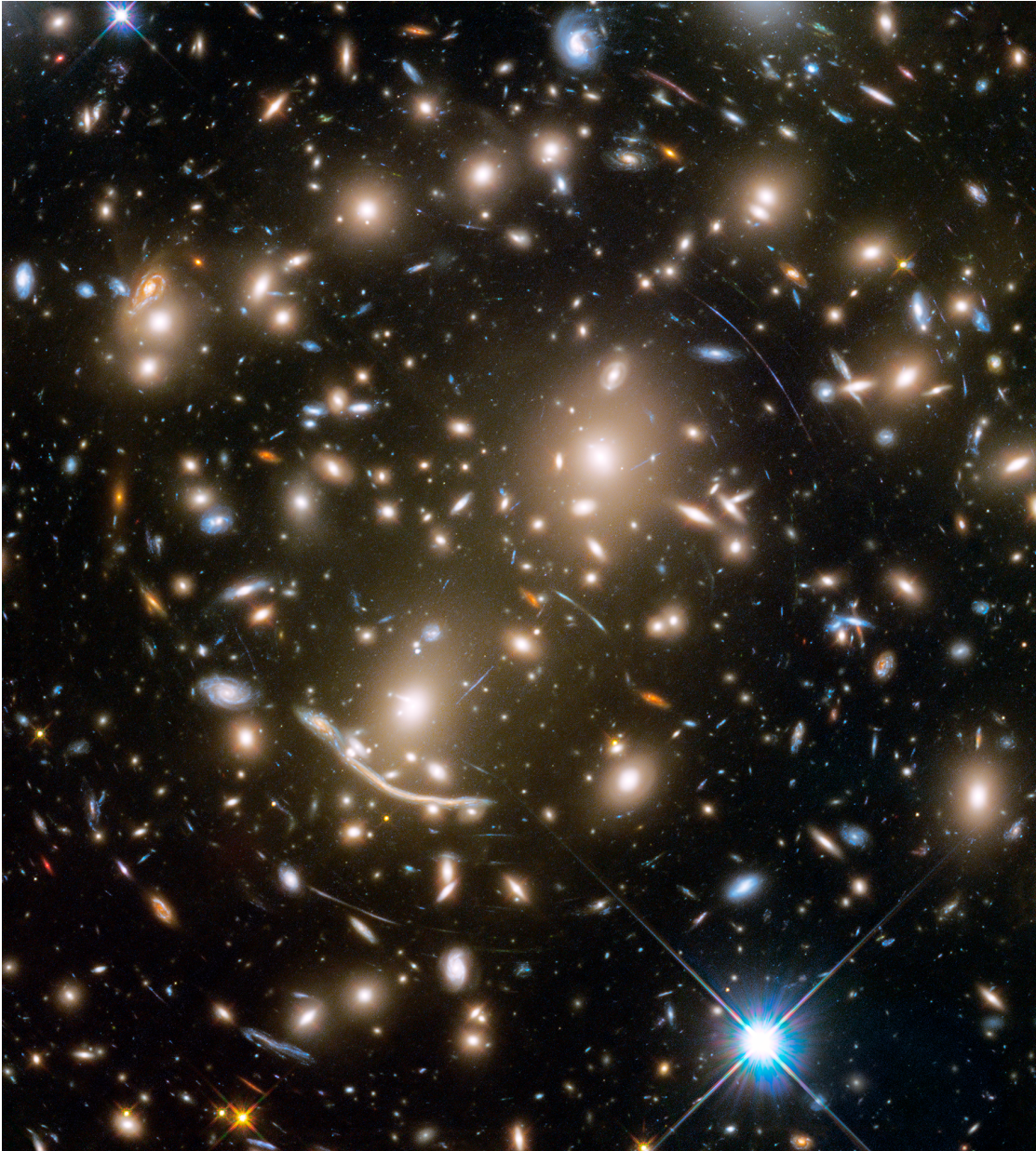


Figure 3.3 – The image shows the strong lensing galaxy cluster Abell 370. It was taken by the HST as part of the HFF program. The cluster members are the yellow elliptical galaxies. Lensed sources closer to the lens center and thus to the SL area have a much stronger arclet shape than those further away, a clear sign that the flexion effect is substantial. (Image credit: NASA / ESA / J. Lotz and the HFF Team / STScI)

and the magnification μ is (Bartelmann and Schneider, 2001)

$$\mu(\boldsymbol{\theta}) = \frac{|\boldsymbol{\theta}|}{|\boldsymbol{\theta} - \boldsymbol{\theta}_E|}, \quad (3.43)$$

so the Einstein ring defines a critical curve. For the two flexions, we obtain (Bacon et al., 2006)

$$\begin{aligned}\mathcal{F} &= -\frac{\theta_E}{2|\boldsymbol{\theta}|^2} e^{i\phi}, \\ \mathcal{G} &= \frac{3\theta_E}{2|\boldsymbol{\theta}|^2} e^{3i\phi}.\end{aligned}\tag{3.44}$$

3.3 Cosmological constraints

A prime cosmological application of the lensing effect is naturally the measurement of the mass in galaxies and clusters. The big advantage of lensing compared to other methods is that it does not depend on the nature of the matter or its state. It is equally sensitive to luminous baryonic matter and to DM, regardless of whether it is in equilibrium or not. Thus it does not need any additional assumptions, which can introduce systematic errors into the measurement (Schneider et al., 2006). For example, Natarajan et al. (2017) investigate the abundance of substructure masses from SL in three HFF clusters and find it to be in agreement with Λ CDM. The significance of such a cosmological test can be greatly improved if the probed substructure is extended to lower-mass halos. The inclusion of flexion measurements are very promising for this purpose (e.g., Lanusse et al., 2016; Bacon et al., 2010a; Leonard et al., 2009) and we study the feasibility of measuring the flexion in chapter 7.

The probability to observe a lensing event depends on the projected number density of lenses, and this density can consequently be studied by investigating well-defined samples of sources and their lensed fraction (Schneider et al., 2006). As a result, we can draw constraints on the redshift evolution of the number of strong lensing galaxies and clusters (Schneider et al., 2006). These can in turn be compared to the number density evolution predicted by cosmology theory and simulation. In a similar spirit, Mantz et al. (2015) use gravitational lensing to calibrate absolute masses in measurements of the number density of massive clusters and its evolution. They use the results to constrain $\sigma_8 \Omega_m^{0.17}$.

Light rays are not only strongly deflected by highly concentrated masses, but they are also continuously slightly lensed by the LSS of the universe. The statistics of the resulting distortions can therefore be used to study the LSS, which in turn depends on the cosmology (Schneider et al., 2006). The first detection of this ‘‘cosmic shear’’ effect was announced in the year 2000 (Van Waerbeke et al., 2000; Bacon et al., 2000; Kaiser et al., 2000; Wittman et al., 2000) and the technique has been steadily improved since. The main observables of this method are the shear correlation functions of e.g. two sheared galaxies,

$$\xi_{\pm}(\theta_{cs}) = \langle \gamma_{t,cs} \gamma_{t,cs} \rangle(\theta_{cs}) \pm \langle \gamma_{x,cs} \gamma_{x,cs} \rangle(\theta_{cs}),\tag{3.45}$$

whereas the mixed correlator

$$\xi_{\times}(\theta_{cs}) = \langle \gamma_{t,cs} \gamma_{\times,cs} \rangle(\theta_{cs}) \quad (3.46)$$

must vanish due to parity symmetry (Schneider et al., 2006). Note that the tangential and cross-component of the shear for this galaxy pair are defined here with respect to the angle ϕ_{cs} of their separation vector θ_{cs} , which is different from the usual definition used in this thesis. The equations are (Schneider et al., 2006)

$$\gamma_{t,cs} = -\text{Re}(\gamma e^{-2i\phi_{cs}}), \quad (3.47)$$

$$\gamma_{\times,cs} = -\text{Im}(\gamma e^{-2i\phi_{cs}}). \quad (3.48)$$

Cosmic shear is to first order sensitive to the cosmological parameter combination (see e.g. Schneider et al., 2006; DES Collaboration et al., 2017)

$$S_8 = \sigma_8 \left(\frac{\Omega_m}{0.3} \right)^{\frac{1}{2}}. \quad (3.49)$$

Like other low- z probes discussed in this section, cosmic shear has the big advantage that it is complementary to CMB cosmology constraints because it measures the respective parameters for recent epochs, whereas the CMB probes high redshifts. In addition, as discussed in section 2.2, the power spectrum P_δ will exhibit non-linear features at recent epochs and on small-scales, so they can be detected by cosmic shear (e.g., Schneider et al., 2006). Indeed, the cosmological constraints from recent surveys might be in tension with Planck CMB results, see further section 3.4.

Already Refsdal (1964b) has suggested to use SL time delays as a cosmological probe. Signal variations in the lensed source will appear first in one multiple image and only after a time interval Δt in the others, because the light had to travel a larger distance along these other geodesics. The time difference is sensitive to the expansion of the universe and thus it probes H_0 . Constraints on the Hubble constant from this method have been successfully derived by e.g. the COSMOGRAIL project (Vuissoz et al., 2008) and more recently by the H0LiCOW collaboration (Bonvin et al., 2017). However, this method requires a very precise lens model. In addition, its results are very sensitive to lensing degeneracies and thus they have to be accounted for (e.g. Schneider and Sluse, 2013, 2014). We will discuss a particular lensing degeneracy, namely the mass-sheet degeneracy, in chapter 4 and present a novel method to break it.

SL galaxy clusters have a large magnification effect μ and this can be exploited to study otherwise unobservably faint objects (see e.g., Kneib et al., 2004; Kneib and Natarajan, 2011; Lotz et al., 2017). Thus they act as “cosmic telescopes”. With this technique, astronomers can study the early universe and derive constraints on its reionization. For example Atek et al. (2015b) and Ishigaki et al. (2015) use HFF cluster lens models (e.g., Jauzac et al., 2014,

Chapter 3. Cosmology from gravitational lensing

2015; Ishigaki et al., 2015; Oguri, 2015; Limousin et al., 2016) to constrain the UV-luminosity function at high redshift. This method demands a very well constrained lens model, as errors in the model will translate into errors on the computed source magnification. In other words, the “cosmic telescope” must be well understood. Consequently, high-quality imaging data is required. In addition, the quality of the lens model also benefits greatly from the availability of spectroscopic redshifts, as we show in chapter 5, and the creation of the models demands efficient software tools, which we discuss and present in chapter 6.

The creation of giant arcs is one of the most spectacular effects of SL, and the number of galaxy clusters which can produce them is sensitive to the initial conditions and the expansion history of the cosmos (see e.g. Kneib and Natarajan, 2011). While it is difficult to distinguish the cosmological influence on this number from cluster physics effects and the properties of the high- z source population, Bartelmann et al. (1998) find that the observed number of giant arcs is much larger than predicted by Λ CDM. Several possible explanations for this discrepancy have been studied, for example cooling and star formation in clusters, feedback from Active Galactic Nuclei, and line-of-sight effects, but they could not resolve this discrepancy (see Meneghetti et al., 2013, for a review). Further work both on the theoretical and on the observational frontier is required to investigate if this effect is truly due to cosmology or whether it has another explanation (Meneghetti et al., 2013). However, if this effect is indeed due to a cosmological influence, it might become a valuable tool once data from DES, Euclid, and the LSST are available (Boldrin et al., 2016).

A similarly spectacular sight is the collision of two galaxy clusters. During these highly energetic events, the baryonic Intracluster Medium (ICM) in the clusters interacts, which is visible in the X-ray part of the electromagnetic spectrum. The individual cluster galaxies, however, have such a low density that they virtually never collide. We can use gravitational lensing to map the distribution of the DM, which dominates the total mass of galaxy clusters. The location of the DM with respect to the collisionless galaxies and the strongly interacting ICM shows that DM is virtually not interacting with itself (e.g., Clowe et al., 2006; Paraficz et al., 2016). Techniques have been developed to compute the DM self-interaction cross-section (e.g., Markevitch et al., 2004; Kim et al., 2017) and upper bounds have been derived (e.g., Markevitch et al., 2004; Randall et al., 2008; Bradač et al., 2008; Harvey et al., 2015). However, the measurement is difficult and the resulting upper bounds are very sensitive to the used observational constraints, see e.g. Wittman et al. (2017).

SL galaxy cluster cosmography constrains Ω_m , Ω_Λ , and the equation-of-state parameter of DE w_Λ , which is defined as the ratio of pressure to energy density and it equals -1 for a cosmological constant Λ (e.g., Jullo et al., 2010; Golse et al., 2002). This technique uses several multiple image families with secure spectroscopic redshifts. Since the lensing strength depends on the angular diameter distance ratio, it will be different for sources at different redshifts and thus different angular diameter distances. The cosmological dependence arises from the sensitivity of the angular diameter distances to the geometry of the universe and

thus to the cosmology. Let us illustrate this by assuming a lens model with fixed parameters and only one lensed source. The lens model shall give a perfect fit to the observational data. Now we vary only the cosmology. This results in a modified angular diameter distance to the source at fixed redshift. This in turn affects the lensing strength and thus the lens model will no longer perfectly fit the data. If we now take several sources at fixed redshifts and compute the lens model for all of them simultaneously, then requiring a good fit to the observational data for all sources leads to constraints on the respective angular diameter distances and thus on the cosmology. We define the “image family ratio”, from which we can extract the cosmological constraints, as (e.g., [Jullo et al., 2010](#); [Kneib and Natarajan, 2011](#))

$$\Xi(z_l, z_{s1}, z_{s2}, \Omega_m, \Omega_\Lambda, w_\Lambda) = \frac{D(z_l, z_{s1}) D(0, z_{s2})}{D(0, z_{s1}) D(z_l, z_{s2})}, \quad (3.50)$$

where z_l, z_{s1}, z_{s2} are the respective redshifts of the lens, source 1, and source 2. Cosmological constraints from the analysis of the very well constrained galaxy clusters Abell 1689 and Abell S1063 were respectively presented in [Jullo et al. \(2010\)](#) and [Caminha et al. \(2016\)](#). [Acebron et al. \(2017\)](#) use cluster simulations to show that not only unimodal clusters like the two aforementioned, but also more complex, multimodal clusters can result in competitive constraints. Further improvements on cosmology constraints from this method require the inclusion of lensing by line-of-sight structure (e.g., [McCully et al., 2017](#); [D’Aloisio and Natarajan, 2011](#); [Jullo et al., 2010](#)) and by massive substructure on the cluster outskirts ([Acebron et al., 2017](#); [Mahler et al., 2017](#)). The cosmological parameter degeneracies of this technique are entirely distinct from those of other cluster methods and [Gilmore and Natarajan \(2009\)](#) and [D’Aloisio and Natarajan \(2011\)](#) show that combining results from several clusters with many multiple image families in each yields a powerful probe of DE. This, however, requires tightly constrained lens models for several observed clusters and the simultaneous availability of spectroscopic redshifts for many lensed background sources in each cluster. Therefore high resolution HST imaging in combination with spectroscopy from MUSE on the VLT would be an ideal combination to tackle such a highly challenging task, see chapter 5. However, it is not necessary to apply for telescope time, as extensive HST data, spectroscopic redshifts, and precise lens models are already available for the six HFF clusters. A big challenge for such a study is the availability of computing resources and software which are efficient enough to simultaneously model six SL clusters while at the same time varying the cosmology to find the best fitting cosmological parameters. Indeed, the computational demands were so prohibitive that [Jullo et al. \(2010\)](#) could not use the more precise lens plane χ^2 computation and had to calculate it in the source plane. These performance challenges should be solved in the near future by LENSTOOL-HPC, see chapter 6.

3.4 The road ahead: Precision cosmology from lensing and its challenges

The previous section clearly demonstrates that gravitational lensing is a very versatile tool for precision cosmology. We now take a look at exciting recent developments in the field to illustrate two major points. First, it is not completely unreasonable to expect strong hints to new physics from gravitational lensing in the near future. Second, this will only be feasible if we carefully study and understand systematic errors in theory, simulation, and observation.

Λ CDM predicts a large number of DM halos and subhalos. The halo and subhalo mass functions can be calculated from N-body simulations and they will be different for CDM, WDM, and self-interacting DM (see e.g. [Mao et al., 2017](#)). The mass function can in principle be directly measured from SL and WL. In a recent publication, [Jauzac et al. \(2016a\)](#) identify seven massive subhalos within a radius of 1 Mpc around the center of the HFF cluster Abell 2744, which has a high-precision lensing mass model ([Jauzac et al., 2015](#)). In a follow-up paper, [Schwinn et al. \(2017\)](#) study simulated clusters in the Millennium XXL simulation and they cannot find a simulated cluster with this much massive substructure. As a result, the authors suggest that Abell 2744 might be a challenge to Λ CDM and lead to new cosmological insights. However, [Natarajan et al. \(2017\)](#) compare the subhalo mass functions of three HFF clusters including Abell 2744 from gravitational lensing with hydrodynamical galaxy cluster simulations from the Illustris Project ([Vogelsberger et al., 2014](#)) and claim an excellent agreement of both the amplitude and shape of the subhalo mass function over four decades in subhalo mass ($10^9 - 10^{13} M_{\odot}$). A possible solution which could reconcile these results was found by [Mao et al. \(2017\)](#). They show that the discrepancy with Λ CDM can be resolved by accounting for the fact that the projected mass measured inside an aperture is not always equal to the actual subhalo mass. Indeed, it can be greatly amplified by mass in the body of the galaxy cluster which happens to be projected along the line-of-sight to the subhalo.

Precise cosmic shear measurements from CFHTLenS (e.g., [Joudaki et al., 2017](#)) and KiDS (e.g., [Hildebrandt et al., 2017](#)) result in cosmological constraints which seem to be in tension with the Λ CDM cosmology values found by the Planck mission ([Planck Collaboration et al., 2016](#)). This is particularly interesting since Planck probes the CMB and thus the early universe at high redshift, while the two cosmic shear surveys probe the late, low- z universe. In addition, [Riess et al. \(2016\)](#) measure a local value of H_0 which is in tension with the findings of Planck for the Hubble constant. There are several possible explanations for these discrepancies, such as internal inconsistencies in the cosmic shear surveys (which are e.g. likely present in KiDS, see [Efstathiou and Lemos \(2017\)](#)) and internal issues in the Planck data (as for example visible in the tensions between cosmological constraints inferred from low-multipole data of the CMB temperature anisotropy power spectrum and those derived from high-multipole data, see [Addison et al. \(2016\)](#)). These have to be carefully studied and accounted for. In addition, the reported tensions are mostly at the $2-3\sigma$ level and the newly released DES Year 1 cosmology results, which are based on galaxy clustering and WL as probes of the low redshift universe,

3.4. The road ahead: Precision cosmology from lensing and its challenges

are consistent with Λ CDM and Planck (DES Collaboration et al., 2017). Nevertheless, these tensions warrant a detailed analysis and if they should still remain after accounting for all systematics, then this might possibly be a first hint at new physics beyond the cosmology standard model.

These recent developments show that we are living in a very interesting time for observational cosmology. In addition, the upcoming next generation of telescopes and surveys like JWST, Euclid, and LSST hold a lot of promise for new insights and discoveries. This, however, requires good knowledge and control of systematics, high precision measurements, a large amount of data to beat statistical errors, and the necessary computing architecture to deliver these results in a finite amount of time. In this thesis, we present contributions to several areas, namely a new theoretical method to break the mass-sheet degeneracy (chapter 4), a software pipeline for the extraction of spectroscopic redshifts from MUSE (chapter 5), an improved cluster lens model from HST and MUSE data (chapter 5), contributions to a High Performance Computing (HPC) version of the lens modeling software LENS`TOOL` (chapter 6), and a measurement pipeline for weak lensing flexion and preliminary application to a HFF cluster (chapter 7).

4 The mass-sheet degeneracy: A limiting factor?

4.1 Theoretical foundation

The mass sheet-degeneracy arises due to the fact that we cannot measure the shear γ and the two flexions \mathcal{F} and \mathcal{G} directly. The quantities which we can measure in observations of lensed images are their sizes and ellipticities, but since we do not know their intrinsic, unlensed sizes and shapes, this is of little immediate value. However, we can assume that galaxies are intrinsically randomly oriented and thus the ellipticity induced by the lensing effect can be inferred by averaging over a set of close-by galaxies (Bartelmann and Schneider, 2001). This is the most commonly used WL technique. In addition, lensing conserves surface brightness and hence the magnification effect increases the size of galaxy images at a fixed surface brightness level. This allows us to observe otherwise unobservably faint sources and thus the local number density of galaxies above a given flux threshold is modified by lensing and can be measured (Bartelmann and Schneider, 2001).

However, the signal-to-noise for the number counts technique is much worse than the one for the ellipticity method. We illustrate this with the following computation presented in Schneider et al. (2000). The local cumulative number counts density $n(\boldsymbol{\theta}, S)$ above a flux limit S is related to the unlensed number counts density $n_0(S)$ as

$$n(\boldsymbol{\theta}, S) = \frac{1}{\mu} n_0\left(\frac{S}{\mu}\right). \quad (4.1)$$

We assume that the counts follow locally a power law of the form

$$n_0 \propto S^{-b} \quad (4.2)$$

and this yields (Schneider et al., 2000)

$$n(\boldsymbol{\theta}) = n_0 \mu^{b-1} \quad (4.3)$$

Chapter 4. The mass-sheet degeneracy: A limiting factor?

for any fixed flux threshold. If the intrinsic counts are flatter than one, then lensed counts will be reduced, and this depletion effect is the signature of lensing which we want to measure (Schneider et al., 2000). We assume that the magnification μ is constant over a solid angle and we expect N_μ galaxies in this area in the absence of lensing depletion. As a result, the magnification signal is (Schneider et al., 2000)

$$|\Delta N| = |\mu^{b-1} - 1|N_\mu \quad (4.4)$$

and the noise is approximately the Poisson noise estimate $\sqrt{N_\mu}$. Consequently, we have (Schneider et al., 2000)

$$\left(\frac{S}{N}\right)_\mu = |\mu^{b-1} - 1|\sqrt{N_\mu}. \quad (4.5)$$

Now we look at the ellipticity measurement method. We assume that the shear γ is constant over a solid angle, that $\kappa \ll 1$ and that there are N_γ background galaxies for which we have measured the ellipticity. The noise is given by $\sigma_\gamma/\sqrt{N_\gamma}$, where σ_γ is the ellipticity dispersion (Schneider et al., 2000). As a result, the signal-to-noise is (Schneider et al., 2000)

$$\left(\frac{S}{N}\right)_\gamma = \frac{|\gamma|\sqrt{N_\gamma}}{\sigma_\gamma}. \quad (4.6)$$

Since we assumed that κ is small, we have

$$\mu \approx 1 + 2\kappa \quad (4.7)$$

and a first order expansion for $|\Delta N_\mu|$ gives (Schneider et al., 2000)

$$|\Delta N_\mu| \approx 2\kappa|1 - b|N_\mu. \quad (4.8)$$

With these results, the ratio of the signal-to-noises is (Schneider et al., 2000)

$$\frac{\left(\frac{S}{N}\right)_\gamma}{\left(\frac{S}{N}\right)_\mu} = \frac{|\gamma|}{\kappa} \frac{1}{2\sigma_\gamma|1 - b|} \sqrt{\frac{N_\gamma}{N_\mu}}. \quad (4.9)$$

For an order-of-magnitude comparison, we can assume that the magnitudes of shear γ and convergence κ are similar, that we have $b \geq 0.5$, that the ellipticity dispersion is $\sigma_\gamma \approx 0.25$, and that $N_\gamma \approx N_\mu/3$, since we can still count galaxies which are too faint to measure their

ellipticity (e.g., [Schneider et al., 2000](#); [Leauthaud et al., 2007](#)). As a result, we have

$$\frac{\left(\frac{s}{N}\right)_\gamma}{\left(\frac{s}{N}\right)_\mu} \approx 2. \quad (4.10)$$

Note that this is actually an underestimate, since real galaxies cluster and hence the number counts error will be larger in practice (e.g., [Bartelmann and Schneider, 2001](#)). Thus the ellipticity measurement is typically the preferred WL technique.

Let us now investigate why the averaged ellipticity does not directly yield the lensing shear γ . This can be seen by rewriting equation 3.23 as ([Bartelmann and Schneider, 2001](#))

$$\mathbf{A} = (1 - \kappa) \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix}, \quad (4.11)$$

where we have used the reduced shear g , which is defined as

$$g(\boldsymbol{\theta}) = \frac{\gamma(\boldsymbol{\theta})}{1 - \kappa(\boldsymbol{\theta})}. \quad (4.12)$$

The prefactor $(1 - \kappa)$ in equation 4.11 does only affect the size of the images and not their shapes. Thus we can only measure the reduced shear g from the observed ellipticity in astronomical data.

We have shown that we cannot directly measure the shear γ and thus the logical next step is to investigate whether we can determine the flexions \mathcal{F} and \mathcal{G} from observational data. Since they are functions of the shear derivatives, the answer is no, and the proof is as follows:

$$\begin{aligned} \mathcal{F} &= \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix} = (1 - \kappa) \begin{pmatrix} g_{1,1} + g_{2,2} \\ g_{2,1} - g_{1,2} \end{pmatrix}, \\ \mathcal{G} &= \begin{pmatrix} \gamma_{1,1} - \gamma_{2,2} \\ \gamma_{2,1} + \gamma_{1,2} \end{pmatrix} = (1 - \kappa) \begin{pmatrix} g_{1,1} - g_{2,2} \\ g_{2,1} + g_{1,2} \end{pmatrix}, \end{aligned} \quad (4.13)$$

and we see that we can only obtain the reduced flexions F and G from our data, which we define as ([Okura et al., 2008](#))

$$\begin{aligned} F &= \frac{\mathcal{F}}{1 - \kappa}, \\ G &= \frac{\mathcal{G}}{1 - \kappa}. \end{aligned} \quad (4.14)$$

Chapter 4. The mass-sheet degeneracy: A limiting factor?

Alternatively, we can use complex notation, $\mathcal{F} = \mathcal{F}_1 + i\mathcal{F}_2$, and define the reduced flexions as follows (Schneider and Er, 2008)

$$\begin{aligned} G_1 &= \frac{\mathcal{F} + g\mathcal{F}^*}{1 - \kappa}, \\ G_3 &= \frac{\mathcal{G} + g\mathcal{F}}{1 - \kappa}, \end{aligned} \quad (4.15)$$

where the asterisk denotes complex conjugation. However, we will not use the latter definition in this thesis and instead stick to the former.

Now we may wonder: Does the fact that we can only measure the reduced quantities actually limit our scientific objectives or is this simply a nuisance? Unfortunately, the answer is that it limits our ability to investigate the lens in a significant way. Let us look at the deflection potential Ψ and transform it by using a constant $\lambda \in \mathbb{R}$ as

$$\Psi(\boldsymbol{\theta}) \rightarrow \Psi'(\boldsymbol{\theta}) = \frac{1 - \lambda}{2} \boldsymbol{\theta}^2 + \lambda \Psi(\boldsymbol{\theta}). \quad (4.16)$$

As a result, the non-reduced lensing quantities transform in the following way,

$$\begin{aligned} \kappa &\rightarrow \kappa' = \lambda\kappa + (1 - \lambda), \\ \gamma &\rightarrow \gamma' = \lambda\gamma, \\ \mathcal{F} &\rightarrow \mathcal{F}' = \lambda\mathcal{F}, \\ \mathcal{G} &\rightarrow \mathcal{G}' = \lambda\mathcal{G}, \\ \mu &\rightarrow \mu' = \frac{\mu}{\lambda^2}, \end{aligned} \quad (4.17)$$

but the reduced quantities are invariant under such a family of transformations,

$$\begin{aligned} g &\rightarrow g' = g, \\ F &\rightarrow F' = F, \\ G &\rightarrow G' = G. \end{aligned} \quad (4.18)$$

Therefore two different surface mass distributions which differ only by a λ -transformation cannot be distinguished by measuring the image distortions. Note that in the limit of λ almost equal to unity, equation 4.17 amounts to adding a constant mass-sheet κ_0 to κ , hence the name mass-sheet degeneracy.

4.2 Impact on lens models and proposed techniques to lift the degeneracy

In principle, the mass-sheet degeneracy could be lifted by requiring that the lensing mass distribution is completely inside our field-of-view, and thus $\kappa = 0$ at the boundary of the image

4.2. Impact on lens models and proposed techniques to lift the degeneracy

(e.g., Bradač et al., 2004). However, even for wide-field cameras this leads to a substantial underestimate of the cluster mass, as cluster density profiles appear to be falling smoothly even out to large radius. If we look at a cluster with virial mass $M_{\text{vir}} = 10^{15} M_{\odot}$ at $z = 0.2$, we expect from N-body simulations that $\kappa \approx 0.005$ at a separation of 15 arcmin from the cluster center (Bradač et al., 2004). Even with a camera with 30 arcmin \times 30 arcmin field-of-view, we would underestimate the virial mass by $\sim 20\%$ if we simply set κ to zero at the boundary (Bradač et al., 2004). Therefore more sophisticated techniques are required.

Over time a few methods have been developed which can lift the mass-sheet degeneracy. In 1995, Broadhurst et al. (1995) propose to use the magnification μ by comparing lensed and unlensed background source counts to reconstruct the non-degenerate mass distribution. This magnification effect is subsequently observed and applied (see e.g. Fort et al., 1997; Taylor et al., 1998; Dye et al., 2002). In the same year, Bartelmann and Narayan (1995) describe a technique which is called “lens parallax method”. Its principle is as follows: The mean sizes of lensed faint blue galaxies are compared with those of unlensed sources in an “empty” field. As previously described, lensing preserves the surface brightness, and thus the magnification μ can be computed by comparing the mean sizes of lensed and unlensed galaxies with the same surface brightness. This information can subsequently be combined with the shear γ to break the mass-sheet degeneracy for the lens model. However, the measurement of the surface brightness from seeing-convolved images is difficult and thus the application of this method is a challenge (Schneider et al., 2006). It is also possible to break the degeneracy by fitting the gravitational lens potential to shear γ and magnification μ data simultaneously. Several fitting and Bayesian methods have been proposed for this purpose (e.g., Umetsu et al., 2011; Schneider et al., 2000; Bartelmann et al., 1996). In 2004, Bradač et al. (2004) demonstrate that it is possible to reconstruct a non-degenerate mass distribution if the individual redshifts of sources at different distances are combined with the shear field. However, this method is only applicable to critical lenses. A different technique is required for sub-critical lenses and when a sufficient number of redshifts is not available.

In the publication which is reprinted in the next section, we propose a new technique which can lift the mass-sheet degeneracy by a simple comparison of aperture masses. We show that taking the difference in mass estimates from convergence information, derived for example from source counts, and shear or flexion data allows us to compute the value of the mass-sheet degeneracy parameter λ . As a result, we can calibrate the masses inferred from gravitational lensing shear or flexion measurements. We demonstrate the feasibility of this technique for both stacked galaxy-galaxy lensing and for lensing by galaxy clusters.

In addition, we also show in the publication for two widely used lens models that the signal-to-noise of shear and flexion in the WL regime is always superior to the number counts, both on galaxy and cluster scales. Consequently we can use the convergence information to lift the mass-sheet degeneracy without sacrificing the quality of the lens reconstruction.

4.3 A new method to break the degeneracy

The following paper, [Rexroth et al. \(2016\)](#), is reprinted in this thesis with the generous permission from the Monthly Notices of the Royal Astronomical Society (E-Mail communication, 15 September 2017).

A new method to break the mass-sheet degeneracy using aperture moments

Markus Rexroth,¹★ Priyamvada Natarajan² and Jean-Paul Kneib^{1,3}

¹Laboratoire d'Astrophysique, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland

²Department of Astronomy, Yale University, 260 Whitney Avenue, New Haven, CT 06511, USA

³Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, F-13388, Marseille, France

Accepted 2016 April 26. Received 2016 April 26; in original form 2015 December 22

ABSTRACT

Mass determinations from gravitational lensing shear and the higher order estimator flexion are both subject to the mass-sheet degeneracy. Mass sheet degeneracy refers to a transformation that leaves the reduced shear and flexion invariant. In general, this transformation can be approximated by the addition of a constant surface mass density sheet. We propose a new technique to break the mass-sheet degeneracy. The method uses mass moments of the shear or flexion fields in combination with convergence information derived from number counts which exploit the magnification bias. The difference between the measured mass moments provides an estimator for the magnitude of the additive constant that is the mass sheet. For demonstrating this, we derive relations that hold true in general for n th order moments and show how they can be employed effectively to break the degeneracy. We investigate the detectability of this degeneracy parameter from our method and find that the degeneracy parameter can be feasibly determined from stacked galaxy–galaxy lensing data and cluster lensing data. Furthermore, we compare the signal-to-noise ratios of convergence information from number counts with shear and flexion for singular isothermal sphere and Navarro–Frenk–White models. We find that the combination of shear and flexion performs best on galaxy and cluster scales and the convergence information can therefore be used to break the mass-sheet degeneracy without quality loss in the mass reconstruction. In summary, there is power in the combination of shear, flexion, convergence and their higher order moments. With the anticipated wealth of lensing data from upcoming and future satellite missions – *EUCLID* and *WFIRST* – this technique will be feasible.

Key words: gravitational lensing; weak – galaxies: clusters: general – galaxies: haloes – dark matter.

1 INTRODUCTION

Many mass reconstructions were successfully obtained utilizing gravitational lensing shear (e.g. Natarajan et al. 1998; Hoekstra, Yee & Gladders 2004) and the feasibility of lens models from flexion has been demonstrated (Leonard et al. 2007; Okura, Umetsu & Futamase 2008). Several mass reconstruction methods have been developed (e.g. Kaiser & Squires 1993; Kaiser 1995; Bartelmann et al. 1996; Schneider, King & Erben 2000); see also Kneib & Natarajan (2011) for a review of galaxy cluster lens reconstructions. However, the masses derived from the shear and flexion fields are only determined up to a constant due to the mass sheet degeneracy (e.g. Schneider & Seitz 1995; Schneider & Er 2008). Several techniques have been proposed to lift the degeneracy. Broadhurst,

Taylor & Peacock (1995) suggested using magnification information by comparing lensed and unlensed background source counts to reconstruct the non-degenerate mass sheet. This effect was subsequently observed and applied (e.g. Fort, Mellier & Dantel-Fort 1997; Taylor et al. 1998; Dye et al. 2002). Bartelmann & Narayan (1995) proposed the lens parallax method. This method compares the mean sizes of lensed faint blue galaxies with those of unlensed sources in an empty control field. While gravitational lensing magnifies the area, it preserves the surface brightness, and thus the magnification can be inferred by comparing the mean sizes of lensed and unlensed galaxies with the same surface brightness. The magnification information can then be combined with the shear measurement to break the degeneracy. However, it is hard to estimate the surface brightness from seeing-convolved images and thus the application of this technique is challenging (Meylan, Jetzer & North 2006). In addition, the degeneracy can also be lifted by fitting the gravitational lens potential to shear and magnification data simultaneously. Several

* E-mail: markus.rexroth@epfl.ch

fitting and Bayesian methods have been proposed for this purpose (see e.g. Bartelmann et al. 1996; Schneider et al. 2000; Umetsu et al. 2011). Bradač, Lombardi & Schneider (2004) showed that it is possible to reconstruct non-degenerate mass distributions if the individual redshifts of sources at different distances are combined with the shear field. However, this technique can only be used for critical lenses. A different method is needed for sub-critical lenses and lenses where a sufficient number of redshifts is not available. In this paper, we present, develop, and explicate the technical framework that can be used to derive the mass-sheet degeneracy using combinations of shape and convergence estimator moments.

We propose a new method that can break the mass-sheet degeneracy by a simple comparison of aperture masses. We show that taking the difference in mass estimates from convergence information, derived for example from source counts, and shear or flexion data allows the determination of the value of the mass sheet, thus enabling the derivation of calibrated masses. We demonstrate its feasibility for both stacked galaxy–galaxy lensing and cluster lensing. Furthermore, we show that for two widely used mass models, the signal-to-noise ratio of shear and flexion in the weak lensing regime is always superior to the number counts, on both galaxy and cluster scales. Therefore we can use the convergence information to break the mass sheet without sacrificing the quality of the lens reconstruction.

We develop the theoretical framework of this method for both shear and flexion, the higher order lensing effect which describes the arciness of the lensed sources. Flexion was introduced into the weak lensing framework, because it provides valuable additional information and is highly sensitive to substructure (Goldberg & Natarajan 2002; Goldberg & Bacon 2005; Bacon et al. 2006; Irwin & Shmakova 2006). Studies have showed that the addition of flexion can improve mass reconstructions significantly (e.g. Leonard et al. 2007; Okura, Umetsu & Futamase 2007). The high-quality data from the *Hubble Frontier Fields* initiative permit flexion measurements with increased accuracy (Rexroth 2015).

The paper is organized as follows. In Section 2, we give a brief description of the weak lensing formalism and the mass-sheet degeneracy. Section 3 presents the equivalent moments and how their transformations can break the degeneracy. We demonstrate this method for a singular isothermal sphere (SIS) halo model in Section 4. In Section 5 we investigate the detectability of the mass-sheet parameter without making assumptions on the halo model. Section 6 shows that the combination of shear and flexion has a higher signal-to-noise ratio than number counts for SIS and Navarro–Frenk–White (NFW) haloes. We conclude in Section 7.

2 WEAK LENSING FORMALISM AND THE MASS-SHEET DEGENERACY

Weak lensing of background sources by massive foreground objects results in the deformation of their shapes. The strength of these distortions is directly related to the surface mass density of the lens. Let us consider a lens with projected surface mass density $\Sigma(\boldsymbol{\theta})$, where the angular coordinate $\boldsymbol{\theta}$ denotes the position in the lens plane. The convergence of the lens κ is defined as

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\text{crit}}}, \quad (1)$$

where the critical surface mass density, defined as

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \left(\frac{D_{\text{os}}}{D_{\text{ls}} D_{\text{ol}}} \right), \quad (2)$$

depends on the angular diameter distances from the lens to the source D_{ls} , observer to source D_{os} , and observer to lens D_{ol} . The lensed and unlensed coordinates for the distorted background sources are related by the mapping

$$A_{ij}(\boldsymbol{\theta}) = \frac{\partial \theta'_i}{\partial \theta_j}, \quad (3)$$

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (4)$$

where $\boldsymbol{\theta}'$ are the unlensed coordinates and γ is the shear. The magnification of the sources is

$$\mu = \frac{1}{(1 - \kappa)^2 - \gamma^2}. \quad (5)$$

Furthermore, we can express the convergence and shear using the deflection potential

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\beta \kappa(\boldsymbol{\beta}) \ln |\boldsymbol{\theta} - \boldsymbol{\beta}| \quad (6)$$

as

$$\kappa = \frac{1}{2}(\psi_{,11} + \psi_{,22}), \quad \gamma = \begin{pmatrix} \frac{1}{2}(\psi_{,11} - \psi_{,22}) \\ \psi_{,12} \end{pmatrix}. \quad (7)$$

The commas denote partial θ derivatives. The reduced shear is defined as

$$g = \frac{\gamma}{1 - \kappa}. \quad (8)$$

If shear and convergence are constant over a source image, the transformation is given by a simple matrix multiplication:

$$\theta'_i = A_{ij} \theta_j. \quad (9)$$

However, if the shear varies over the image, we have to expand to include flexion terms:

$$\theta'_i \simeq A_{ij} \theta_j + \frac{1}{2} D_{ijk} \theta_j \theta_k, \quad (10)$$

where $D_{ijk} = A_{ij,k}$. The first and second flexion are given in terms of derivatives of the two shear components (as per the notation in Bacon et al. 2006)

$$\mathcal{F} = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}, \quad \mathcal{G} = \begin{pmatrix} \gamma_{1,1} - \gamma_{2,2} \\ \gamma_{2,1} + \gamma_{1,2} \end{pmatrix}. \quad (11)$$

The mass-sheet degeneracy arises due to the fact that the observed quantities are not the shear and flexions themselves, but the reduced shear and the reduced first and second flexion. The latter two can be compactly written if we use a complex notation, i.e. $\mathcal{F} = \mathcal{F}_1 + i\mathcal{F}_2$,

$$G_1 = \frac{\mathcal{F} + g\mathcal{F}^*}{1 - \kappa}, \quad G_3 = \frac{\mathcal{G} + g\mathcal{F}}{1 - \kappa}, \quad (12)$$

where the asterisk denotes complex conjugation (Schneider & Er 2008). Alternatively, the reduced flexions can be defined as (Okura et al. 2008)

$$F = \frac{\mathcal{F}}{1 - \kappa}, \quad G = \frac{\mathcal{G}}{1 - \kappa}, \quad (13)$$

and we will use this definition in this paper. If the deflection potential is transformed for any constant λ as

$$\psi(\boldsymbol{\theta}) \rightarrow \psi'(\boldsymbol{\theta}) = \frac{1 - \lambda}{2} \boldsymbol{\theta}^2 + \lambda \psi(\boldsymbol{\theta}), \quad (14)$$

we have

$$\begin{aligned} \kappa &\rightarrow \kappa' = \lambda\kappa + (1 - \lambda), \gamma \rightarrow \gamma' = \lambda\gamma, \\ \mathcal{F} &\rightarrow \mathcal{F}' = \lambda\mathcal{F}, \mathcal{G} \rightarrow \mathcal{G}' = \lambda\mathcal{G}, \mu \rightarrow \mu' = \frac{\mu}{\lambda^2}, \end{aligned} \quad (15)$$

but the reduced shear and reduced flexions are invariant under such a family of transformations,

$$g \rightarrow g' = g, F \rightarrow F' = F, G \rightarrow G' = G. \quad (16)$$

Therefore, two different surface mass distributions that differ by λ cannot be distinguished by measuring only the image distortions. Note that in the limit of λ almost equal to unity, equation (15) amounts to adding a constant mass sheet κ_0 to κ .

In principle, the mass-sheet degeneracy could be lifted by requiring that $\kappa = 0$ at the boundary of the image. However, even for wide-field cameras this leads to a substantial underestimate of the cluster mass as cluster density profiles appear to be falling smoothly even out to large radius. For a cluster with virial mass $M_{\text{vir}} = 10^{15} M_{\odot}$ at redshift $z = 0.2$, we expect from N -body simulations that $\kappa \approx 0.005$ at 15 arcmin from the cluster centre and with a 30×30 arcmin² camera field of view, we would underestimate the virial mass by ~ 20 per cent if we simply set κ to zero at the boundary (Bradač et al. 2004). Therefore more sophisticated techniques are required.

3 PROPOSED METHOD: MOMENTS OF THE CONVERGENCE, SHEAR AND FLEXION FIELDS

Our goal is to derive a method that is independent of spectroscopic redshifts and broadly applicable to both critical and sub-critical lenses. To develop such a method, we first investigate n th order moments of the shear, convergence, and flexions and study their transformation properties with a view to understand the explicit dependence on λ . Schneider & Bartelmann (1997) (hereafter SB97) show the equivalence of shear and convergence mass moments $M^{(n)}$ and multipole moments $Q^{(n)}$. We extend these equivalences to mass and multipole moments using flexion. Subsequently, we show that the mass-sheet degeneracy transformations destroy these equivalences, as they give rise to different surface terms. These permit us to calculate the mass-sheet degeneracy parameter λ .

Our method comes in two variants. One uses the reconstructed, unreduced shear or flexion fields. We can obtain these by multiplying the reduced quantities with $1 - \kappa_{\text{rec}}$, where κ_{rec} is the convergence of the reconstructed mass sheet. Many techniques for obtaining κ_{rec} from shear or flexion data have been successfully developed and Er, Li & Schneider (2010) present a method that reconstructs the mass sheet by combining the two. The second variant uses the reduced shear and flexion fields, thus it does not require the reconstructed convergence.

3.1 Mass and multipole moments of shear, flexion, and convergence

The detailed derivation of the different moments and the proof of their equivalences are shown in Appendix A1. We do not explicitly denote (x_0, φ_0) or (x, φ) dependencies of $\kappa, \kappa_{\text{rec}}, \gamma, \mathcal{F}$, and \mathcal{G} to keep the notation compact. The resulting equivalent moments

are:

$$M_{\kappa}^{(n)} = \int_0^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi \kappa, \quad (17)$$

$$M_{\mathcal{F}}^{(n)} = \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi (1 - \kappa_{\text{rec}}) F_{\text{t}}, \quad (18)$$

$$\begin{aligned} M_{\gamma}^{(n)} &= \int_0^{\infty} dx \left[2W(x) - x^{n+1} w(x) \right] \\ &\times \int_0^{2\pi} d\varphi (1 - \kappa_{\text{rec}}) g_{\text{t}}, \end{aligned} \quad (19)$$

$$\begin{aligned} M_{\mathcal{F}, \mathcal{G}}^{(n)} &= - \int_0^{\infty} dx \left[V(x) - \frac{1}{2} x W(x) \right] \\ &\times \int_0^{2\pi} d\varphi \left[(1 - \kappa_{\text{rec}}) F_{\text{t}} + (1 - \kappa_{\text{rec}}) G_{\text{t}} \right]. \end{aligned} \quad (20)$$

The formula for $M_{\mathcal{F}}^{(n)}$ in the case that $n = 0$ was also derived in Leonard, King & Wilkins (2009), but with stronger assumptions on the weighting function. We used the definitions

$$x W(x) = \int_0^x dy y^{n+1} w(y), \quad (21)$$

$$V(x) = \int_0^x dy W(y), \quad (22)$$

and the reduced tangential and radial shear and flexions,

$$g_{\text{t}} = -[g_1(x, \varphi) \cos(2\varphi) + g_2(x, \varphi) \sin(2\varphi)], \quad (23)$$

$$g_{\text{r}} = -[g_2(x, \varphi) \cos(2\varphi) - g_1(x, \varphi) \sin(2\varphi)], \quad (24)$$

$$F_{\text{t}} = -[F_1(x, \varphi) \cos(\varphi) + F_2(x, \varphi) \sin(\varphi)], \quad (25)$$

$$F_{\text{r}} = -[F_2(x, \varphi) \cos(\varphi) - F_1(x, \varphi) \sin(\varphi)], \quad (26)$$

$$G_{\text{t}} = -[G_1(x, \varphi) \cos(3\varphi) + G_2(x, \varphi) \sin(3\varphi)], \quad (27)$$

$$G_{\text{r}} = -[G_2(x, \varphi) \cos(3\varphi) - G_1(x, \varphi) \sin(3\varphi)]. \quad (28)$$

We require that the boundary term $x W(x) \kappa(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$. The derivation of M_{γ} demands in addition that the surface terms $x W(x) \gamma_{\text{t}}(x, \varphi)$ vanishes for the same limits and the derivation of $M_{\mathcal{F}, \mathcal{G}}$ demands in addition to the former two that $[2V(x) - x W(x)] \gamma_{\text{t}}(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$. This can be achieved by using a suitable weighting function.

Analogously, we derived the equivalences between the multipole moments

$$Q_{\kappa}^{(n)} = \int_0^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} \kappa, \quad (29)$$

$$Q_{\mathcal{F}}^{(n)} = \int_0^{\infty} dx x W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}) F_{\text{t}}, \quad (30)$$

$$\begin{aligned} Q_{\gamma}^{(n)} &= \int_0^{\infty} dx \left[2W(x) - x^{n+1} w(x) \right] \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}) g_{\text{t}} \\ &- in \int_0^{\infty} dx W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}) g_{\text{r}}, \end{aligned} \quad (31)$$

$$\begin{aligned}
 Q_{\mathcal{F},\mathcal{G}}^{(n)} = & - \int_0^\infty dx \left[V(x) - \frac{1}{2}x W(x) \right] \int_0^{2\pi} d\varphi \\
 & e^{in\varphi} \left[(1 - \kappa_{\text{rec}})F_t + (1 - \kappa_{\text{rec}})G_t \right] \\
 & + i \frac{n}{2} \int_0^\infty dx V(x) \int_0^{2\pi} d\varphi e^{in\varphi} \\
 & \left[(1 - \kappa_{\text{rec}})F_r + (1 - \kappa_{\text{rec}})G_r \right]. \quad (32)
 \end{aligned}$$

Again we require that the term $xW(x)\kappa(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$. The derivation of Q_γ demands in addition that the surface term $xW(x)\gamma_t(x, \varphi)$ vanishes and the derivation of $Q_{\mathcal{F},\mathcal{G}}$ demands in addition to the former two that $[2V(x) - xW(x)]\gamma_t(x, \varphi)$ and $V(x)\gamma_r(x, \varphi)$ vanish for $x \rightarrow 0, \infty$.

In some cases, it is useful to compute the moments only on a ring which excludes the central part of the field of view, for example if we want to exclude the strong lensing area. Therefore we extended the identities shown in this section to rings. The results are given in Appendix A1.1.

3.2 Mass and multipole moments of reduced shear, reduced flexion, and K

The derivation of the moments and their equivalences is given in Appendix B. The equivalent reduced moments are:

$$M_K^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi K, \quad (33)$$

$$M_F^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi F_t, \quad (34)$$

$$\begin{aligned}
 M_g^{(n)} = & \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_t F_t + g_r F_r) \\
 & + \int_R^\infty dx (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} d\varphi g_t, \quad (35)
 \end{aligned}$$

$$\begin{aligned}
 M_{F,G}^{(n)} = & \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_r F_r + \frac{1}{2}(F_t + G_t)) \\
 & - \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi (F_t + G_t - 2g_t F_t). \quad (36)
 \end{aligned}$$

We defined analogous to Cain, Schechter & Bautz (2011)

$$K = -\ln(1 - \kappa) \quad (37)$$

and in addition

$$xW_R(x) = \int_R^x dy y^{n+1} w(y), \quad (38)$$

$$V_R(x) = \int_R^x dy W_R(y). \quad (39)$$

We choose the lower integral limit for our ring, R , such that $\kappa < 1 \forall x \geq R, \forall \varphi$. Thus K is well-defined and finite. We require that $xW_R(x)K(x, \varphi)$ vanishes for $x \rightarrow R, \infty$. The M_g moment requires in addition that $xW_R(x)g_t(x, \varphi)$ vanishes for $x \rightarrow R, \infty$ and the $M_{F,G}$ moment demands that additionally to the two requirements also the surface term $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ vanishes for the same limits.

For the multipole moments, we have the following equivalent moments:

$$Q_K^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} K, \quad (40)$$

$$Q_F^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} F_t, \quad (41)$$

$$\begin{aligned}
 Q_g^{(n)} = & \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t F_t + g_r F_r) \\
 & + \int_R^\infty dx (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} d\varphi e^{in\varphi} g_t \\
 & - in \int_R^\infty dx W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} g_r, \quad (42)
 \end{aligned}$$

$$\begin{aligned}
 Q_{F,G}^{(n)} = & \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t F_t + g_r F_r) \\
 & - \int_R^\infty dx (2V_R(x) - xW_R(x)) \int_0^{2\pi} d\varphi e^{in\varphi} \\
 & \times \left(\frac{1}{2}(F_t + G_t) - g_t F_t \right) \\
 & + in \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} \left(\frac{1}{2}(F_t + G_t) - g_t F_t \right). \quad (43)
 \end{aligned}$$

We have the same requirement on R as for the mass moments and we demand that the resulting surface term $xW_R(x)K(x, \varphi)$ vanishes for $x \rightarrow R, \infty$. For the Q_g moment we require additionally that $xW_R(x)g_t(x, \varphi)$ vanishes for the same limits. The $Q_{F,G}$ moment requires in addition to the two that the terms $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ and $V_R(x)g_r(x, \varphi)$ vanish for $x \rightarrow R$ and $x \rightarrow \infty$.

We note that the g moments require additional terms involving the flexion, which is not the case for the γ moments. Similarly, the F, G moments require shear information, which is not required for the \mathcal{F}, \mathcal{G} moments.

3.3 Mass moment transformations

3.3.1 Mass moments of shear, flexion, and convergence

The mass-sheet degeneracy transformations in equation (15) destroy the mass moment equivalences shown in the previous section. They hold only for $\lambda = 1$. The n th order mass moment of the convergence for the transformed field κ' is

$$M_{\kappa'}^{(n)} = \int_0^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi (\lambda\kappa + (1 - \lambda)) \quad (44)$$

and can be written explicitly in terms of the corresponding n th order mass moments of κ ,

$$M_{\kappa'}^{(n)} = \lambda M_\kappa^{(n)} + 2\pi(1 - \lambda) \int_0^\infty dx x^{n+1} w(x). \quad (45)$$

The integral on the right-hand side of the equation above is a Γ function for the choice of a Gaussian window function,

$$w(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad (46)$$

therefore we have

$$M_{\kappa'}^{(n)} = \lambda M_{\kappa}^{(n)} + 2^{\frac{n+1}{2}} \sqrt{\pi} (1 - \lambda) \sigma^{n+2} \Gamma\left(1 + \frac{n}{2}\right). \quad (47)$$

Now let us examine the transformation properties for the mass moments computed using the shear field γ . In the limit that λ goes to unity, $M_{\kappa'}^{(n)}$ is equivalent to $M_{\gamma'}^{(n)}$. However, we show here that this is not the case in general, since the surface terms do not vanish identically. The moments for a transformed shear field $\gamma'_t = g_t(1 - \kappa'_{\text{rec}}) = \lambda g_t(1 - \kappa_{\text{rec}}) = \lambda \gamma_t$ are

$$M_{\gamma'}^{(n)} = \int_0^\infty dx [2W(x) - x^{n+1}w(x)] \int_0^{2\pi} d\varphi \gamma'_t = \lambda M_{\gamma}^{(n)}. \quad (48)$$

Similarly, the mass moments in terms of flexion and convergence are equivalent in the limit $\lambda \rightarrow 1$, but they differ in general. For the flexion fields $\mathcal{F}'_t = F_t(1 - \kappa'_{\text{rec}}) = \lambda \mathcal{F}_t$ and $\mathcal{G}'_t = G_t(1 - \kappa'_{\text{rec}}) = \lambda \mathcal{G}_t$ we have

$$M_{\mathcal{F}'}^{(n)} = \int_0^\infty dx x W(x) \int_0^{2\pi} d\varphi \mathcal{F}'_t = \lambda M_{\mathcal{F}}^{(n)}, \quad (49)$$

$$\begin{aligned} M_{\mathcal{F}', \mathcal{G}'}^{(n)} &= - \int_0^\infty dx \left[V(x) - \frac{1}{2} x W(x) \right] \int_0^{2\pi} d\varphi \left[\mathcal{F}'_t + \mathcal{G}'_t \right] \\ &= \lambda M_{\mathcal{F}, \mathcal{G}}^{(n)}. \end{aligned} \quad (50)$$

Therefore, by comparing the mass moments of a given order derived from the convergence field with that derived from the shear or flexion fields, λ can be estimated and hence the mass degeneracy can be lifted. A comparison of the shear and flexion mass moments cannot be used to infer λ , because they are equivalent for every λ .

As shown in Appendix A1, the requirements on the weighting function for the use of $M_{\mathcal{F}, \mathcal{G}}^{(n)}$ are stricter than those for $M_{\mathcal{F}}^{(n)}$. The application of the former demands in the case of a simple Gaussian weighting function that the tangential shear decreases faster than $1/\log(x)$ for $x \rightarrow \infty$ to ensure that equation (A31) holds, while the use of the latter does not require such constraints. Shear typically decreases as $1/x$ and both identities are therefore usable, but it can be advantageous to design a better suited weighting function.

3.3.2 Mass moments of reduced shear, reduced flexion, and K

The mass-sheet degeneracy transformations also break the equivalences of the mass moments of the reduced quantities. We have

$$K' = -\ln(1 - (\lambda\kappa + 1 - \lambda)) = -\ln(\lambda) - \ln(1 - \kappa) \quad (51)$$

and therefore the corresponding mass moment transforms as

$$M_{K'}^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi [-\ln(\lambda) - \ln(1 - \kappa)] \quad (52)$$

$$= M_K^{(n)} - 2\pi \ln(\lambda) \int_R^\infty dx x^{n+1} w(x). \quad (53)$$

For a Gaussian weighting function, this leads to

$$M_{K'}^{(n)} = M_K^{(n)} - 2^{\frac{n+1}{2}} \sqrt{\pi} \ln(\lambda) \sigma^{n+2} \Gamma\left(1 + \frac{n}{2}, \frac{R^2}{2\sigma^2}\right), \quad (54)$$

where $\Gamma(a, x)$ is the incomplete upper gamma function,

$$\Gamma(a, x) = \int_x^\infty dt t^{a-1} e^{-t}. \quad (55)$$

The remaining mass moments depend only on reduced quantities which do not change under this type of transformation, thus

we have

$$M_{g'}^{(n)} = M_g^{(n)}, \quad (56)$$

$$M_{F'}^{(n)} = M_F^{(n)}, \quad (57)$$

$$M_{F', G'}^{(n)} = M_{F, G}^{(n)}. \quad (58)$$

As we showed in the previous section, for $\lambda \rightarrow 1$ all moments of a given order are equal and we can therefore compare one of the latter three to the K mass moment to determine λ .

As seen in Appendix B, the requirements for the use of $M_{F, G}^{(n)}$ are again stricter than those for $M_F^{(n)}$. For a Gaussian weighting function, the reduced tangential shear has to decrease faster than $1/\log(x)$ for $x \rightarrow \infty$, which is typically the case. Designing a faster declining weight can avoid the requirement on the reduced shear.

3.4 Multipole moment transformations

3.4.1 Multipole moments of shear, flexion, and convergence

We show that the multipole moments are equivalent under the mass-sheet transformations for all λ . Therefore, they cannot be used to determine its value. This is due to the additional exponential factor $e^{in\varphi}$. For the following calculation, we will assume $n \geq 1$, as the multipole moments are identical to the mass moments for $n = 0$. We have

$$\begin{aligned} Q_{\kappa'}^{(n)} &= \int_0^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} [\lambda\kappa + (1 - \lambda)] \\ &= \lambda Q_{\kappa}^{(n)}. \end{aligned} \quad (59)$$

The shear and flexion multipole moments transform like their mass moments,

$$Q_{\gamma'}^{(n)} = \lambda Q_{\gamma}^{(n)}, \quad (60)$$

$$Q_{\mathcal{F}'}^{(n)} = \lambda Q_{\mathcal{F}}^{(n)}, \quad (61)$$

$$Q_{\mathcal{F}', \mathcal{G}'}^{(n)} = \lambda Q_{\mathcal{F}, \mathcal{G}}^{(n)}. \quad (62)$$

3.4.2 Multipole moments of reduced shear, reduced flexion, and K

The multipole moments obtained from the reduced fields are also equivalent for every choice of λ and thus they cannot be used to break the mass-sheet degeneracy. As before, we restrict ourselves to the multipole moments with $n \geq 1$, because for $n = 0$ they are identical to the mass moments. The K multipole moment transforms as

$$\begin{aligned} Q_{K'}^{(n)} &= \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} [-\ln(\lambda) - \ln(1 - \kappa)] \\ &= Q_K^{(n)}. \end{aligned} \quad (63)$$

Like the mass moments, the remaining multipole moments do not change under a transformation of the mass sheet:

$$Q_{g'}^{(n)} = Q_g^{(n)}, \quad (64)$$

$$Q_{F'}^{(n)} = Q_F^{(n)}, \quad (65)$$

$$Q_{F', G'}^{(n)} = Q_{F, G}^{(n)}. \quad (66)$$

4 APPLICATION TO A SIS MODEL

We apply our theoretical framework to a SIS model to illustrate the method. The projected surface mass density is given by (Bartelmann & Schneider 2001):

$$\Sigma_{\text{SIS}}(\xi) = \frac{\sigma_v^2}{2G\xi}, \quad (67)$$

where ξ is the separation from the lens centre in the projected lens plane and σ_v is the line-of-sight velocity dispersion of the particles. The convergence is given by $\kappa(\xi) = \Sigma(\xi)/\Sigma_{\text{crit}}$ and defining the angular distance $\theta = \xi/D_{\text{ol}}$ and the Einstein deflection angle

$$\theta_E = 4\pi \left(\frac{\sigma_v}{c} \right)^2 \frac{D_{\text{ls}}}{D_{\text{os}}} \quad (68)$$

leads to

$$\kappa_{\text{SIS}}(\theta) = \frac{\theta_E}{2\theta}. \quad (69)$$

The shear for a SIS at a vectorial angular separation θ is (Bartelmann & Schneider 2001)

$$\gamma_{\text{SIS}}(\theta) = -\frac{\theta_E}{2\theta} e^{2i\varphi} \quad (70)$$

and for flexion, we have (Bacon et al. 2006)

$$\mathcal{F}_{\text{SIS}}(\theta) = -\frac{\theta_E}{2\theta^2} e^{i\varphi}, \quad \mathcal{G}_{\text{SIS}}(\theta) = \frac{3\theta_E}{2\theta^2} e^{3i\varphi}, \quad (71)$$

where we used the complex shear notation $\gamma = \gamma_1 + i\gamma_2$ and the analogous notation for flexion. Therefore, the tangential and radial components are

$$\gamma_{t,\text{SIS}}(\theta) = \frac{\theta_E}{2\theta}, \quad \gamma_{r,\text{SIS}}(\theta) = 0, \quad (72)$$

$$\mathcal{F}_{t,\text{SIS}}(\theta) = \frac{\theta_E}{2\theta^2}, \quad \mathcal{F}_{r,\text{SIS}}(\theta) = 0, \quad (73)$$

$$\mathcal{G}_{t,\text{SIS}}(\theta) = -\frac{3\theta_E}{2\theta^2}, \quad \mathcal{G}_{r,\text{SIS}}(\theta) = 0. \quad (74)$$

Measuring higher order mass moments tends to be a noisy affair, therefore we examine for simplicity the lowest order mass moment $M^{(0)}$. Since the 0th order mass and multipole moments are identical, we will also restrict our treatment to the mass moment. We investigate the non-reduced moment method, as these moments give short analytic expressions and the reduced moment technique is exactly analogous. Writing the tangential components in terms of θ and performing the integrals using the Gaussian weighting function, we have for the SIS

$$M_{\kappa,\text{SIS}}^{(0)} = M_{\gamma,\text{SIS}}^{(0)} = M_{\mathcal{F},\text{SIS}}^{(0)} = M_{\mathcal{F},\mathcal{G},\text{SIS}}^{(0)} = \frac{\pi\sigma\theta_E}{2}. \quad (75)$$

For the transformed moments, we have

$$M_{\kappa',\text{SIS}}^{(0)} = \lambda \frac{\pi\sigma\theta_E}{2} + \sqrt{2\pi}(1-\lambda)\sigma^2, \quad (76)$$

$$M_{\gamma',\text{SIS}}^{(0)} = M_{\mathcal{F}',\text{SIS}}^{(0)} = M_{\mathcal{F}',\mathcal{G}',\text{SIS}}^{(0)} = \lambda \frac{\pi\sigma\theta_E}{2}. \quad (77)$$

Therefore we can estimate the value of $(1-\lambda)$ by evaluating $(M_{\kappa',\text{SIS}}^{(0)} - M_{\gamma',\text{SIS}}^{(0)})/\sqrt{2\pi}\sigma^2$, where x represents a shear or flexion mass moment.

5 DETECTABILITY

In this section, we estimate the reliability of λ detections from the mass moment measurements *without assuming a halo model*. For

this purpose, we compare the moments with the moment standard deviations in the absence of lensing. The latter are calculated in Appendix C. We assume that the convergence information is derived from number counts exploiting the magnification bias and that the unlensed number counts follow a power law $n_0 \propto S^{-\beta}$ with flux limit S and $\beta = 0.5$ (e.g. Schneider et al. 2000). The results are

$$\sigma_{M,\kappa} = \sqrt{\frac{2\pi}{n_\kappa}} \left(\int_0^\infty dx x^{2n+1} w(x)^2 \right)^{\frac{1}{2}}, \quad (78)$$

$$\sigma_{M,\gamma} = \sqrt{\frac{2\pi}{n_\gamma}} \sigma_\gamma \left(\int_0^\infty dx \frac{1}{x} [2W(x) - x^{n+1}w(x)]^2 \right)^{\frac{1}{2}}, \quad (79)$$

$$\sigma_{M,\mathcal{F}} = \sqrt{\frac{2\pi}{n_{\mathcal{F}}}} \sigma_{\mathcal{F}} \left(\int_0^\infty dx x W(x)^2 \right)^{\frac{1}{2}}, \quad (80)$$

$$\sigma_{M,\mathcal{F},\mathcal{G}} = \sqrt{2\pi} \sqrt{\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}}} \left(\int_0^\infty dx \frac{1}{x} \left[V(x) - \frac{1}{2}xW(x) \right]^2 \right)^{\frac{1}{2}} \quad (81)$$

for the non-reduced moments and

$$\sigma_{M,K} = \left(\frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 A^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2 \right)^{\frac{1}{2}}, \quad (82)$$

$$\sigma_{M,g} = \left(\frac{4\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \sigma_\gamma^2 \left(\int_R^\infty dx x W_R(x)^2 \right) + \frac{2\pi}{n_\gamma} \sigma_\gamma^2 \int_R^\infty dx \frac{1}{x} (2W_R(x) - x^{n+1}w(x))^2 \right)^{\frac{1}{2}}, \quad (83)$$

$$\sigma_{M,F} = \sqrt{\frac{2\pi}{n_{\mathcal{F}}}} \sigma_{\mathcal{F}} \left(\int_R^\infty dx x W_R(x)^2 \right)^{\frac{1}{2}}, \quad (84)$$

$$\begin{aligned} \sigma_{M,F,G} = \sqrt{2\pi} \left(\left(\frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{F}}^2}{4n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{4n_{\mathcal{G}}} \right) \int_R^\infty dx x W_R(x)^2 \right. \\ \left. - \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \right) \int_R^\infty dx W_R(x) V_R(x) \right. \\ \left. + \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} + 4 \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \right) \int_R^\infty dx \frac{1}{x} V_R(x)^2 \right)^{\frac{1}{2}} \quad (85) \end{aligned}$$

for the reduced moments. Here A denotes our total field of view area. Using the two techniques described in this paper, we can measure

Table 1. Detectability of $1 - \lambda$ (upper part) and $\ln(\lambda)$ (lower part) for different moment combinations. The results are shown for galaxy lenses, 1000 stacked galaxy lenses, and cluster lenses for different σ values of the Gaussian $w(x)$ weighting function. The moment combinations listed after the first one involve the W or V weighting function types, which drop off very slowly. Thus, their detectability can be substantially improved by designing a faster declining weighting function (see text).

σ	Galaxy lensing			Stacking			Cluster lensing		
	2 arcsec	5 arcsec	10 arcsec	2 arcsec	5 arcsec	10 arcsec	20 arcsec	50 arcsec	100 arcsec
$\kappa + \gamma$	0.4	0.9	1.9	11.8	29.5	59.0	3.7	9.3	18.7
$\kappa + \mathcal{F}$	0.1	0.2	0.2	4.1	4.9	5.6	0.1	0.2	0.2
$\kappa + \mathcal{F}, \mathcal{G}$	0.04	0.05	0.08	1.2	1.7	2.6	0.04	0.05	0.08
$K + g$	0.2	0.4	0.4	7.6	11.1	13.4	0.3	0.4	0.4
$K + F$	0.1	0.1	0.2	3.8	4.5	5.1	0.1	0.1	0.2
$K + F, G$	0.03	0.05	0.07	1.0	1.5	2.1	0.04	0.05	0.07

$1 - \lambda$ or $\ln(\lambda)$ and their respective detectabilities are given by:

$$\left(\frac{S}{N}\right)_{1-\lambda}^{(n)} = \left| \frac{M_{\kappa'}^{(n)} - M_{x'}^{(n)}}{\left(\sigma_{M^{(n)},\kappa'}^2 + \sigma_{M^{(n)},x'}^2\right)^{\frac{1}{2}}} \right|, \quad (86)$$

$$\left(\frac{S}{N}\right)_{\ln(\lambda)}^{(n)} = \left| \frac{M_{K'}^{(n)} - M_{X'}^{(n)}}{\left(\sigma_{M^{(n)},K'}^2 + \sigma_{M^{(n)},X'}^2\right)^{\frac{1}{2}}} \right|, \quad (87)$$

where x and X represent the non-reduced and reduced shear and flexion moments used to break the degeneracy. We calculate the detectabilities for the lowest order mass moment combinations. We assume a λ of 1.2 and that we can measure the shear in 100 sources/arcmin², as in the *Hubble Frontier Fields*. We further choose $n_{\kappa} = 300/\text{arcmin}^2$, $\sigma_{\gamma} = 0.26$ (Leauthaud et al. 2007), and $n_{\mathcal{F}} = n_{\mathcal{G}} = 25/\text{arcmin}^2$. We have $\sigma_{F a_{\text{gal}}} \approx \sigma_{\mathcal{F} a_{\text{gal}}}$ and $\sigma_{G a_{\text{gal}}} \approx \sigma_{\mathcal{G} a_{\text{gal}}}$, where a_{gal} is the semi-major of the lensed source, and typical values for the intrinsic flexion of 0.03 and 0.04, respectively (Goldberg & Bacon 2005; Goldberg & Leonard 2007). As flexion can more reliably be measured for larger sources, we assume a lower source density and $a_{\text{gal}} = 0.2$ arcsec. This leads to $\sigma_{\mathcal{F}} = 0.15$ 1/arcsec and $\sigma_{\mathcal{G}} = 0.2$ 1/arcsec. For the galaxy lens, we assume an Einstein radius of 0.2 arcsec and for the cluster 20 arcsec, but note that our results do not require any assumptions on the halo shape, as the difference between the mass moments solely depends on lambda and the weighting function. We choose the lower integration boundary R to be the Einstein radius and for the noise integrals involving intrinsic flexion, we choose an upper integration boundary of 3 arcmin for the galaxy lens and of 30 arcmin for the cluster instead of infinity. We can do this because flexion drops off very quickly, typically as $1/x^2$, and therefore we can assume that the flexion and thus its noise is 0 outside of a certain, generously chosen area. For example, for an SIS, we have that the \mathcal{F} and \mathcal{G} signals drop to less than 10^{-5} 1/arcsec at these distances. If we do not make this simplification, we would integrate the noise out to infinity, as the W weighting function does not drop off quickly enough for a Gaussian w function, even though there is obviously no signal in this region. In addition, we assume that we have an image area of 11 arcmin², which corresponds to the field of view of the *Hubble Space Telescope Advanced Camera for Surveys*.

For the stacking calculations, we assume that we stack 1000 galaxy-scale lenses to obtain the convergence map of an averaged galaxy lens. In this case, we can compute the λ parameter and lift the mass-sheet degeneracy for the stacked, i.e. averaged, lensing galaxy.

Table 1 shows the detectability of $1 - \lambda$ and $\ln(\lambda)$ for a Gaussian weighting function w and several choices of σ . A larger σ leads to a wider weighting function that encompasses more of the lensing information and thus results in a larger detectability. If the convergence and non-reduced shear moments are combined, we clearly detect λ and we can break the mass-sheet degeneracy for stacked galaxy lenses and cluster lenses. Similarly, we can estimate λ for stacking if we combine K and reduced shear or if we use the first flexion moments, even though the latter are much noisier.

However, in several cases the detectability is too low to reliably measure λ . The reason is that these moment combinations involve the flexion and thus the weighting function $W(x) = 1/x \int_0^x dy y^{n+1} w(y)$ or $W_R(x) = 1/x \int_R^x dy y^{n+1} w(y)$, respectively. For a Gaussian weight $w(x)$ and $n = 0$, the integral rapidly approaches a constant value. Thus $W(x)$ and $W_R(x)$ drop off very slowly as $1/x$, while the flexion drops off as $1/x^2$. As a result, a large area which contains no signal is highly weighted. This results in a low detectability. If a faster declining W function is designed, e.g. by using a $w(x)$ which becomes negative at a certain distance from the lens centre, the detectability of these moments can be substantially improved. This is in particular true for the mass moments using first and second flexion, as these require in addition the $V(x)$ and $V_R(x)$ weights, which even increase with distance. The mass moments remain finite, as the flexion drops off faster than the weight increases, but a quickly decreasing weight would increase the detectability significantly.

A second reason for using fast declining W and V functions is the following: In our calculation, we could safely neglect very small flexion signals without affecting the resulting mass moment difference, because we assumed a value for λ . However, depending on σ and the mass moment in question, these very small flexion values can be amplified by the slow drop off of the weighting function and neglecting them can lead to a potentially substantial error in the mass moment estimate.

Several optimized aperture mass weighting functions have been proposed (e.g. SB97; Leonard et al. 2009), but they typically assume that the weight w is compensated, i.e. $\int_0^\infty x w(x) dx = 0$. We did not impose this condition on the weighting function in our derivations, and indeed a compensated w cannot be used to break the mass-sheet degeneracy using the non-reduced moments of lowest order, which are expected to have the minimal noise, as the additional term in equation (45) would vanish.

Maturi et al. (2005) propose a non-compensated, optimized weighting function for shear that takes into account the shape of the signal and the noise power spectrum. This approach is promising, but it is complicated by the fact that for our technique the weights w , W , and V are related to each other. Thus it is not possible to optimize the weighting functions independently. In fact, we must

optimize at least two inter-dependent weighting functions for very different signal and noise shapes simultaneously. This optimization is outside the scope of this paper and is therefore left for a future publication.

Lensing measurements in observational data give discrete sets rather than smooth functions of observables. An application of the aperture mass technique to observational data can therefore be achieved by either binning and approximating the integral with a Riemann sum or fitting the results to a smooth function.

6 CONVERGENCE, SHEAR, AND FLEXION SIGNAL-TO-NOISE RATIO

We compare the signal-to-noise ratios for SIS and NFW profiles (Navarro, Frenk & White 1996, 1997). We show that the combination of shear and flexion has a higher signal-to-noise ratio than the number counts method for both galaxy and cluster lensing. Therefore, including number counts in the mass-sheet reconstruction does not improve the lens model. Thus we can use this information to break the mass-sheet degeneracy without sacrificing the accuracy of the model.

We can restrict ourselves to the non-reduced shear and flexion, because the extra $1/(1 - \kappa)$ term in the reduced quantities would only boost their signal-to-noise ratio. The convergence, shear, and flexion formulas for a NFW profile are given in Appendix D. In addition, we only investigate the intrinsic noise of each method and neglect possible additional measurement errors.

We can make flexion dimensionless by multiplying it with the semi-major of the observed galaxy a_{gal} (Goldberg & Bacon 2005). The signal-to-noise ratios are

$$\left(\frac{S}{N}\right)_{F a_{\text{gal}}} = \frac{|F| a_{\text{gal}} \sqrt{N_F}}{\sigma_{F a_{\text{gal}}}}, \quad (88)$$

$$\left(\frac{S}{N}\right)_{G a_{\text{gal}}} = \frac{|G| a_{\text{gal}} \sqrt{N_G}}{\sigma_{G a_{\text{gal}}}}, \quad (89)$$

where $\sigma_{F a_{\text{gal}}}$ and $\sigma_{G a_{\text{gal}}}$ are the dispersions for dimensionless reduced flexion and we average over N_F and N_G sources. We can infer the signal-to-noise ratio of $\mathcal{F} a_{\text{gal}}$ and $\mathcal{G} a_{\text{gal}}$ by replacing the reduced quantities with these expressions. We have again $\sigma_{F a_{\text{gal}}} \approx \sigma_{\mathcal{F} a_{\text{gal}}}$ and $\sigma_{G a_{\text{gal}}} \approx \sigma_{\mathcal{G} a_{\text{gal}}}$ with typical values of 0.03 and 0.04, respectively (Goldberg & Bacon 2005; Goldberg & Leonard 2007). The signal-to-noise ratio of the shear and number counts is (Schneider et al. 2000)

$$\left(\frac{S}{N}\right)_g = \frac{|g| \sqrt{N_g}}{\sigma_\epsilon}, \quad (90)$$

$$\left(\frac{S}{N}\right)_\kappa = |\mu^{\beta-1} - 1| \sqrt{N_\kappa} \approx 2\kappa |1 - \beta| \sqrt{N_\kappa}, \quad (91)$$

where we have used the first-order weak lensing expansion. We average our number counts over N_κ sources and β is the number counts exponent (see Appendix C). As we investigate the non-reduced quantities, we replace g with γ . We choose $\sigma_\epsilon = 0.26$ (Leauthaud et al. 2007) and the remaining parameters analogously to Schneider et al. (2000), $\beta = 0.5$ and $N_\kappa = 3.5 N_g$. The latter is a slightly optimistic estimate for N_κ , since for real observations, we will lose several sources due to colour cuts. A lower estimate would decrease the signal-to-noise ratio of the number counts method, so we will

keep this value. Using the expressions for the SIS, we have

$$\frac{(S/N)_\gamma}{(S/N)_\kappa} = 2.1, \quad (92)$$

$$\frac{(S/N)_{\mathcal{F} a_{\text{gal}}}}{(S/N)_\kappa} = \frac{a_{\text{gal}}}{0.03\theta} \sqrt{\frac{N_F}{N_\kappa}}, \quad (93)$$

$$\frac{(S/N)_{\mathcal{G} a_{\text{gal}}}}{(S/N)_\kappa} = \frac{3a_{\text{gal}}}{0.04\theta} \sqrt{\frac{N_G}{N_\kappa}}. \quad (94)$$

The radial dependence of the ratios is shown in Fig. 1 for $a_{\text{gal}} = 0.2$ arcmin and $N_F = N_G$ values of N_g , $0.5 N_g$, and $0.1 N_g$. Note that the signal-to-noise ratios do not depend on the Einstein radius θ_E , which is cancelled in the calculation. However, the distance of the weak lensing regime from the halo centre does depend on the Einstein radius. Therefore we show κ for Einstein radii of 0.2, 10, and 20 arcsec. For larger Einstein radii, shear by itself has always the best signal-to-noise ratio in the weak lensing regime. For smaller θ_E , shear and flexion together give the best results. Note that even in the case of large Einstein radii, the flexion information is not redundant. It is much more sensitive to small-scale structures than shear and can therefore be used to reconstruct them (Leonard, King & Wilkins 2009; Bacon, Amara & Read 2010).

The analytic expressions for the NFW profile are complex and therefore it is more instructing to plot the signal-to-noise ratios. These are independent of ρ_{crit} , δ_c , and Σ_{crit} . We show the radial dependence for three angular scale radii $\theta_s = r_s/D_{\text{ol}}$ in Fig. 2. The first plot corresponds to a galaxy at redshift $z = 0.35$ with $r_s = 16h^{-1}$ kpc and sources at redshift $z = 0.6$, which agrees e.g. with the observations in Hoekstra et al. (2004). The second and third correspond to clusters with $r_s = 250h^{-1}$ kpc and $r_s = 70h^{-1}$ kpc at the same redshift, comparable to the NFW models for one component of Abell 370 and for the cluster MS 2137 in Shu et al. (2008). For the first cluster, we only model one component to obtain a NFW profile with intermediate Einstein radius for our comparison. Assuming a standard, flat Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $h = 0.7$, we calculate the angular diameter distances using the Wright (2006) cosmology calculator and astropy (Astropy Collaboration et al. 2013) and obtain the corresponding angular scale radii 4.6 arcsec, 72 arcsec, and 20 arcsec. For the galaxy, we further set $\delta_c = 2.4 \times 10^4$ and for the two clusters $\kappa_s = \rho_{\text{crit}} \delta_c r_s / \Sigma_{\text{crit}} = 0.16$ and 0.66, in agreement with the values cited in the respective publications. The size of the semi-major and the number of measured sources are the same as for the SIS. The effect of a different choice of these parameters can be visualized by a shift of the corresponding graphs, because the scale in the figures is logarithmic.

Fig. 2 shows that shear by itself has always the best signal-to-noise ratio in the weak lensing regime of the two clusters. For the galaxy, a combination of flexion and shear provides the best results, as the flexion dominates at small separations from the centre. This illustrates also the value of flexion for the reconstruction of substructures.

7 DISCUSSION AND CONCLUSIONS

We have derived mass and multipole moment formulas in terms of flexion and showed that they are equivalent to the mass and multipole moments of shear and convergence. Furthermore, we have showed the equivalence of mass and multipole moments in terms of reduced shear, reduced flexion, and K , which is a quantity derived

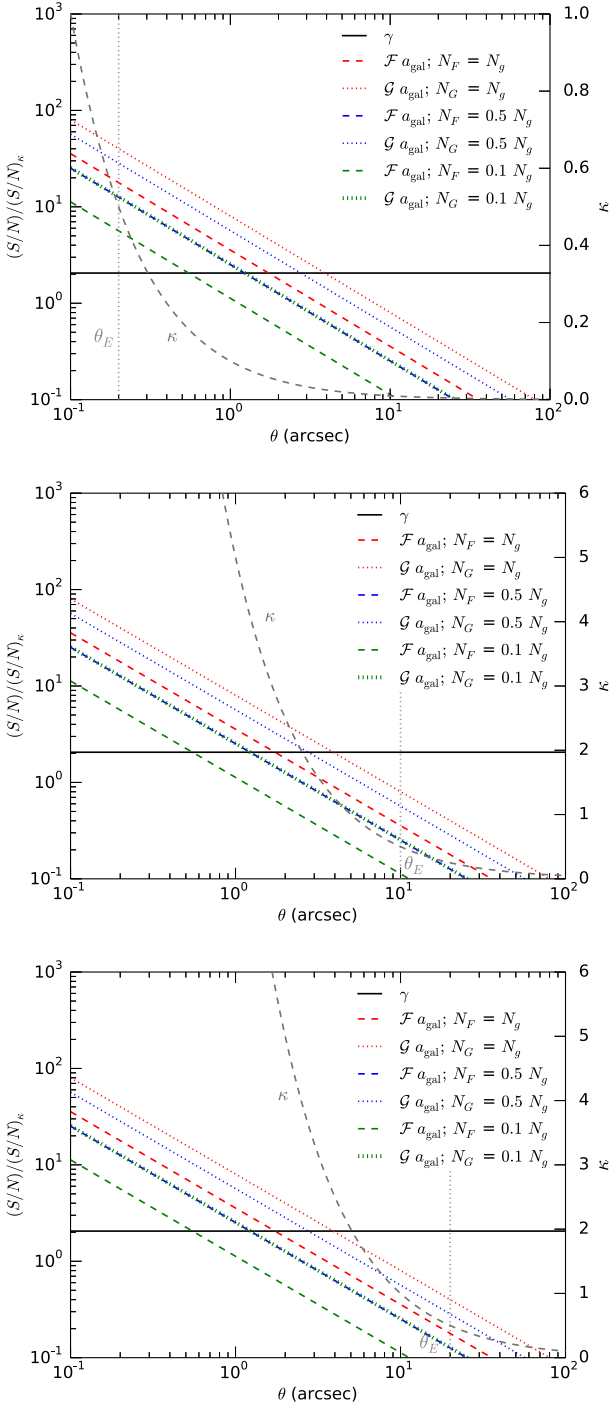


Figure 1. Comparison of the signal-to-noise ratios of γ , $\mathcal{F}a_{\text{gal}}$, and $\mathcal{G}a_{\text{gal}}$ normalized to $(S/N)_k$, the signal-to-noise ratio of number counts, for the SIS. The Einstein radii are 0.2 arcsec (top), 10 arcsec (middle), and 20 arcsec (bottom). We show the flexion results for source densities of 100 per cent, 50 per cent, and 10 per cent of the shear source density and plot the convergence value as a reference.

from the convergence. We investigated the moment behaviour with respect to the mass-sheet degeneracy transformation and found that the equivalences are broken for the mass moments, but preserved for the multipole moments. The resulting surface terms can be used to break the mass-sheet degeneracy.

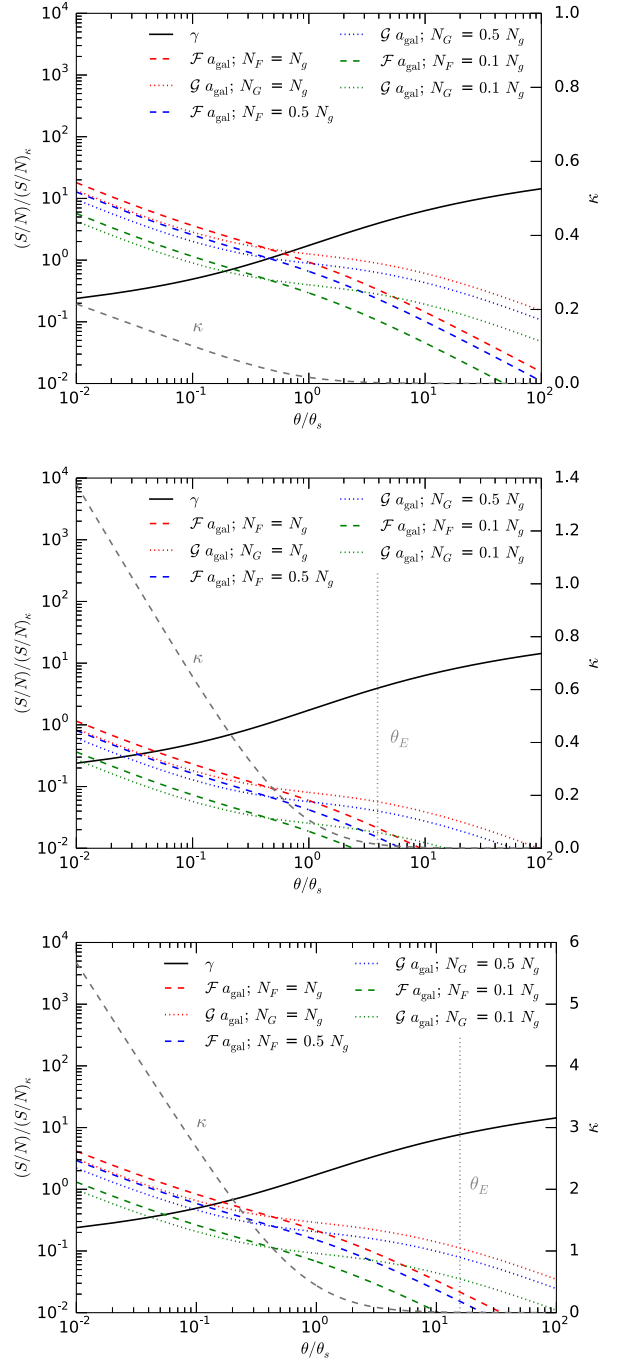


Figure 2. Comparison of (S/N) ratios of γ , $\mathcal{F}a_{\text{gal}}$, and $\mathcal{G}a_{\text{gal}}$ normalized to $(S/N)_k$, the signal-to-noise ratio of number counts, for NFW profiles with $\theta_s = 4.6$ arcsec (top), 72 arcsec (middle), and 20 arcsec (bottom). The flexion is shown for source densities of 100 per cent, 50 per cent, and 10 per cent of the shear source density. The convergence and the Einstein radius are plotted as a reference. The Einstein radius of the galaxy is too small to be shown.

We demonstrated the new theoretical framework by applying it to a SIS mass model. In addition, we investigated the detectability of the mass-sheet parameter λ without assuming a halo model. We found that we can break the mass-sheet degeneracy for stacked galaxy–galaxy and cluster lensing. Combinations of the shear and convergence moments have a much higher detectability than

combinations using the flexions. This is due to their weighting function, which drops off too slowly. A fine-tuned weight will significantly improve their performance.

Finally, we investigated the signal-to-noise ratios of shear, flexion, and convergence information from number counts for SIS and NFW haloes in the weak lensing regime. It was assumed that the noise in γ , \mathcal{F} , and \mathcal{G} is dominated by the intrinsic signal of the lensed sources and that the number counts are dominated by Poisson noise. The estimates do not include measurement errors e.g. due to pixel noise. This will have an additional effect, in particular on flexion and if sources with insufficient signal-to-noise ratio are used (Rowe et al. 2013). With high-quality data, e.g. from the *Hubble Frontier Fields*, the reliability of the flexion measurements should improve. Under these assumptions, we demonstrated that the combination of shear and flexion always gives better results than other combinations using the number counts. Using an estimate of shear and number count signal-to-noise ratios, e.g. Schneider et al. (2000) also found that shear is the superior method. However, Van Waerbeke et al. (2010) and Hildebrandt et al. (2013) caution that we can still count sources at redshifts for which shear measurements are no longer feasible, thus reducing the advantage of shear. We find that the convergence information can thus be used to break the mass-sheet degeneracy without a loss in the quality of the lens model.

ACKNOWLEDGEMENTS

MR thanks the Department of Astronomy at Yale University for hosting him and acknowledges support by the German National Academic Foundation. He thanks Thibault Kuntzer, Austen Groener, and Justin Bird for fruitful discussions and Alexandre Refregier for his support. MR and PN acknowledge useful discussions with David Goldberg and Richard Massey on flexions. MR and JPK acknowledge support from the ERC advanced grant LIDA. PN gratefully acknowledges support from an NSF theory program via the grant AST-1044455 and a theory grant from the Space Telescope Science Institute HST-AR-12144.01-A.

REFERENCES

Astropy Collaboration et al., 2013, *A&A*, 558, A33
 Bacon D. J., Goldberg D. M., Rowe B. T. P., Taylor A. N., 2006, *MNRAS*, 365, 414
 Bacon D. J., Amara A., Read J. I., 2010, *MNRAS*, 409, 389
 Bartelmann M., 1996, *A&A*, 313, 697

Bartelmann M., Narayan R., 1995, *ApJ*, 451, 60
 Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291
 Bartelmann M., Narayan R., Seitz S., Schneider P., 1996, *ApJ*, 464, L115
 Bradač M., Lombardi M., Schneider P., 2004, *A&A*, 424, 13
 Broadhurst T. J., Taylor A. N., Peacock J. A., 1995, *ApJ*, 438, 49
 Cain B., Schechter P. L., Bautz M. W., 2011, *ApJ*, 736, 43
 Dye S. et al., 2002, *A&A*, 386, 12
 Er X., Li G., Schneider P., 2010, preprint ([arXiv:1008.3088](https://arxiv.org/abs/1008.3088))
 Fort B., Mellier Y., Dantel-Fort M., 1997, *A&A*, 321, 353
 Goldberg D. M., Bacon D. J., 2005, *ApJ*, 619, 741
 Goldberg D. M., Leonard A., 2007, *ApJ*, 660, 1003
 Goldberg D. M., Natarajan P., 2002, *ApJ*, 564, 65
 Goodman L. A., 1960, *J. Am. Stat. Assoc.*, 55, 708
 Hildebrandt H. et al., 2013, *MNRAS*, 429, 3230
 Hoekstra H., Yee H. K. C., Gladders M. D., 2004, *ApJ*, 606, 67
 Irwin J., Shmakova M., 2006, *ApJ*, 645, 17
 Kaiser N., 1995, *ApJ*, 439, L1
 Kaiser N., Squires G., 1993, *ApJ*, 404, 441
 Kneib J.-P., Natarajan P., 2011, *A&AR*, 19, 47
 Leauthaud A. et al., 2007, *ApJS*, 172, 219
 Leonard A., Goldberg D. M., Haaga J. L., Massey R., 2007, *ApJ*, 666, 51
 Leonard A., King L. J., Wilkins S. M., 2009, *MNRAS*, 395, 1438
 Maturi M., Meneghetti M., Bartelmann M., Dolag K., Moscardini L., 2005, *A&A*, 442, 851
 Meylan G., Jetzer P., North P., Schneider P., Kochanek C. S., Wambsgans J., 2006, *Gravitational Lensing: Strong, Weak and Micro*. Springer, Berlin
 Natarajan P., Kneib J.-P., Smail I., Ellis R. S., 1998, *ApJ*, 499, 600
 Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
 Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
 Okura Y., Umetsu K., Futamase T., 2007, *ApJ*, 660, 995
 Okura Y., Umetsu K., Futamase T., 2008, *ApJ*, 680, 1
 Rexroth M., 2015, *IAU Proc. Astronomy in Focus*, submitted
 Rowe B., Bacon D., Massey R., Heymans C., Häußler B., Taylor A., Rhodes J., Mellier Y., 2013, *MNRAS*, 435, 822
 Schneider P., Bartelmann M., 1997, *MNRAS*, 286, 696 (SB97)
 Schneider P., Er X., 2008, *A&A*, 485, 363
 Schneider P., Seitz C., 1995, *A&A*, 294, 411
 Schneider P., King L., Erben T., 2000, *A&A*, 353, 41
 Shu C., Zhou B., Bartelmann M., Comerford J. M., Huang J.-S., Mellier Y., 2008, *ApJ*, 685, 70
 Taylor A. N., Dye S., Broadhurst T. J., Benitez N., van Kampen E., 1998, *ApJ*, 501, 539
 Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Hsu L.-Y., 2011, *ApJ*, 729, 127
 Van Waerbeke L., Hildebrandt H., Ford J., Milkeraitis M., 2010, *ApJ*, 723, L13
 Wright E. L., 2006, *PASP*, 118, 1711
 Wright C. O., Brainerd T. G., 2000, *ApJ*, 534, 34

APPENDIX A: DERIVATION OF THE MASS AND MULTIPOLE MOMENT EQUIVALENCES FOR CONVERGENCE, SHEAR, AND FLEXION

A1 Moments on the full field of view

We compute the general n th order moments within an aperture defined by the weighting function $w(x)$. This allows the application to finite field observational data. The moments of the mass distribution in terms of κ are defined as (SB97)

$$M_{\kappa}^{(n)} = \int_0^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi \kappa(x, \varphi, x_0, \varphi_0) \quad (\text{A1})$$

and the corresponding multipole moments are defined as

$$Q_{\kappa}^{(n)} = \int_0^{\infty} dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} \kappa(x, \varphi, x_0, \varphi_0). \quad (\text{A2})$$

From now on we will no longer write the explicit dependence on x_0 and φ_0 to simplify the notation. As in [SB97](#), we can integrate equation (A1) by parts to obtain

$$M_\kappa^{(n)} = - \int_0^\infty dx x W(x) \int_0^{2\pi} d\varphi \frac{\partial \kappa(x, \varphi)}{\partial x}, \quad (\text{A3})$$

where

$$x W(x) = \int_0^x dy y^{n+1} w(y) \quad (\text{A4})$$

and we require that the boundary term $x W(x) \kappa(x, \varphi)$ vanishes for $x \rightarrow 0$ and $x \rightarrow \infty$. This can be achieved by picking a weighting function which drops off sufficiently fast. Expressing the radial derivative in Cartesian coordinates and using $\mathcal{F} = \nabla \kappa$ ([Bacon et al. 2006](#)), we can write the mass moments in terms of first flexion:

$$M_{\mathcal{F}}^{(n)} = \int_0^\infty dx x W(x) \int_0^{2\pi} d\varphi \mathcal{F}_i(x, \varphi), \quad (\text{A5})$$

where

$$\mathcal{F}_i(x, \varphi) = - [F_1(x, \varphi) \cos(\varphi) + F_2(x, \varphi) \sin(\varphi)], \quad (\text{A6})$$

$$F_i(x, \varphi) = - [F_1(x, \varphi) \cos(\varphi) + F_2(x, \varphi) \sin(\varphi)], \quad (\text{A7})$$

defines the non-reduced and reduced tangential first flexion. We use a mass-sheet reconstruction κ_{rec} to find

$$M_{\mathcal{F}}^{(n)} = \int_0^\infty dx x W(x) \int_0^{2\pi} d\varphi (1 - \kappa_{\text{rec}}(x, \varphi)) F_i(x, \varphi). \quad (\text{A8})$$

Note that κ_{rec} is different from the convergence that is used to calculate the moments in equations (A1) and (A2). Applying the same operations to equation (A2), we can also express the multipole moments using first flexion:

$$Q_{\mathcal{F}}^{(n)} = \int_0^\infty dx x W(x) \int_0^{2\pi} d\varphi e^{in\varphi} \mathcal{F}_i(x, \varphi) \quad (\text{A9})$$

$$= \int_0^\infty dx x W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}(x, \varphi)) F_i(x, \varphi). \quad (\text{A10})$$

The moments can also be expressed using a combination of \mathcal{F} and \mathcal{G} . We define the tangential and radial shear and the corresponding reduced quantities by

$$\gamma_t(x, \varphi) = - [\gamma_1(x, \varphi) \cos(2\varphi) + \gamma_2(x, \varphi) \sin(2\varphi)], \quad (\text{A11})$$

$$\gamma_r(x, \varphi) = - [\gamma_2(x, \varphi) \cos(2\varphi) - \gamma_1(x, \varphi) \sin(2\varphi)], \quad (\text{A12})$$

$$g_t(x, \varphi) = - [g_1(x, \varphi) \cos(2\varphi) + g_2(x, \varphi) \sin(2\varphi)], \quad (\text{A13})$$

$$g_r(x, \varphi) = - [g_2(x, \varphi) \cos(2\varphi) - g_1(x, \varphi) \sin(2\varphi)], \quad (\text{A14})$$

and use the relation shown in [Kaiser \(1995\)](#)

$$\nabla \kappa = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}. \quad (\text{A15})$$

Inserting the radial part of this relation transformed to polar coordinates into equation (A3), using trigonometric identities, integrating by parts and demanding that the surface terms $x W(x) \gamma_t(x, \varphi)$ vanish for $x \rightarrow 0, \infty$ leads to an expression of the mass moments using shear ([SB97](#)),

$$M_\gamma^{(n)} = \int_0^\infty dx [2W(x) - x^{n+1} w(x)] \int_0^{2\pi} d\varphi \gamma_t(x, \varphi) = \int_0^\infty dx [2W(x) - x^{n+1} w(x)] \int_0^{2\pi} d\varphi (1 - \kappa_{\text{rec}}(x, \varphi)) g_t(x, \varphi). \quad (\text{A16})$$

We define

$$V(x) = \int_0^x dy W(y) \quad (\text{A17})$$

and require that $[2V(x) - x W(x)] \gamma_t(x, \varphi) \rightarrow 0$ for $x \rightarrow 0, x \rightarrow \infty$. Then the boundary terms resulting from a partial integration with respect to x vanish and we find

$$M_\gamma^{(n)} = - \int_0^\infty dx [2V(x) - x W(x)] \int_0^{2\pi} d\varphi \frac{\partial \gamma_t(x, \varphi)}{\partial x}. \quad (\text{A18})$$

We further define the radial first flexion, the tangential and radial second flexion and their reduced counterparts by

$$\mathcal{F}_r(x, \varphi) = -[\mathcal{F}_2(x, \varphi) \cos(\varphi) - \mathcal{F}_1(x, \varphi) \sin(\varphi)], \quad (\text{A19})$$

$$\mathcal{G}_t(x, \varphi) = -[\mathcal{G}_1(x, \varphi) \cos(3\varphi) + \mathcal{G}_2(x, \varphi) \sin(3\varphi)], \quad (\text{A20})$$

$$\mathcal{G}_r(x, \varphi) = -[\mathcal{G}_2(x, \varphi) \cos(3\varphi) - \mathcal{G}_1(x, \varphi) \sin(3\varphi)], \quad (\text{A21})$$

$$F_r(x, \varphi) = -[F_2(x, \varphi) \cos(\varphi) - F_1(x, \varphi) \sin(\varphi)], \quad (\text{A22})$$

$$G_t(x, \varphi) = -[G_1(x, \varphi) \cos(3\varphi) + G_2(x, \varphi) \sin(3\varphi)], \quad (\text{A23})$$

$$G_r(x, \varphi) = -[G_2(x, \varphi) \cos(3\varphi) - G_1(x, \varphi) \sin(3\varphi)]. \quad (\text{A24})$$

Now we can express the radial derivative in Cartesian coordinates, apply trigonometric identities and use the relations between flexion and shear derivatives in equation (11) to obtain the mass moments in terms of first and second flexion:

$$M_{\mathcal{F}, \mathcal{G}}^{(n)} = -\int_0^\infty dx \left[V(x) - \frac{1}{2}xW(x) \right] \int_0^{2\pi} d\varphi [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)] \quad (\text{A25})$$

$$= -\int_0^\infty dx \left[V(x) - \frac{1}{2}xW(x) \right] \int_0^{2\pi} d\varphi [(1 - \kappa_{\text{rec}}(x, \varphi))F_t(x, \varphi) + (1 - \kappa_{\text{rec}}(x, \varphi))G_t(x, \varphi)]. \quad (\text{A26})$$

The multipole moments using \mathcal{F} and \mathcal{G} can be derived in an analogous way. We integrate equation (A2) by parts and require that the boundary term $xW(x)\kappa(x, \varphi)$ vanishes for $x \rightarrow 0$ and $x \rightarrow \infty$. Inserting the radial part of equation (A15) transformed to polar coordinates, integrating by parts and demanding that $xW(x)\gamma_t(x, \varphi)$ vanishes for $x \rightarrow 0, \infty$ leads to the multipole moments in terms of shear (SB97),

$$Q_\gamma^{(n)} = \int_0^\infty dx \left[2W(x) - x^{n+1}w(x) \right] \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_t(x, \varphi) - in \int_0^\infty dx W(x) \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_t(x, \varphi) \quad (\text{A27})$$

$$= \int_0^\infty dx \left[2W(x) - x^{n+1}w(x) \right] \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}(x, \varphi))g_t(x, \varphi) - in \int_0^\infty dx W(x) \int_0^{2\pi} d\varphi e^{in\varphi} (1 - \kappa_{\text{rec}}(x, \varphi))g_r(x, \varphi). \quad (\text{A28})$$

Partial integration with respect to x and applying the same transformations and identities as before lead to

$$Q_{\mathcal{F}, \mathcal{G}}^{(n)} = -\int_0^\infty dx \left[V(x) - \frac{1}{2}xW(x) \right] \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)] + i\frac{n}{2} \int_0^\infty dx V(x) \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_r(x, \varphi) + \mathcal{G}_r(x, \varphi)] \quad (\text{A29})$$

$$\begin{aligned} Q_{\mathcal{F}, \mathcal{G}}^{(n)} = & -\int_0^\infty dx \left[V(x) - \frac{1}{2}xW(x) \right] \int_0^{2\pi} d\varphi e^{in\varphi} [(1 - \kappa_{\text{rec}}(x, \varphi))F_t(x, \varphi) + (1 - \kappa_{\text{rec}}(x, \varphi))G_t(x, \varphi)] \\ & + i\frac{n}{2} \int_0^\infty dx V(x) \int_0^{2\pi} d\varphi e^{in\varphi} [(1 - \kappa_{\text{rec}}(x, \varphi))F_r(x, \varphi) + (1 - \kappa_{\text{rec}}(x, \varphi))G_r(x, \varphi)], \end{aligned} \quad (\text{A30})$$

where we required that

$$[2V(x) - xW(x)]\gamma_t(x, \varphi) \rightarrow 0 \quad \text{for } x \rightarrow 0, x \rightarrow \infty, \quad (\text{A31})$$

$$V(x)\gamma_r(x, \varphi) \rightarrow 0 \quad \text{for } x \rightarrow 0, x \rightarrow \infty, \quad (\text{A32})$$

so that the boundary terms in the partial integration vanish.

A1.1 Moments on rings

We now extend our moment formulas to rings, i.e. we omit the innermost part of the lens in the integration, because we typically do not have weak lensing information in this area. The derivation is exactly analogous to the one presented in the previous section and we will therefore only state the resulting equivalent formulas and the respective requirements for the weighting function. Furthermore, we will not explicitly denote the reconstruction of shear and flexion using κ_{rec} to keep this section brief. We define

$$M_{R;\kappa}^{(n)} = \int_R^\infty dx x^{n+1}w(x) \int_0^{2\pi} d\varphi \kappa(x, \varphi) \quad (\text{A33})$$

and the corresponding multipole moments

$$\mathcal{Q}_{R;x}^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} \kappa(x, \varphi). \quad (\text{A34})$$

We further define

$$x W_R(x) = \int_R^x dy y^{n+1} w(y) \quad (\text{A35})$$

and demand that $xW_R(x)\kappa(x, \varphi)$ vanishes for $x \rightarrow R, \infty$. Note that this is simply a different choice of the integration constant. We obtain

$$M_{R;\mathcal{F}}^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi \mathcal{F}_t(x, \varphi), \quad (\text{A36})$$

$$\mathcal{Q}_{R;\mathcal{F}}^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} \mathcal{F}_t(x, \varphi). \quad (\text{A37})$$

We further demand that the term $xW_R(x)\gamma_t(x, \varphi)$ vanishes for $x \rightarrow R, \infty$ and note that the other surface terms cancel due to periodicity. We have now

$$M_{R;\gamma}^{(n)} = \int_R^\infty dx [2W_R(x) - x^{n+1}w(x)] \int_0^{2\pi} d\varphi \gamma_t(x, \varphi), \quad (\text{A38})$$

$$\mathcal{Q}_{R;\gamma}^{(n)} = \int_R^\infty dx [2W_R(x) - x^{n+1}w(x)] \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_t(x, \varphi) - in \int_R^\infty dx W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} \gamma_t(x, \varphi). \quad (\text{A39})$$

We define

$$V_R(x) = \int_R^x dy W_R(y) \quad (\text{A40})$$

and require that $[2V_R(x) - xW_R(x)]\gamma_t(x, \varphi)$ and, for the multipole moments, $V_R(x)\gamma_t(x, \varphi)$ vanish for $x \rightarrow R, \infty$ to find

$$M_{R;\mathcal{F},\mathcal{G}}^{(n)} = - \int_R^\infty dx [V_R(x) - \frac{1}{2}xW_R(x)] \int_0^{2\pi} d\varphi [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)], \quad (\text{A41})$$

$$\mathcal{Q}_{R;\mathcal{F},\mathcal{G}}^{(n)} = - \int_R^\infty dx [V_R(x) - \frac{1}{2}xW_R(x)] \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)] + i\frac{n}{2} \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} [\mathcal{F}_t(x, \varphi) + \mathcal{G}_t(x, \varphi)]. \quad (\text{A42})$$

If weak lensing data are only available in a small part of the ring, we can further restrict the integration to a partial ring. For this purpose, we modify the mass and multipole moment equations for κ and \mathcal{F} by replacing 2π with the desired maximum angle ϕ and their equivalence still holds. However, they are now no longer equivalent to the γ and \mathcal{F}, \mathcal{G} moments. The integration area of these moments cannot be restricted in general, since the derivation of the equivalences uses the periodicity due to the 2π boundary. However, in the case of the mass moments $M_{R;\gamma}^{(n)}$ and $M_{R;\mathcal{F},\mathcal{G}}^{(n)}$, replacing 2π with π also preserves the required periodicity and thus it is possible to restrict these moments to one half of a ring. Therefore, all mass moments are also equivalent for $\phi = \pi$.

APPENDIX B: DERIVATION OF THE MASS AND MULTIPOLE MOMENT EQUIVALENCES FOR κ , REDUCED SHEAR, AND REDUCED FLEXIONS

We derive the moment equivalence relations for the reduced quantities. Following Cain et al. (2011), we define

$$K(x, \varphi, x_0, \varphi_0) = -\ln(1 - \kappa(x, \varphi, x_0, \varphi_0)) \quad (\text{B1})$$

and thus

$$K_{,i} = \frac{1}{1 - \kappa} \kappa_{,i} = F_i. \quad (\text{B2})$$

We want $\kappa < 1$ and choose the lower integral limit for our ring, R , such that $\kappa(x, \varphi) < 1 \forall x \geq R, \forall \varphi$. Therefore $K(x, \varphi)$ is well-defined and finite and we can define

$$M_K^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi K(x, \varphi, x_0, \varphi_0), \quad (\text{B3})$$

$$\mathcal{Q}_K^{(n)} = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi e^{in\varphi} K(x, \varphi, x_0, \varphi_0). \quad (\text{B4})$$

In the following, we will again not explicitly denote x_0 and φ_0 to keep the notation simple. We require that $xW_R(x)K(x, \varphi)$ vanishes for $x \rightarrow R$ and $x \rightarrow \infty$ and integrate equation (B3) by parts with respect to x to find

$$M_K^{(n)} = - \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi \frac{\partial K(x, \varphi)}{\partial x}. \quad (\text{B5})$$

We now express the radial derivative in Cartesian coordinates, apply equation (B2) and use the definition of F_t in equation (A7). Then we have

$$M_F^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi F_t(x, \varphi). \quad (\text{B6})$$

In the exact same way we find for the multipole moments

$$Q_F^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} F_t(x, \varphi). \quad (\text{B7})$$

Now we would like to express the moments using the reduced shear. Therefore, we rewrite the K derivatives by using equation (A15) and the definition of g :

$$K_{,1} = g_{1,1} - g_1 K_{,1} + g_{2,2} - g_2 K_{,2}, \quad (\text{B8})$$

$$K_{,2} = g_{2,1} - g_2 K_{,1} - g_{1,2} + g_1 K_{,2}. \quad (\text{B9})$$

Furthermore, we use $K_{,i} = F_i$ to find

$$\frac{\partial K}{\partial x} = -F_t, \quad (\text{B10})$$

$$\frac{\partial K}{\partial \varphi} = -x F_r. \quad (\text{B11})$$

Transforming the Cartesian derivatives on the right hand side of equations (B8) and (B9) into polar coordinates, using trigonometric identities and the definitions of the tangential and radial reduced shear and reduced flexion, we have

$$\frac{\partial K}{\partial x} = \cos(\varphi) K_{,1} + \sin(\varphi) K_{,2} = - \left[\frac{\partial g_t}{\partial x} - \frac{\cos(2\varphi)}{x} \frac{\partial g_2}{\partial \varphi} + \frac{\sin(2\varphi)}{x} \frac{\partial g_1}{\partial \varphi} + g_t F_t + g_r F_r \right]. \quad (\text{B12})$$

We insert this expression into equation (B5), integrate the term with the x -derivative by parts and demand that $x W_R(x) g_t(x, \varphi)$ vanishes for $x \rightarrow R$ and $x \rightarrow \infty$. Then, the surface term vanishes and we have

$$M_F^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_t F_t + g_r F_r) - \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi g_t - \int_R^\infty dx W_R(x) \int_0^{2\pi} d\varphi \left(\cos(2\varphi) \frac{\partial g_2}{\partial \varphi} - \sin(2\varphi) \frac{\partial g_1}{\partial \varphi} \right). \quad (\text{B13})$$

Integrating the φ derivatives by parts and noting that the surface terms vanish due to their periodicity, we find

$$M_g^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) + \int_R^\infty dx (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} d\varphi g_t(x, \varphi). \quad (\text{B14})$$

Using the same steps, we can derive the multipole moments in terms of reduced shear and reduced flexion,

$$Q_g^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) + \int_R^\infty dx (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} d\varphi e^{in\varphi} g_t(x, \varphi) - in \int_R^\infty dx W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} g_r(x, \varphi). \quad (\text{B15})$$

We can integrate the second term in equation (B14) by parts with respect to x and demand that the surface term $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ vanishes for $x \rightarrow R$ and $x \rightarrow \infty$. We have now

$$M_g^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) - \int_R^\infty dx (2V_R(x) - xW_R(x)) \int_0^{2\pi} d\varphi \frac{\partial g_t(x, \varphi)}{\partial x}. \quad (\text{B16})$$

Again applying a transformation to Cartesian coordinates, trigonometric identities, $\nabla \kappa = \mathcal{F}$, and the relations from equation (11), we find for the derivative

$$\frac{\partial g_t}{\partial x} = \frac{1}{2}(F_t + G_t) - g_t F_t \quad (\text{B17})$$

and thus we have for the mass moment

$$M_{F,G}^{(n)} = \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi (g_r(x, \varphi) F_r(x, \varphi) + \frac{1}{2}(F_t(x, \varphi) + G_t(x, \varphi))) - \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi (F_t(x, \varphi) + G_t(x, \varphi) - 2g_t(x, \varphi) F_t(x, \varphi)). \quad (\text{B18})$$

Analogously, we can integrate the last two terms in equation (B15) by parts and demand that the resulting surface terms $(2V_R(x) - xW_R(x))g_t(x, \varphi)$ and $V_R(x)g_r(x, \varphi)$ vanish for $x \rightarrow R$ and $x \rightarrow \infty$. In the same way as above, we can calculate the derivative as

$$\frac{\partial g_r}{\partial x} = \frac{1}{2}(F_r + G_r) - g_r F_t \quad (\text{B19})$$

and use our result for $\partial g_i/\partial x$ to find

$$\begin{aligned} Q_{F,G}^{(n)} = & \int_R^\infty dx x W_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} (g_t(x, \varphi) F_t(x, \varphi) + g_r(x, \varphi) F_r(x, \varphi)) - \int_R^\infty dx (2V_R(x) - xW_R(x)) \int_0^{2\pi} d\varphi e^{in\varphi} \\ & \cdot \left(\frac{1}{2}(F_t(x, \varphi) + G_t(x, \varphi)) - g_t(x, \varphi) F_t(x, \varphi) \right) + in \int_R^\infty dx V_R(x) \int_0^{2\pi} d\varphi e^{in\varphi} \cdot \left(\frac{1}{2}(F_r(x, \varphi) + G_r(x, \varphi)) - g_r(x, \varphi) F_r(x, \varphi) \right). \end{aligned} \quad (\text{B20})$$

If weak lensing data are only available in a small angular window, we can restrict the integration area to a partial ring. We modify the mass and multipole moment equations for K and F by replacing 2π with the maximum angle ϕ and they will still be equivalent. However, they are then no longer equivalent to the g and F, G moments. The integration area of these moments cannot be further restricted, because otherwise the cancellation of the surface terms due to the periodicity would no longer hold.

APPENDIX C: MASS MOMENT DISPERSIONS

C1 Shear, flexion, and convergence moments

We calculate the mass moment dispersions in the absence of lensing. For the convergence, we have $\langle M_\kappa^{(n)} \rangle = 0$ and

$$\sigma_{M,\kappa}^2 = \langle |M_\kappa^{(n)}|^2 \rangle \quad (\text{C1})$$

$$= \int_0^\infty \int_0^\infty dx dy x^{n+1} y^{n+1} w(x)w(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle \kappa(x, \varphi) \kappa(y, \varphi') \rangle. \quad (\text{C2})$$

Following Schneider et al. (2000), the magnification signal for number counts is $|\Delta n_\kappa| = |\mu^{\beta-1} - 1|n_\kappa$, where μ is the magnification and n_κ is the source density. The Poisson noise is $\sqrt{N_{\kappa,a}}$, where $N_{\kappa,a}$ is the number of sources over which we average to determine κ in a given area a . We neglect the additional noise due to source clustering. In the weak lensing regime and assuming $\beta = 0.5$, we have $\mu \approx 1 + 2\kappa$ and thus $|\kappa| \approx |\Delta n_\kappa|/n_\kappa$.

We assume a discrete distribution of κ over an area A and measure the convergence by averaging our source counts in an area a . Therefore, our measurement points of κ are uncorrelated. Thus, we have for the sum of the convergence variance over the area A :

$$\frac{n_\kappa^2 a^2}{n_\kappa^2} \sum_{i=1}^{N_\kappa/n_\kappa a} \sum_{j=1}^{N_\kappa/n_\kappa a} \langle \kappa_i \kappa_j \rangle = a^2 \sum_{i=1}^{N_\kappa/n_\kappa a} \frac{1}{n_\kappa a} = \frac{A}{n_\kappa}, \quad (\text{C3})$$

$$\langle \kappa_i \kappa_j \rangle = \frac{1}{n_\kappa a} \delta_{ij}. \quad (\text{C4})$$

Extending this to integration, we have

$$\int \int dx^2 dy^2 \langle \kappa(\mathbf{x}) \kappa(\mathbf{y}) \rangle = \frac{1}{n_\kappa} \int dx^2 = \frac{A}{n_\kappa} \quad (\text{C5})$$

and thus

$$\langle \kappa(\mathbf{x}) \kappa(\mathbf{y}) \rangle = \frac{1}{n_\kappa} \delta(\mathbf{x} - \mathbf{y}) \quad (\text{C6})$$

with the Dirac delta distribution $\delta(\mathbf{x} - \mathbf{y})$.

Therefore we have

$$\sigma_{M,\kappa}^2 = \frac{2\pi}{n_\kappa} \int_0^\infty dx x^{2n+1} w(x)^2 \quad (\text{C7})$$

and thus the standard deviation

$$\sigma_{M,\kappa} = \sqrt{\frac{2\pi}{n_\kappa} \left(\int_0^\infty dx x^{2n+1} w(x)^2 \right)^{\frac{1}{2}}}. \quad (\text{C8})$$

Before calculating the shear and flexion moments, we make the following observation. In the case of a discrete distribution over an area A and no lensing, all our measurements of the shear are uncorrelated and we have

$$\frac{1}{n_\gamma^2} \sum_{i=1}^N \sum_{j=1}^N \langle \gamma_i \gamma_j \rangle = \frac{1}{n_\gamma^2} \sum_{i=1}^N \sigma_\gamma^2 = \frac{A}{n_\gamma} \sigma_\gamma^2, \quad (\text{C9})$$

$$\langle \gamma_i \gamma_j \rangle = \sigma_\gamma^2 \delta_{ij} \quad (\text{C10})$$

with the number density $n_\gamma = N_\gamma/A$, where N_γ is the number of sources where we measured γ . Extending this to integration, we have

$$\iint dx^2 dy^2 \langle \gamma(\mathbf{x}) \gamma(\mathbf{y}) \rangle = \frac{1}{n_\gamma} \int dx^2 \sigma_\gamma^2 = \frac{A}{n_\gamma} \sigma_\gamma^2 \quad (\text{C11})$$

and thus

$$\langle \gamma(\mathbf{x}) \gamma(\mathbf{y}) \rangle = \frac{\sigma_\gamma^2}{n_\gamma} \delta(\mathbf{x} - \mathbf{y}) \quad (\text{C12})$$

with the Dirac delta distribution $\delta(\mathbf{x} - \mathbf{y})$.

The errors on κ_{rec} and g_t generally depend on the measurement technique. However, we are assuming the absence of a lens for the calculation of the standard deviation, so we can look at the mass moment in terms of γ instead of κ_{rec} and g_t and we neglect the measurement noise. Therefore we have $\langle M_\gamma^{(n)} \rangle = 0$ and for the variance:

$$\sigma_{M,\gamma}^2 = \langle |M_\gamma^{(n)}|^2 \rangle \quad (\text{C13})$$

$$= \int_0^\infty \int_0^\infty dx dy [2W(x) - x^{n+1}w(x)] [2W(y) - y^{n+1}w(y)] \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle \gamma_t(x, \varphi) \gamma_t(y, \varphi') \rangle. \quad (\text{C14})$$

We use

$$\langle \gamma_t(x, \varphi) \gamma_t(y, \varphi') \rangle = \frac{\sigma_\gamma^2}{n_\gamma} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \quad (\text{C15})$$

where we have written the Dirac delta distribution in polar coordinates, and get

$$\sigma_{M,\gamma}^2 = \int_0^\infty dx \frac{1}{x} [2W(x) - x^{n+1}w(x)]^2 \int_0^{2\pi} d\varphi \frac{\sigma_\gamma^2}{n_\gamma} \quad (\text{C16})$$

and for the standard deviation

$$\sigma_{M,\gamma} = \sqrt{\frac{2\pi}{n_\gamma}} \sigma_\gamma \left(\int_0^\infty dx \frac{1}{x} [2W(x) - x^{n+1}w(x)]^2 \right)^{\frac{1}{2}}. \quad (\text{C17})$$

Similarly, we have $\langle M_{\mathcal{F}}^{(n)} \rangle = 0$ and

$$\sigma_{M,\mathcal{F}}^2 = \langle |M_{\mathcal{F}}^{(n)}|^2 \rangle \quad (\text{C18})$$

$$= \int_0^\infty \int_0^\infty dx dy x W(x) y W(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle \mathcal{F}_t(x, \varphi) \mathcal{F}_t(y, \varphi') \rangle. \quad (\text{C19})$$

Using

$$\langle \mathcal{F}_t(x, \varphi) \mathcal{F}_t(y, \varphi') \rangle = \frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \quad (\text{C20})$$

we have

$$\sigma_{M,\mathcal{F}}^2 = \int_0^\infty dx x W(x)^2 \int_0^{2\pi} d\varphi \frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \quad (\text{C21})$$

and consequently

$$\sigma_{M,\mathcal{F}} = \sqrt{\frac{2\pi}{n_{\mathcal{F}}}} \sigma_{\mathcal{F}} \left(\int_0^\infty dx x W(x)^2 \right)^{\frac{1}{2}}. \quad (\text{C22})$$

In the same way, we get $\langle M_{\mathcal{F},\mathcal{G}}^{(n)} \rangle = 0$ and

$$\sigma_{M,\mathcal{F},\mathcal{G}}^2 = \langle |M_{\mathcal{F},\mathcal{G}}^{(n)}|^2 \rangle \quad (\text{C23})$$

$$= \int_0^\infty \int_0^\infty dx dy \left[V(x) - \frac{1}{2}xW(x) \right] \left[V(y) - \frac{1}{2}yW(y) \right] \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle \mathcal{F}_t(x, \varphi)\mathcal{F}_t(y, \varphi') \rangle + \langle \mathcal{F}_t(x, \varphi)\mathcal{G}_t(y, \varphi') \rangle + \langle \mathcal{G}_t(x, \varphi)\mathcal{F}_t(y, \varphi') \rangle + \langle \mathcal{G}_t(x, \varphi)\mathcal{G}_t(y, \varphi') \rangle). \quad (\text{C24})$$

We use equation (C20) and

$$\langle \mathcal{G}_t(x, \varphi)\mathcal{G}_t(y, \varphi') \rangle = \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \frac{1}{y} \delta(x-y)\delta(\varphi-\varphi') \quad (\text{C25})$$

and assume that

$$\langle \mathcal{F}_t(x, \varphi)\mathcal{G}_t(y, \varphi') \rangle = 0 \quad (\text{C26})$$

to get

$$\sigma_{M,\mathcal{F},\mathcal{G}}^2 = \int_0^\infty dx \frac{1}{x} \left[V(x) - \frac{1}{2}xW(x) \right]^2 \int_0^{2\pi} d\varphi \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}} \right) \quad (\text{C27})$$

and the standard deviation

$$\sigma_{M,\mathcal{F},\mathcal{G}} = \sqrt{2\pi} \sqrt{\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{G}}^2}{n_{\mathcal{G}}}} \left(\int_0^\infty dx \frac{1}{x} \left[V(x) - \frac{1}{2}xW(x) \right]^2 \right)^{\frac{1}{2}}. \quad (\text{C28})$$

C2 Reduced shear, reduced flexion, and K moments

We have

$$\langle M_K^{(n)} \rangle = \int_R^\infty dx x^{n+1} w(x) \int_0^{2\pi} d\varphi \langle K(x, \varphi) \rangle. \quad (\text{C29})$$

In the absence of lensing, κ is typically everywhere smaller than 1, so we can use the Mercator series,

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n \text{ for } |x| < 1, \quad (\text{C30})$$

to find

$$\langle K(x, \varphi) \rangle = (-1) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (-1)^n \langle \kappa(x, \varphi)^n \rangle. \quad (\text{C31})$$

As κ is small, we can restrict ourselves to the two lowest order terms,

$$\langle \kappa(x, \varphi) \rangle = 0, \quad (\text{C32})$$

$$\langle \kappa(x, \varphi)\kappa(x, \varphi) \rangle = \frac{1}{n_{\kappa} a}, \quad (\text{C33})$$

where we again average the sources over the area a , and get

$$\langle K(x, \varphi) \rangle = \frac{1}{2n_{\kappa} a}. \quad (\text{C34})$$

Therefore, we have

$$\langle M_K^{(n)} \rangle = \frac{\pi}{n_{\kappa} a} \int_R^\infty dx x^{n+1} w(x). \quad (\text{C35})$$

The variance is

$$\sigma_{M,K}^2 = \langle |M_K^{(n)}|^2 \rangle - \langle M_K^{(n)} \rangle^2. \quad (\text{C36})$$

We have

$$\langle |M_K^{(n)}|^2 \rangle = \int_R^\infty \int_R^\infty dx dy x^{n+1} y^{n+1} w(x)w(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle K(x, \varphi)K(y, \varphi') \rangle \quad (\text{C37})$$

and

$$\langle K(x, \varphi)K(y, \varphi') \rangle = \langle \ln(1 - \kappa(x, \varphi)) \ln(1 - \kappa(y, \varphi')) \rangle \quad (\text{C38})$$

and using again the Mercator series, we get

$$\langle K(x, \varphi)K(y, \varphi') \rangle = \left\langle \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (-1)^n \kappa^n(x, \varphi) \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} (-1)^m \kappa^m(y, \varphi') \right\rangle \quad (\text{C39})$$

$$= \left\langle \sum_{n=1}^{\infty} \frac{\kappa^n(x, \varphi)}{n} \sum_{m=1}^{\infty} \frac{\kappa^m(y, \varphi')}{m} \right\rangle. \quad (\text{C40})$$

As κ is small, we can ignore the higher order terms to get

$$\langle K(x, \varphi)K(y, \varphi') \rangle \approx \langle \kappa(x, \varphi)\kappa(y, \varphi') \rangle = \frac{1}{n_\kappa} \frac{1}{y} \delta(x-y)\delta(\varphi-\varphi'). \quad (\text{C41})$$

Thus we have

$$\sigma_{M,K}^2 = \frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 a^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2, \quad (\text{C42})$$

$$\sigma_{M,K} = \left(\frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 a^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2 \right)^{\frac{1}{2}}. \quad (\text{C43})$$

The variance is typically well behaved. However, as the first term depends on the source density and the second on the number of sources over which we average to obtain the convergence, $n_\kappa a$, it is theoretically possible to construct unreasonable combinations. For example, obtaining the convergence from only one source per bin while having a high source density would lead to an unreasonable result, as the contribution from the poorly constrained expectation value would dominate the other uncertainties. Naturally, such a combination would be avoided in real applications.

As we treat the case of no lensing, $\langle \kappa \rangle = 0$ and we can thus average over the whole area A to obtain our convergence estimate. Thus we have

$$\sigma_{M,K} = \left(\frac{2\pi}{n_\kappa} \int_R^\infty dx x^{2n+1} w(x)^2 - \frac{\pi^2}{n_\kappa^2 A^2} \left(\int_R^\infty dx x^{n+1} w(x) \right)^2 \right)^{\frac{1}{2}}. \quad (\text{C44})$$

We have $\langle M_g^{(n)} \rangle = 0$ and

$$\begin{aligned} \sigma_{M,g}^2 &= \langle |M_g^{(n)}|^2 \rangle \\ &= \int_R^\infty \int_R^\infty dx dy x W_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle + \langle g_t(x, \varphi) F_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle \\ &\quad + \langle g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle + \langle g_r(x, \varphi) F_r(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle) \\ &\quad + \int_R^\infty \int_R^\infty dx dy x W_R(x) (2W_R(y) - y^{n+1} w(y)) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') \rangle + \langle g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') \rangle) \\ &\quad + \int_R^\infty \int_R^\infty dx dy y W_R(y) (2W_R(x) - x^{n+1} w(x)) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle g_t(y, \varphi') F_t(y, \varphi') g_t(x, \varphi) \rangle + \langle g_r(y, \varphi') F_r(y, \varphi') g_t(x, \varphi) \rangle) \\ &\quad + \int_R^\infty \int_R^\infty dx dy (2W_R(x) - x^{n+1} w(x)) (2W_R(y) - y^{n+1} w(y)) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle g_t(x, \varphi) g_t(y, \varphi') \rangle. \end{aligned} \quad (\text{C46})$$

In the absence of lensing, we expect the flexion and the shear to be uncorrelated. Therefore, we have, using the relation shown in Goodman (1960),

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle \quad (\text{C47})$$

$$= (\langle g_t(x, \varphi) g_t(x, \varphi) \rangle \langle F_t(x, \varphi) F_t(x, \varphi) \rangle + \langle g_t(x, \varphi) g_t(x, \varphi) \rangle \langle F_t(x, \varphi) \rangle^2) \quad (\text{C48})$$

$$+ \langle F_t(x, \varphi) F_t(x, \varphi) \rangle \langle g_t(x, \varphi) \rangle^2 \frac{1}{n_{\gamma, \mathcal{F}}} \frac{1}{y} \delta(x-y)\delta(\varphi-\varphi') \quad (\text{C49})$$

$$= \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\gamma, \mathcal{F}}} \frac{1}{y} \delta(x-y)\delta(\varphi-\varphi'), \quad (\text{C50})$$

where $n_{\nu, \mathcal{F}}$ is the number density of sources for which we have both shear and flexion information. As flexion is typically much harder to measure than shear, we can make the approximation $n_{\nu, \mathcal{F}} \approx n_{\mathcal{F}}$. Using this approximation and making a similar calculation in the other cases, we have

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle = \frac{\sigma_{\mathcal{F}}^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x-y) \delta(\varphi - \varphi'), \quad (\text{C51})$$

$$\langle g_r(x, \varphi) F_r(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle = \frac{\sigma_{\mathcal{F}}^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x-y) \delta(\varphi - \varphi'), \quad (\text{C52})$$

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle = \langle g_t(x, \varphi) \rangle \langle g_r(y, \varphi') \rangle \langle F_t(x, \varphi) \rangle \langle F_r(y, \varphi') \rangle = 0, \quad (\text{C53})$$

$$\langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') \rangle = \langle g_t(x, \varphi) g_t(y, \varphi') \rangle \langle F_t(x, \varphi) \rangle = 0, \quad (\text{C54})$$

$$\langle g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') \rangle = \langle g_r(x, \varphi) g_t(y, \varphi') \rangle \langle F_r(x, \varphi) \rangle = 0, \quad (\text{C55})$$

and thus

$$\sigma_{M,g}^2 = \int_R^\infty dx x W_R(x)^2 \int_0^{2\pi} d\varphi 2 \frac{\sigma_{\mathcal{F}}^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \int_R^\infty dx \frac{1}{x} (2W_R(x) - x^{n+1}w(x))^2 \int_0^{2\pi} d\varphi \frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \quad (\text{C56})$$

and we have the standard deviation

$$\sigma_{M,g} = \left(\frac{4\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \sigma_{\mathcal{F}}^2 \left(\int_R^\infty dx x W_R(x)^2 \right) + \frac{2\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \int_R^\infty dx \frac{1}{x} (2W_R(x) - x^{n+1}w(x))^2 \right)^{\frac{1}{2}}. \quad (\text{C57})$$

We have $\langle M_F^{(n)} \rangle = 0$ and

$$\sigma_{M,F}^2 = \langle |M_F^{(n)}|^2 \rangle \quad (\text{C58})$$

$$= \int_R^\infty \int_R^\infty dx dy x W_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' \langle F_t(x, \varphi) F_t(y, \varphi') \rangle \quad (\text{C59})$$

and using

$$\langle F_t(x, \varphi) F_t(y, \varphi') \rangle = \frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \frac{1}{y} \delta(x-y) \delta(\varphi - \varphi') \quad (\text{C60})$$

we have

$$\sigma_{M,F}^2 = \frac{2\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \int_R^\infty dx x W_R(x)^2 \quad (\text{C61})$$

and the standard deviation

$$\sigma_{M,F} = \sqrt{\frac{2\pi}{n_{\mathcal{F}}} \sigma_{\mathcal{F}}^2 \left(\int_R^\infty dx x W_R(x)^2 \right)^{\frac{1}{2}}}. \quad (\text{C62})$$

We have $\langle M_{F,G}^{(n)} \rangle = 0$ and

$$\sigma_{M,F,G}^2 = \langle |M_{F,G}^{(n)}|^2 \rangle \quad (\text{C63})$$

$$\begin{aligned} &= \int_R^\infty \int_R^\infty dx dy x W_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle g_r(x, \varphi) F_r(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle \\ &\quad + \frac{1}{2} \langle g_r(x, \varphi) F_r(x, \varphi) F_t(y, \varphi') \rangle + \frac{1}{2} \langle g_r(x, \varphi) F_r(x, \varphi) G_t(y, \varphi') \rangle + \frac{1}{2} \langle F_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle \\ &\quad + \frac{1}{2} \langle G_t(x, \varphi) g_r(y, \varphi') F_r(y, \varphi') \rangle + \frac{1}{4} \langle F_t(x, \varphi) F_t(y, \varphi') \rangle + \frac{1}{4} \langle F_t(x, \varphi) G_t(y, \varphi') \rangle \\ &\quad + \frac{1}{4} \langle G_t(x, \varphi) F_t(y, \varphi') \rangle + \frac{1}{4} \langle G_t(x, \varphi) G_t(y, \varphi') \rangle) \\ &\quad - \int_R^\infty \int_R^\infty dx dy x W_R(x) V_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle g_r(x, \varphi) F_r(x, \varphi) F_t(y, \varphi') \rangle \\ &\quad + \langle g_r(x, \varphi) F_r(x, \varphi) G_t(y, \varphi') \rangle - 2 \langle g_r(x, \varphi) F_r(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle \\ &\quad + \frac{1}{2} \langle F_t(x, \varphi) F_t(y, \varphi') \rangle + \frac{1}{2} \langle F_t(x, \varphi) G_t(y, \varphi') \rangle - \langle F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle) \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2} \langle G_t(x, \varphi) F_t(y, \varphi') \rangle + \frac{1}{2} \langle G_t(x, \varphi) G_t(y, \varphi') \rangle - \langle G_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle \\
 & - \int_R^\infty \int_R^\infty dx dy V_R(x) y W_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle g_r(y, \varphi') F_r(y, \varphi') F_t(x, \varphi) \rangle \\
 & + \langle g_r(y, \varphi') F_r(y, \varphi') G_t(x, \varphi) \rangle - 2 \langle g_r(y, \varphi') F_r(y, \varphi') g_t(x, \varphi) F_t(x, \varphi) \rangle \\
 & + \frac{1}{2} \langle F_t(y, \varphi') F_t(x, \varphi) \rangle + \frac{1}{2} \langle F_t(y, \varphi') G_t(x, \varphi) \rangle - \langle F_t(y, \varphi') g_t(x, \varphi) F_t(x, \varphi) \rangle \\
 & + \frac{1}{2} \langle G_t(y, \varphi') F_t(x, \varphi) \rangle + \frac{1}{2} \langle G_t(y, \varphi') G_t(x, \varphi) \rangle - \langle G_t(y, \varphi') g_t(x, \varphi) F_t(x, \varphi) \rangle \\
 & + \int_R^\infty \int_R^\infty dx dy V_R(x) V_R(y) \int_0^{2\pi} \int_0^{2\pi} d\varphi d\varphi' (\langle F_t(x, \varphi) F_t(y, \varphi') \rangle + \langle F_t(x, \varphi) G_t(y, \varphi') \rangle \\
 & - 2 \langle F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle + \langle G_t(x, \varphi) F_t(y, \varphi') \rangle + \langle G_t(x, \varphi) G_t(y, \varphi') \rangle \\
 & - 2 \langle G_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle - 2 \langle g_t(x, \varphi) F_t(x, \varphi) F_t(y, \varphi') \rangle \\
 & - 2 \langle g_t(x, \varphi) F_t(x, \varphi) G_t(y, \varphi') \rangle + 4 \langle g_t(x, \varphi) F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle).
 \end{aligned} \tag{C64}$$

We again assume that the first and second flexions and the shear are uncorrelated in the absence of lensing, thus we can use the previously derived relations and we also have

$$\langle G_t(x, \varphi) G_t(y, \varphi') \rangle = \frac{\sigma_G^2}{n_G} \frac{1}{y} \delta(x - y) \delta(\varphi - \varphi'), \tag{C65}$$

$$\langle g_r(x, \varphi) F_r(x, \varphi) F_t(y, \varphi') \rangle = \langle g_r(x, \varphi) F_r(x, \varphi) G_t(y, \varphi') \rangle = \langle F_t(x, \varphi) G_t(y, \varphi') \rangle = 0, \tag{C66}$$

$$\langle F_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle = \langle g_t(y, \varphi') \rangle \langle F_t(x, \varphi) F_t(y, \varphi') \rangle = 0, \tag{C67}$$

$$\langle G_t(x, \varphi) g_t(y, \varphi') F_t(y, \varphi') \rangle = 0. \tag{C68}$$

This gives us

$$\begin{aligned}
 \sigma_{M,F,G}^2 = 2\pi \left(\frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{F}}^2}{4n_{\mathcal{F}}} + \frac{\sigma_G^2}{4n_G} \right) \int_R^\infty dx x W_R(x)^2 - 2\pi \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_G^2}{n_G} \right) \int_R^\infty dx W_R(x) V_R(x) \\
 + 2\pi \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_G^2}{n_G} + 4 \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \right) \int_R^\infty dx \frac{1}{x} V_R(x)^2,
 \end{aligned} \tag{C69}$$

$$\begin{aligned}
 \sigma_{M,F,G} = \sqrt{2\pi} \left(\left(\frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_{\mathcal{F}}^2}{4n_{\mathcal{F}}} + \frac{\sigma_G^2}{4n_G} \right) \int_R^\infty dx x W_R(x)^2 - \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_G^2}{n_G} \right) \int_R^\infty dx W_R(x) V_R(x) \right. \\
 \left. + \left(\frac{\sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} + \frac{\sigma_G^2}{n_G} + 4 \frac{\sigma_\gamma^2 \sigma_{\mathcal{F}}^2}{n_{\mathcal{F}}} \right) \int_R^\infty dx \frac{1}{x} V_R(x)^2 \right)^{\frac{1}{2}}.
 \end{aligned} \tag{C70}$$

APPENDIX D: CONVERGENCE, SHEAR, AND FLEXION FOR NFW PROFILES

Navarro, Frenk, and White showed that spherically averaged cold dark matter (CDM) halo density profiles can be fitted over two decades in radius by the NFW density profile (Navarro et al. 1996, 1997)

$$\frac{\rho(r)}{\rho_{\text{crit}}} = \frac{\delta_c}{(r/r_s)(1+r/r_s)^2}, \tag{D1}$$

where r_s is a scale radius, δ_c is a characteristic dimensionless density, and $\rho_{\text{crit}} = 3H^2/8\pi G$ is the critical density for closure. This profile leads to the convergence formula (Bartelmann 1996; Wright & Brainerd 2000)

$$\kappa_{\text{NFW}}(y) = \frac{2\rho_{\text{crit}}\delta_c r_s}{\Sigma_{\text{crit}}} \frac{f(y)}{y^2 - 1}, \tag{D2}$$

where $y = \xi/r_s$ and

$$f(y) = \begin{cases} 1 - \frac{2}{\sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right), & y < 1 \\ \frac{y^2-1}{3}, & y = 1 \\ 1 - \frac{2}{\sqrt{y^2-1}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{y+1}} \right), & y > 1. \end{cases} \quad (\text{D3})$$

The NFW profile is spherically symmetric and therefore $\gamma_{r,\text{NFW}} = 0$ (see e.g. Meylan et al. 2006). The tangential shear is (Wright & Brainerd 2000)

$$\gamma_{t,\text{NFW}} = \frac{\rho_{\text{crit}} \delta_c r_s}{\Sigma_{\text{crit}}} g(y), \quad (\text{D4})$$

where

$$g(y) = \begin{cases} \frac{8}{y^2 \sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right) + \frac{4}{y^2} \log \left(\frac{y}{2} \right) - \frac{2}{y^2-1} + \frac{4}{(y^2-1)\sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right), & y < 1 \\ \frac{10}{3} + 4 \log \left(\frac{1}{2} \right), & y = 1 \\ \frac{8}{y^2 \sqrt{y^2-1}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{1+y}} \right) + \frac{4}{y^2} \log \left(\frac{y}{2} \right) - \frac{2}{y^2-1} + \frac{4}{(y^2-1)^{3/2}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{1+y}} \right), & y > 1. \end{cases} \quad (\text{D5})$$

We can use the expressions for first and second flexion for a NFW profile derived in Bacon et al. (2006) to obtain

$$\mathcal{F}_{t,\text{NFW}} = \frac{2\rho_{\text{crit}} \delta_c D_{\text{ol}}}{\Sigma_{\text{crit}}(y^2-1)^2} [2yf(y) - h(y)], \quad (\text{D6})$$

$$\mathcal{G}_{t,\text{NFW}} = -\frac{2\rho_{\text{crit}} \delta_c D_{\text{ol}}}{\Sigma_{\text{crit}}} \left[\frac{8}{y^3} \log \left(\frac{y}{2} \right) + \frac{3(1-2y^2)/y + l(y)}{(y^2-1)^2} \right], \quad (\text{D7})$$

where

$$h(y) = \begin{cases} \frac{2y}{\sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right) - \frac{1}{y}, & y < 1 \\ \frac{2y(y^2-1)}{3} - 0.4(y^2-1)^2, & y = 1 \\ \frac{2y}{\sqrt{y^2-1}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{y+1}} \right) - \frac{1}{y}, & y > 1, \end{cases} \quad (\text{D8})$$

$$l(y) = \begin{cases} \left(\frac{8}{y^3} - \frac{20}{y} + 15y \right) \frac{2}{\sqrt{1-y^2}} \operatorname{arctanh} \left(\sqrt{\frac{1-y}{1+y}} \right), & y < 1 \\ \frac{1}{y} \left(\frac{94}{15} y(y^2-1)^2 - 3(1-2y^2) \right), & y = 1 \\ \left(\frac{8}{y^3} - \frac{20}{y} + 15y \right) \frac{2}{\sqrt{y^2-1}} \operatorname{arctan} \left(\sqrt{\frac{y-1}{y+1}} \right), & y > 1. \end{cases} \quad (\text{D9})$$

The radial flexion components are both zero.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.

5 Improving lens models with spectroscopic redshifts

5.1 Photometric vs spectroscopic redshifts

There are two different methods to determine the redshift of an observed image. The first is the so-called “photometric redshift” technique. It uses the brightness of an image in different filters and thus at different colors to compute the most likely redshift. There are two main techniques for computing the photo- z , machine learning and template fitting techniques. The latter take a set of physically motivated spectral energy distributions and find the best match to the data (see e.g. [Brinchmann et al., 2017](#), for a comparison of photo- z from fitting techniques to spectroscopic redshifts). The second method to determine an image redshift is the “spectroscopic redshift” method. It uses the observed spectrum of an image to identify spectral features, for example a Lyman- α emission line, and subsequently compares the observed wavelength of the feature with the rest-frame wavelength measured in the laboratory to compute the redshift. As a result, the spec- z s are extremely reliable and have tiny errors, while the photo- z s are less accurate. While photo- z algorithms are improving (e.g. [Molino et al., 2017](#); [Brinchmann et al., 2017](#)), they can occasionally be completely wrong, leading to so-called “catastrophic errors”. However, spectroscopy and thus spec- z s require a lot of observing time while photo- z s can be easily computed from multi-band surveys. As a result, the number of objects with photo- z is much too large to be efficiently followed up with spectroscopy (e.g., [Brinchmann et al., 2017](#)). Consequently, astronomers have to distinguish between scenarios which benefit greatly from the availability of spec- z s and those which would only marginally improve and allocate the spectroscopy time accordingly.

Cluster lensing benefits greatly from safe spectroscopic redshifts, as demonstrated e.g. in [Johnson et al. \(2014\)](#), who show that the magnification μ can vary beyond statistical errors when redshift information is added for a single system, and in [Jullo et al. \(2010\)](#) and [Caminha et al. \(2016\)](#), who require secure redshifts for each multiple image family used in their cluster cosmography analysis.

Last year, [Johnson and Sharon \(2016\)](#) published a detailed study of the impact of spec-

troscopic redshifts on cluster lens modeling. They found that the fit of the resulting lens models to the data improves most efficiently if more multiple images with spec- z are added. In addition, they demonstrate that the availability of secure redshifts is of particular importance during the construction phase of the lens model. During this period, lens modelers often rely on preliminary mass models to predict the locations of new multiple images and thus to identify new sets. As [Johnson and Sharon \(2016\)](#) show, at least 10 spec- z s are needed to distinguish between multiple image candidates based on their model-predicted location. In general, it is most important that at least a single reliable redshift is used when modeling any cluster lens. If no redshift at all is known, then the lensing mass is degenerate with redshift (e.g., [Johnson and Sharon, 2016](#)).

Let us now assume that we have a lens model with a fixed set of multiple image systems. If we add spec- z s for each of these systems, how does the quality of the magnification prediction change? As [Johnson and Sharon \(2016\)](#) show, this significantly improves the accuracy of the prediction. This finding holds as long as less than half of the used multiple image families already have secure redshifts. Once this fraction is satisfied, the effect of adding more spec- z s is much less pronounced. Another interesting question is whether it makes a difference for the accuracy of the magnification prediction if we add a new multiple image system with spectroscopic redshift or one without. The answer is that it will improve in a similar way as long as we already have at least one spec- z included in the model ([Johnson and Sharon, 2016](#)). In addition, [Johnson and Sharon \(2016\)](#) find that adding spec- z s can improve the accuracy of the mass profile, but only if the model did not have a single spectroscopic redshift before.

Recently, [Acebron et al. \(2017\)](#) have shown that the inference of cosmological parameters from cluster cosmography can be biased if only a restricted redshift range is available from using spectroscopic redshifts. They investigate if the cosmology estimation can be improved by including photometric redshifts covering a broad redshift range. They find that adding this complimentary data set to their spectroscopic redshifts indeed reduces the bias.

In a nutshell, cluster lens models benefit significantly if at least a few secure spectroscopic redshifts are available. However, there is a plethora of spectrographs available and each has its own advantages and downsides. Which one is best suited for cluster lensing? Luckily, there is a simple answer to this question: MUSE. We will describe the reasons for this answer in section 5.2.2. Before doing so, let us take a closer look at the MUSE instrument.

5.2 MUSE: A spectroscopic redshift machine

5.2.1 The MUSE instrument

The Multi Unit Spectroscopic Explorer (MUSE) is a panoramic Integral-Field Spectrograph (IFS) for the visible wavelength range installed on the unit telescope number four of the Very Large Telescope (VLT) in the Chilean desert, see figure 5.1. It has a wide field-of-view of

5.2. MUSE: A spectroscopic redshift machine

1×1 arcmin² sampled at a pixel size of 0.2×0.2 arcsec² (Bacon et al., 2010b). The instrument consists of 24 identical high performance integral-field units. Each unit is composed of an advanced image slicer, a spectrograph, and a high-resolution detector (Bacon et al., 2010b). The MUSE provides medium resolution spectroscopy over the full 4750 – 9350 Å domain with a dispersion of 1.25 Å per pixel and a resolution increasing from $R \approx 2000$ to $R \approx 4000$ between the blue and the red parts of the observed spectrum (e.g., Richard et al., 2015). The resulting datacube, which provides a full spectrum for each pixel in the image, is shown in figure 5.2.

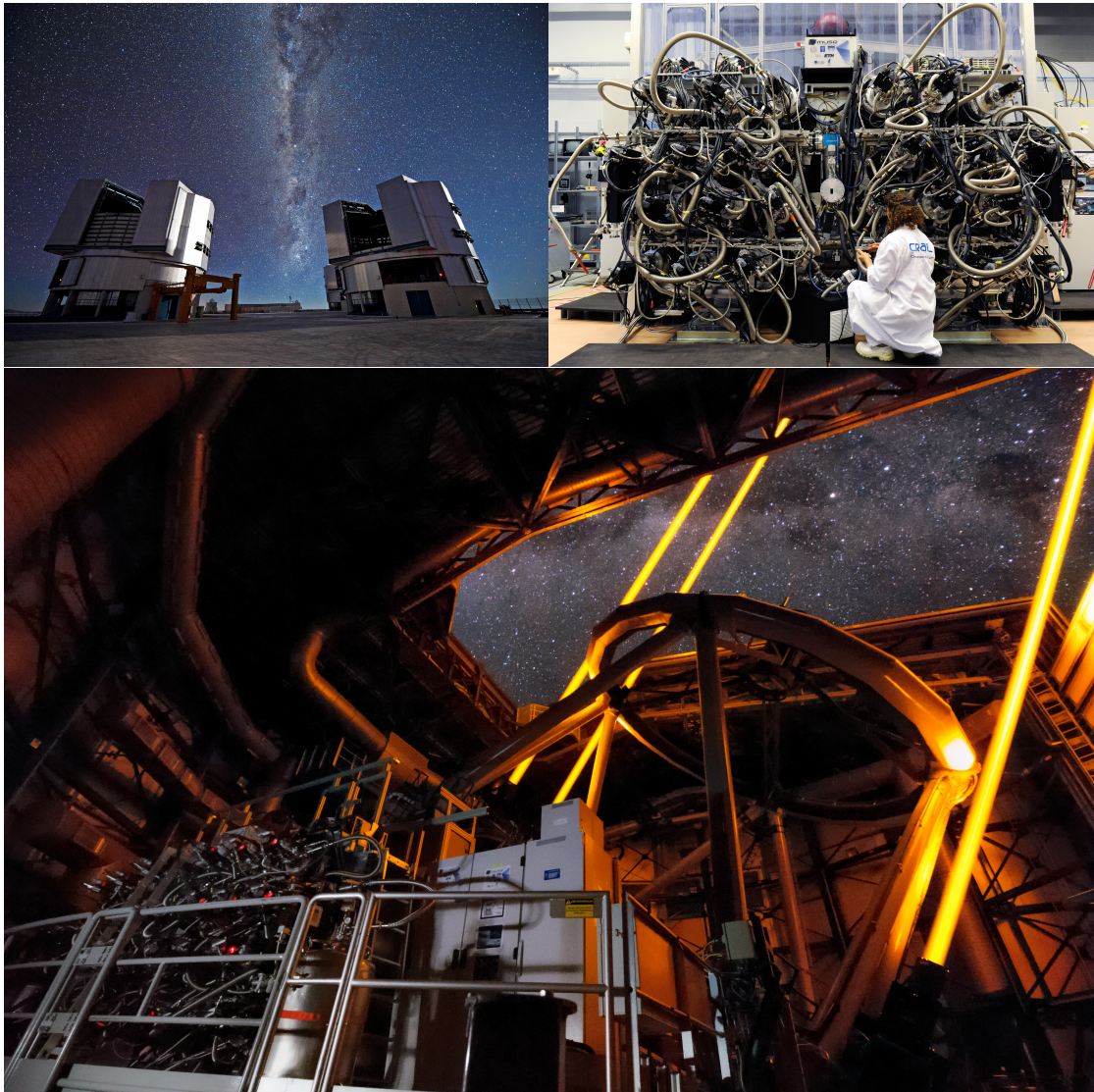


Figure 5.1 – One of the four unit telescopes of the VLT (top left image) is equipped with the MUSE instrument (top right image). Bottom image: MUSE on the VLT at work. The lasers of the recently installed adaptive optics system are clearly visible. (Image credit: ESO, see photographic credits)

The raw data produced by MUSE must be processed by a data reduction pipeline, which

Chapter 5. Improving lens models with spectroscopic redshifts

produces the final data cube. The pipeline is the result of a substantial software development effort and due to the size and the quality of the MUSE raw data, it is recommended to run the pipeline on a machine with 24 physical Central Processing Unit (CPU) cores, 64 GB of Random Access Memory (RAM), and 4 TB of available hard disk space. Additional software has been developed by the MUSE consortium to fully exploit the high-quality data products, for example a principal component analysis based sky subtraction tool named Zurich Atmosphere Purge (ZAP) (Soto et al., 2016), an advanced data reduction pipeline (Conseil et al., 2016), and a tool which allows a detailed analysis of the produced datacubes, the MUSE Python Data Analysis Framework (MPDAF) (Piqueras et al., 2017).

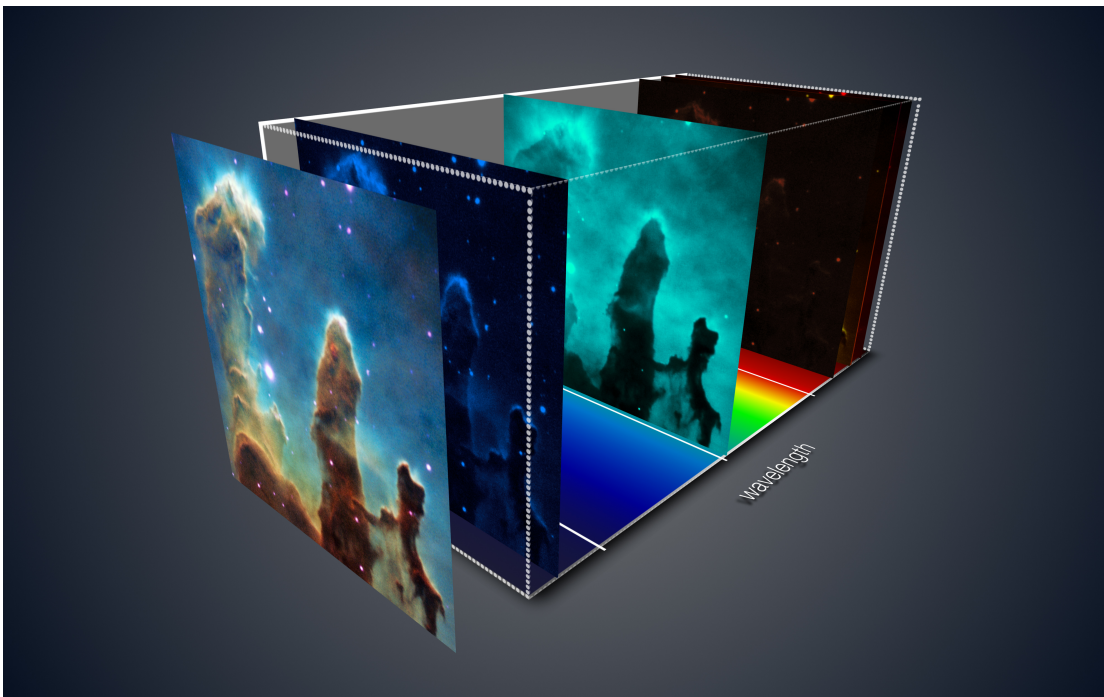


Figure 5.2 – A datacube produced by the IFS MUSE. Each pixel in the image has full spectroscopic information for the visible wavelengths. As a result, we can view the image in different colors. The signal strength of the image will vary with the wavelength and thus we can investigate the chemical composition of the object. (Image credit: ESO)

However, at the time when we obtained the first raw data from our MUSE cluster lensing program described in the next section, MPDAF was not publicly available. The use of this software was restricted to the MUSE consortium. We thus faced the challenge that we had obtained fantastic data from a brand-new instrument, but we had no means to exploit it. As a result, we have developed a public, open source software named Integral-Field Spectrograph Redshift Extractor (IFS-RedEx) (Rexroth et al., 2017). It permits a quick and efficient extraction of redshifts from IFS datacubes. In addition, it allows the user to combine the IFS data with high resolution data, e.g. from HST, to easily detect and use faint sources which might otherwise be missed. A key feature of the software is the wavelet-based spectrum cleaner, which identifies reliable spectrum features and reconstructs their shapes

while removing the spectrum noise. A test shows that it can detect spectrum peaks down to a signal-to-noise of 8 while reporting no fake detections. We met with Roland Bacon, the Principal Investigator of MUSE, during his visit to the Geneva observatory and we were asked if we consented to the inclusion of our wavelet cleaner into MPDAF, which had been publicly released in the meantime. We were glad to do so and as a result, the wavelet cleaner presented in this thesis is now part of the official data analysis framework of MUSE, MPDAF. We will present the IFS-RedEx publication in section 5.3. Before we do this, however, we first present our MUSE observing program of cluster lenses in section 5.2.2. An improved lens model from combining MUSE and HST data, which is the topic of Rexroth et al. (2017b, in preparation), is presented in section 5.4.

5.2.2 MUSE observations of cluster lenses

The typical magnitudes for cluster-lensed sources found with HST are $\approx 23 - 25$ AB (e.g., Richard et al., 2009; Jauzac et al., 2014). As a result, deep spectroscopy is required to obtain reliable spectra and thus spectroscopic redshifts (Richard et al., 2015). Large campaigns with multi-object spectrographs have been conducted to determine spec- z s for lensed images in galaxy cluster lenses (e.g., Johnson et al., 2014; Grillo et al., 2015). However, they had to rely on image pre-selection based on deep HST multiband data and they are inefficient since all multiple images are typically crowded in the cluster core (Richard et al., 2015).

MUSE (Bacon et al., 2010b) changes this situation dramatically. It is evident from the instrument specifications in the last section that the MUSE field-of-view perfectly catches the central area of most SL clusters, where most of the multiple images are hidden (Richard et al., 2015). MUSE can identify faint emission lines of high equivalent width, in particular Lyman- α between $z = 2.8$ and $z = 6.7$ (Richard et al., 2015). For example, Richard et al. (2015) show that MUSE helps to identify many multiple images based on their spec- z and that these would not have been found with short exposure HST imaging. The impressive redshift capabilities of the instrument are also demonstrated in e.g. Bacon et al. (2015), who use it to observe the HST Deep Field South and find 189 redshifts in the 1 arcmin² field-of-view in a long total exposure time of 27 hours. This increased the number of known spec- z by an order of magnitude. In addition, MUSE also facilitates the detection and identification of cluster members. It is important for the quality of the lens model to include all cluster members in the modeling process, as these can affect the multiple image locations (Harvey et al., 2016).

The potential of MUSE for observations of SL clusters was further demonstrated in recent publications which present improved lens models (Mahler et al., 2017; Lagattuta et al., 2017; Caminha et al., 2017). For example, Richard et al. (2015) increase the precision of the measured cumulative mass at the best constrained radius by a factor of five compared to the previous lens model before MUSE data was available. Jauzac et al. (2016b) and Grillo et al. (2016) use lens models from HST and spectroscopic data, including MUSE, to predict the re-appearance of the strongly lensed supernova “Refsdal” in the galaxy cluster

MACS J1149.6+2223.

Due to this potential, we applied for a significant amount of observing time with MUSE and our program “MUSE critical line mapping of the HST best studied cluster lenses” (Principal Investigator: Jean-Paul Kneib) was accepted. Unfortunately, the observing conditions have not been ideal over the last semesters, and thus the data was slow to arrive. Complete observational data are available for four galaxy clusters, namely MACS J1931.8-2635 (henceforth called MACSJ1931), MACS J2129.4-0741, RX J2129.7+0005, and MS 0451-03. The data reduction has been completed for all clusters and we have extracted the redshifts for each cluster with our IFS-RedEx software. Subsequently, we have combined HST imaging with the MUSE data to model the lenses. The results for MACSJ1931 will be presented in Rexroth et al. (2017b, in preparation) and are detailed in section 5.4. The redshift distributions and first lens model results for the other three clusters have been obtained by Baptiste Klein, who wrote his Master’s thesis under our supervision (Klein, 2017), and will be presented in Klein et al. (in preparation).

5.3 IFS-RedEx: A redshift extraction pipeline with wavelet filtering

The following paper, [Rexroth et al. \(2017\)](#), presents the IFS-RedEx pipeline for Integral-Field Spectrograph datacubes.

IFS-RedEx, a redshift extraction software for integral-field spectrographs: Application to MUSE data

Markus Rexroth¹, Jean-Paul Kneib^{1,2}, Rémy Joseph¹, Johan Richard³, Romaric Her^{1,4}

¹ *Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland*

² *Aix Marseille Université, CNRS, LAM (Laboratoire d’Astrophysique de Marseille) UMR 7326, 13388, Marseille, France*

³ *Univ Lyon, Univ Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France*

⁴ *ISAE-SUPAERO, Université de Toulouse, 10 Avenue Edouard Belin, 31400 Toulouse, France*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present IFS-RedEx, a spectrum and redshift extraction pipeline for integral-field spectrographs. A key feature of the tool is a wavelet-based spectrum cleaner. It identifies reliable spectral features, reconstructs their shapes, and suppresses the spectrum noise. This gives the technique an advantage over conventional methods like Gaussian filtering, which only smears out the signal. As a result, the wavelet-based cleaning allows the quick identification of true spectral features. We test the cleaning technique with degraded MUSE spectra and find that it can detect spectrum peaks down to $S/N \approx 8$ while reporting no fake detections. We apply IFS-RedEx to MUSE data of the strong lensing cluster MACSJ1931.8-2635 and extract 54 spectroscopic redshifts. We identify 29 cluster members and 22 background galaxies with $z \geq 0.4$. IFS-RedEx is open source and publicly available.

Key words: Techniques: Imaging spectroscopy – Techniques: Image processing – Galaxies: clusters: individual: MACSJ1931.8-2635 – Galaxies: high-redshift

1 INTRODUCTION

Astrophysical research has benefited greatly from publicly available open source software and programs like SExtractor (Bertin & Arnouts 1996) and Astropy (Astropy Collaboration et al. 2013) have become standard tools for many astronomers. Their public availability allows researchers to focus on the science and to reduce the programming overhead, while the open source nature facilitates the code’s further development and adaptation. In this spirit, we developed the Integral-Field Spectrograph Redshift Extractor (IFS-RedEx), an open source software for the efficient extraction of spectra and redshifts from integral-field spectrographs¹. The software can also be used as a complement to other tools such as the Multi Unit Spectroscopic Explorer (MUSE) Python Data Analysis Framework (mpdaf)².

Our redshift extraction tool includes a key feature, a wavelet-based spectrum cleaning tool which removes spurious peaks and reconstructs a cleaned spectrum. Wavelet

transformations are well suited for astrophysical image and data processing (see e.g. Starck & Murtagh 2006 for an overview) and have been successfully applied to a variety of astronomical research projects. To name only a few recent examples, wavelets have been used for source deblending (Joseph et al. 2016), gravitational lens modeling (Lanusse et al. 2016) and the removal of contaminants to facilitate the detection of high redshift objects (Livermore et al. 2016).

The paper is designed as follows: Sections 2 and 3 present the spectrum and redshift extraction routines of IFS-RedEx. In section 4, we describe and test the wavelet-based spectrum cleaning tool. In section 5, we illustrate the use of our software by applying it to MUSE data of the strong lensing cluster MACSJ1931.8-2635 (henceforth called MACSJ1931). We summarize our results in section 6.

2 SPECTRUM EXTRACTION & CATALOG CLEANING

It is advantageous to combine Integral-Field Unit (IFU) data cubes with high resolution imaging, as this allows us to detect small, faint sources which might remain undetected if we used the image obtained by collapsing the data cube

¹ The software can be downloaded at <http://lastro.epfl.ch/software>

² Available at <https://git-cral.univ-lyon1.fr/MUSE/mpdaf>

along the wavelength axis (henceforth called white-light image) for source detection. For example, [Bacon et al. \(2015\)](#) used this combination in their analysis of MUSE observations of the *Hubble* Deep Field South. Therefore we exploit this case in the following, but in principle the software can be used without high resolution data. IFS-RedEx uses the center positions of stars provided by the user to align the IFU and high resolution images. It utilizes a SExtractor ([Bertin & Arnouts 1996](#)) catalog of the high-resolution data to extract the spectra and the associated standard deviation noise estimate for each source from the data cube. It extracts the signal in an area with a radius of 3 to 5 data cube pixels, depending on the SExtractor full width at half maximum (FWHM) estimate. Sources with FWHM < 2 high resolution pixels are discarded as these are typically spurious detections, e.g. due to cosmic rays.

IFS-RedEx shows the user each source and extraction radius overplotted on the high resolution image and the IFU data cube. The user can now quickly examine each detection and decide to either keep it in the database or to remove it, for example because it is too close to the data cube boundary and suffers from edge effects.

The tool also supports line emission and continuum emission catalogs. These are for example created by the MUSELET³ software, which uses narrow-band images to perform a blind search for the respective signal. IFS-RedEx displays the detected sources and their extraction radius of 3 pixels on the IFU data cube. The user labels sources which cannot be used, e.g. because the signal is only a spurious detection in one pixel or it is too close to the image boundary. The spectra and noise of the good sources are automatically extracted.

Finally, the cleaned SExtractor, line emission, and continuum emission catalogs are merged into a master catalog. In this step, the sources are displayed on the high-resolution image so that the user can decide if the MUSELET and SExtractor detections are part of the same source. This visual inspection is more reliable than an automatic association and the number of sources is typically small enough for a manual inspection in reasonable time.

3 REDSHIFT EXTRACTION

Each 1D spectrum is displayed in an interactive plot and a second window shows the corresponding high resolution image, see figure 1. The position of sky lines with a flux $\geq 50 \times 10^{-20} \text{ ergs}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$ are labeled in green. The sky line fluxes are taken from [Cosby et al. \(2006\)](#). IFS-RedEx also lists the emission line identifications from MUSELET if available.

The user can now adjust the position of the emission and absorption line template by changing the source

³ MUSELET is part of the mpdaf package. A tutorial and the documentation are available at <http://mpdaf.readthedocs.io/en/latest/muselet.html>

redshift. Once the template matches the source spectrum, the right redshift is found. IFS-RedEx has several features to facilitate the correct identification of spectral features. The user can zoom in and out, overplot the noise on the spectrum, smooth the signal with a Gaussian filter and perform a wavelet-based spectrum cleaning, see figure 1. When IFS-RedEx plots the noise, it shows the standard deviation around an offset. The offset is calculated by smoothing the spectrum signal with a Gaussian with $\sigma = 100$ pixels. Thus the noise is centered on the smoothed signal and it follows signal drifts. The wavelet cleaning is described in detail in the next section. As can be seen in figure 1, it reconstructs the shape of the reliable spectrum features and suppresses the noise. The Gaussian filter only smears out the signal. Thus the wavelet-based reconstruction makes it easier to distinguish true from false peaks.

Finally, the user can fit a Gaussian to the most prominent spectral line. IFS-RedEx combines the error of the fitted center position with the wavelength calibration error from the IFU data reduction pipeline into the final statistical redshift error. The software creates a final catalog with all source redshifts and errors. In addition, it produces a document with all spectral feature identifications and high resolution images for later use, e.g. for verification by a colleague.

4 WAVELET-BASED SPECTRUM RECONSTRUCTION

4.1 Wavelet transform algorithms

The wavelet-based cleaning algorithm reconstructs only spectral features above a given significance threshold. For this purpose, we use the “à trous” wavelet transform with a B_3 -spline scaling function of the coordinate $x \in \mathbb{R}$,

$$\phi(x) = \frac{1}{12} (|x-2|^3 - 4|x-1|^3 + 6|x|^3 - 4|x+1|^3 + |x+2|^3), \quad (1)$$

which is well suited for isotropic signals such as emission lines ([Starck et al. 2007](#); [Starck & Murtagh 2006](#); [Holschneider et al. 1989](#)). In contrast to a Fourier transform, wavelets possess both frequency and location information. We note that the measured spectrum signal is discrete and not continuous and we denote the unprocessed, noisy spectrum data \mathbf{c}_0 , where the subscript indicates the scale s , and its value at pixel position l with $c_{0,l}$. We assume that $c_{0,l}$ is the scalar product of the continuous spectrum function $f(x)$ and $\phi(x)$ at pixel l . Now we can filter this data, where each filtering step increases s by one and leads to \mathbf{c}_{s+1} , which no longer includes the highest frequency information from \mathbf{c}_s . The filtered data for each scale is calculated by using a convolution. The coefficients of the convolution mask \mathbf{h} derive from the scaling function,

$$\frac{1}{2}\phi\left(\frac{x}{2}\right) = \sum_l h(l)\phi(x-l), \quad (2)$$

and they are (1/16, 1/4, 3/8, 1/4, 1/16) ([Starck & Murtagh 2006](#)). By noting that $h(k)$ is symmetric ([Starck et al. 2007](#)), we have

$$c_{s,l} = \sum_k h(k)c_{s-1,l+2^{s-1}k} \quad (3)$$

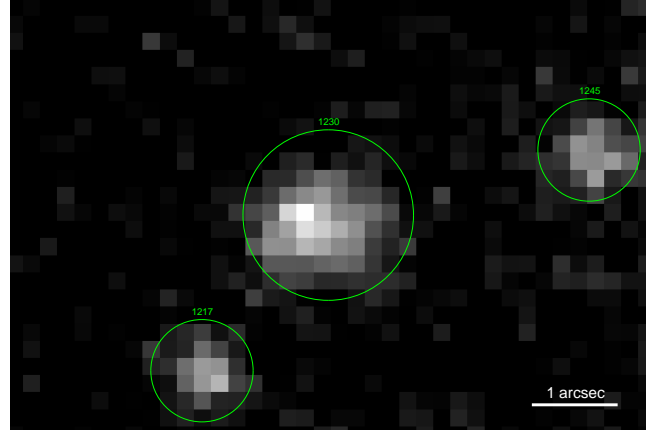
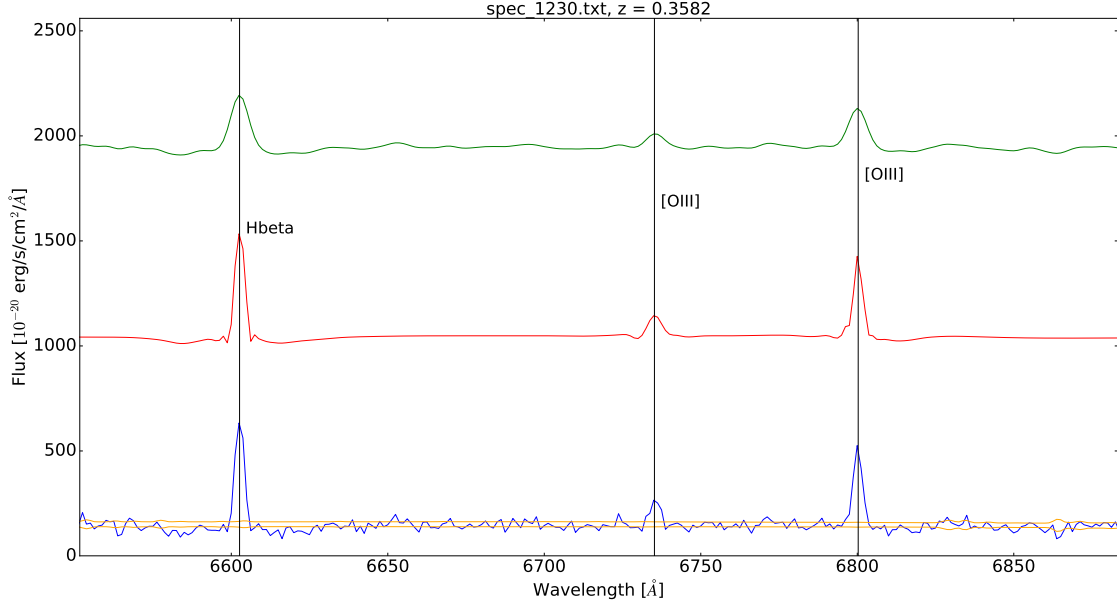


Figure 1. Top: Interactive spectrum plot. The user can apply Gaussian filtering (green), wavelet cleaning (red) and plot the noise (yellow) to distinguish real from spurious features of the data (blue). The offsets of the plots can be adjusted. The noise shows the upper and lower standard deviation around the smoothed signal (see text). Bottom left: IFS-RedEx displays the high resolution color image for each source to facilitate the redshift extraction. The respective source is always at the image center and labeled with the number of the source’s spectrum file, here 1230. Bottom right: MUSE data cube slice at 6799.97 Angstrom corresponding to the high resolution image on the left. In the spectrum fitting step the data cube is typically not needed and thus it is not displayed by default, but it can be quickly loaded via the DS9 interface if required.

and we define the double-convolved data on the same scale by

$$cd_{s,l} = \sum_k h(k)c_{s,l+2^{s-1}k}. \quad (4)$$

The wavelet coefficients are now given by

$$w_{s,l} = c_{s-1,l} - cd_{s,l}, \quad (5)$$

and they include the information between these two scales (Starck et al. 2016). A low scale s implies high frequencies and vice versa. The final wavelet transform is the set $\{\mathbf{w}_1, \dots, \mathbf{w}_L, \mathbf{c}_L\}$, where L is the highest scale level we use, and it includes the full spectrum information. We impose an upper limit for L depending on the spectrum wavelength

range and resolution: $L \leq \log_2((P-1)/(H-1))$, where P is the number of pixels of the spectrum signal and H the length of \mathbf{h} , which is in our case $H = 5$. Otherwise s could become so large that the filtering equation 3 would require data outside of the wavelength range. We compute the wavelet transform according to algorithm 1 and we transform back into real space by using algorithm 2 (Starck et al. 2016).

The cleaning in wavelet space is performed following Starck & Murtagh (2006): We transform a discretized Dirac δ -distribution to obtain the wavelet set $\{\mathbf{w}_1^\delta, \dots, \mathbf{w}_L^\delta\}$. Subsequently, we convolve each squared \mathbf{w}_s^δ with the squared standard deviation spectrum noise extracted from

Algorithm 1 Transform the spectrum into wavelet space

Require: Spectrum \mathbf{c}_0 (= set of discrete spectrum pixels $\{c_{0,l}\}$), highest scale level L , convolution mask \mathbf{h}

Output: Wavelet transform of spectrum $\{\mathbf{w}_1, \dots, \mathbf{w}_L, \mathbf{c}_L\}$

- 1: **Procedure** WAVELET_TRANSFORM(\mathbf{c}_0, L):
- 2: $s \leftarrow 0$
- 3: **while** $s < L$ **do**
- 4: $s \leftarrow s + 1$
- 5: $c_{s,l} \leftarrow \sum_k h(k) c_{s-1, l+2^{s-1}k} \forall l$
- 6: $cd_{s,l} \leftarrow \sum_k h(k) c_{s, l+2^{s-1}k} \forall l$
- 7: $w_{s,l} \leftarrow c_{s-1, l} - cd_{s,l} \forall l$
- 8: **end while**
- 9: **return** $\{\mathbf{w}_1, \dots, \mathbf{w}_L, \mathbf{c}_L\}$

the IFU data cube and take the square root of the result. This gives us the noise coefficients in wavelet space.

In the next step, we build the multiresolution support \mathbf{M} , which is a $(L+1) \times P$ matrix. We compare the absolute value of the signal and noise wavelet coefficients at each pixel, $w_{s,l}$ and $w_{s,l}^N$. We take a threshold T set by the user, for example 5 for a 5σ cleaning in wavelet space, and set the corresponding matrix entry in \mathbf{M} to 1 if $|w_{s,l}| \geq T|w_{s,l}^N|$, and 0 otherwise. Note that for $s = 1$, we use a higher threshold of $T + 1$, as this wavelet scale corresponds to high frequencies, where we expect the noise to dominate. The matrix coefficients for the smoothed signal \mathbf{c}_L are automatically set to 1.

Now we perform the cleaning: We set all $w_{s,l}$ associated with a vanishing \mathbf{M} value to zero and transform back into real space to obtain a first clean spectrum. However, there is still some signal to be harnessed in the residuals. Therefore we subtract the clean spectrum from the full spectrum to obtain the residual spectrum, and we compare its standard deviation, σ_{res} , with the standard deviation of the full spectrum (in the first iteration) or of the residual used in the previous iteration (all subsequent iterations), which we indicate in both cases with σ_{prev} . If $|(\sigma_{\text{prev}} - \sigma_{\text{res}})/\sigma_{\text{res}}| > \epsilon$, we transform the residual spectrum into wavelet space, set wavelets with vanishing \mathbf{M} values to zero, transform back into real space, and add the resulting signal to obtain our new clean signal. Note that the same multiresolution support as before is used. Subsequently, we calculate again the residual and continue until the ϵ criterion is no longer fulfilled and all the signal has been extracted. The value of ϵ is set by the user and must satisfy the condition $0 < \epsilon < 1$. Algorithm 3 summarizes this cleaning procedure.

4.2 Testing the wavelet-based reconstruction

To test our software, we use the spectrum of the brightest cluster galaxy (BCG) from our MUSE data set described in the next section. MUSE provides both the spectrum signal and a noise estimate over the full wavelength range. The original spectrum can be considered clean due to its

Algorithm 2 Transformation from wavelet to real space

Require: Wavelet transform of spectrum $\{\mathbf{w}_1, \dots, \mathbf{w}_L, \mathbf{c}_L\}$, highest scale level L , number of spectrum pixels P , convolution mask \mathbf{h}

Output: Spectrum in real space \mathbf{c}_0 (= set of discrete spectrum pixels $\{c_{0,l}\}$)

- 1: **Procedure** WAVELET_BACKTRANSFORM($\{\mathbf{w}_1, \dots, \mathbf{c}_L\}$):
- 2: $\mathbf{S} \leftarrow \mathbf{c}_L$
- 3: **for all** $s \in \{1, \dots, L\}$ **do**
- 4: **for all** $l \in \{1, \dots, P\}$ **do**
- 5: $C_l \leftarrow \sum_k h(k) S_{l+2^{L-s}k}$
- 6: **end for**
- 7: $\mathbf{S} \leftarrow \mathbf{C} + \mathbf{w}_{L+1-s}$
- 8: **end for**
- 9: $\mathbf{c}_0 \leftarrow \mathbf{S}$
- 10: **return** \mathbf{c}_0

Algorithm 3 Signal cleaning in wavelet space

Require: Spectrum \mathbf{c}_0 (= set of discrete spectrum pixels $\{c_{0,i}\}$), σ_{spec} (= vector with standard deviation noise estimate for each spectrum pixel), highest scale level L , number of spectrum pixels P , cleaning threshold T , cleaning parameter ϵ ($0 < \epsilon < 1$)

Output: Cleaned spectrum $\mathbf{S}_{\text{clean}}$

- 1: **Procedure** CLEAN_SIGNAL($\mathbf{c}_0, \sigma_{\text{spec}}, L, T, \epsilon$):
- 2: $\{\mathbf{w}_1^\delta, \dots, \mathbf{w}_L^\delta, \mathbf{c}_L^\delta\} \leftarrow \text{WAVELET_TRANSFORM}(\delta\text{-dist.}, L)$
- 3: **for all** $\mathbf{w}_s^\delta \in \{\mathbf{w}_1^\delta, \dots, \mathbf{w}_L^\delta\}$ **do**
- 4: $w_{s,l}^N \leftarrow \sqrt{w_{s,l}^{\delta 2} * \sigma_{\text{spec}}^2}$
- 5: **end for**
- 6: $\{\mathbf{w}_1, \dots, \mathbf{w}_L, \mathbf{c}_L\} \leftarrow \text{WAVELET_TRANSFORM}(\mathbf{c}_0, L)$
- 7: $\mathbf{M} \leftarrow \mathbf{0}_{L+1, P}$ // Multiresolution support matrix
- 8: **for all** $s \in \{1, \dots, L+1\}, l \in \{1, \dots, P\}$ **do**
- 9: **if** $s == 1$ **and** $|w_{s,l}| \geq (T+1)|w_{s,l}^N|$ **then**
- 10: $M_{sl} \leftarrow 1$
- 11: **else if** $1 < s \leq L$ **and** $|w_{s,l}| \geq T|w_{s,l}^N|$ **then**
- 12: $M_{sl} \leftarrow 1$
- 13: **else if** $s == L+1$ **then**
- 14: $M_{sl} \leftarrow 1$
- 15: **end if**
- 16: **end for**
- 17: $\mathbf{S}_{\text{clean}} \leftarrow \mathbf{0}_P, \sigma_{\text{prev}} \leftarrow 0, \text{res} \leftarrow \mathbf{c}_0$
- 18: $\sigma_{\text{res}} \leftarrow \text{std}(\text{res})$
- 19: **while** $|(\sigma_{\text{prev}} - \sigma_{\text{res}})/\sigma_{\text{res}}| > \epsilon$ **do**
- 20: $\{\mathbf{w}_1^{\text{res}}, \dots, \mathbf{c}_L^{\text{res}}\} \leftarrow \text{WAVELET_TRANSFORM}(\text{res}, L)$
- 21: **for all** $s \in \{1, \dots, L\}, l \in \{1, \dots, P\}$ **do**
- 22: **if** $M_{sl} == 0$ **then**
- 23: $w_{s,l}^{\text{res}} \leftarrow 0$
- 24: **end if**
- 25: **end for**
- 26: $\text{res}_{\text{clean}} \leftarrow \text{WAVELET_BACKTRANSFORM}(\{\mathbf{w}_1^{\text{res}}, \dots, \mathbf{c}_L^{\text{res}}\})$
- 27: $\mathbf{S}_{\text{clean}} \leftarrow \mathbf{S}_{\text{clean}} + \text{res}_{\text{clean}}$
- 28: $\text{res} \leftarrow \mathbf{c}_0 - \mathbf{S}_{\text{clean}}$
- 29: $\sigma_{\text{prev}} \leftarrow \sigma_{\text{res}}$
- 30: $\sigma_{\text{res}} \leftarrow \text{std}(\text{res})$
- 31: **end while**
- 32: **return** $\mathbf{S}_{\text{clean}}$

very high signal-to-noise. We rescale it to simulate fainter sources at low signal-to-noise. We calculate the rescaling factor R by looking at the highest spectrum signal peak and dividing the associated MUSE noise estimate by this signal. This results in $R \approx 0.0015$. We investigate three cases, namely a good, an intermediate, and a low signal-to-noise case, where we rescale the full signal spectrum by $10R$, $5R$, and $2R$ respectively. Subsequently we add Gaussian noise simulating the real noise estimate of the MUSE data cube. For each spectral wavelength pixel l we obtain the realized noise by drawing from a Gaussian probability distribution with a standard deviation equal to the MUSE standard deviation noise estimate at this pixel. We repeat this process 10 times to obtain spectra with different noise realizations.

We calculate the signal-to-noise of six emission lines by summing over their respective wavelength ranges,

$$\frac{S}{N} = \frac{\sum_l \text{signal}_l}{\sqrt{\sum_{l'} \text{std}_{l'}^2}}, \quad (6)$$

where std_l is the MUSE standard deviation noise estimate at pixel l . We will refer to the lines according to their wavelength order, i.e. the first line is situated at the lowest wavelength and the last line at the highest. We apply our wavelet cleaning software to the spectra using the MUSE noise estimate and different wavelet parameters as input. We investigate 5σ and 3σ cleaning and ϵ parameters of 0.1, 0.01, and 0.001. The cleaning procedure is fast and takes about 1 second per spectrum on a laptop. Figure 2 shows reconstructed spectra for the three different signal-to-noise cases. Note that the last two emission lines in the true spectrum are actually comprised of merged individual lines. As can be seen in figure 2, the wavelet tool can detect if a line consists of two merged lines and reconstruct them correctly if their signal-to-noise is high enough. If it is too low, it will reconstruct them as a single line.

For all 90 spectra which we analyzed with a 5σ wavelet reconstruction, we find no fake detections of emission lines. For signal-to-noise larger than 20, all 6 test emission lines are detected. For S/N between 10 and 20, all emission lines but the third are found. The third peak is no longer recovered due to its proximity to the fourth peak, which has typically a twice larger S/N value. In general, the wavelet software might reconstruct two close-by peaks as a single peak unless they have each a sufficiently large signal. When the signal-to-noise of both peaks was similar, both the third and the fourth emission line were detected and reconstructed. For emission lines with low signal-to-noise values between 5 and 10, we can reconstruct the stronger lines with $S/N \gtrsim 8$, while the weaker peaks remain typically undetected. However, as the bottom plot in figure 2 shows, even weaker peaks can occasionally be reconstructed.

Emission lines modeled with a wavelet reconstruction do sometimes not reach the full peak height of the signal, in particular for high ϵ values, and their tails can suffer from ringing effects which might be due to the wavelet shape, see for example the first emission line of the intermediate S/N case in figure 2. For low signal-to-noise emission lines

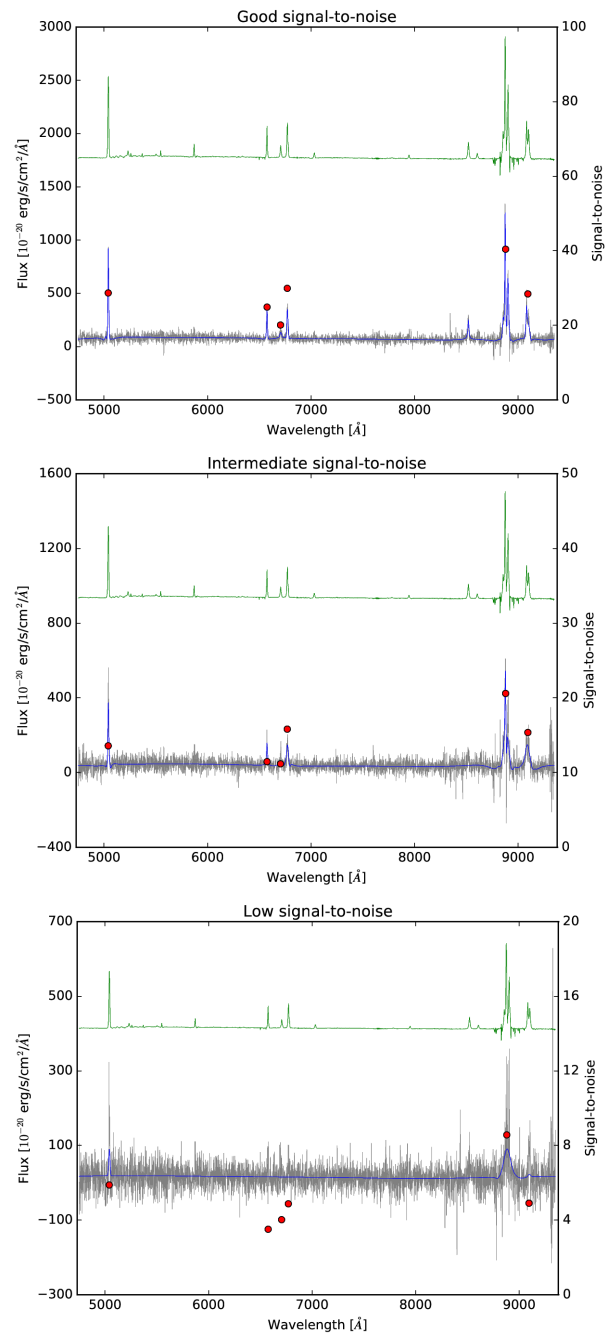


Figure 2. Reconstructed spectra for three different signal-to-noise scenarios using 5σ cleaning and $\epsilon = 0.01$. The true spectrum is offset and shown in green, the noisy spectrum is displayed in gray and the reconstructed signal in blue. Red dots indicate the signal-to-noise of the respective emission lines. The wavelet tool detects peaks with $S/N \gtrsim 8$. However, if a low S/N emission line is located very close to a high S/N line, it is possible that it will not be recovered (middle plot, 3rd emission line from the left). The wavelet reconstruction can occasionally even find lines with $S/N < 8$ (bottom plot, 1st and 6th emission line), but typically they will not be detected (bottom plot, 2nd - 4th emission line).

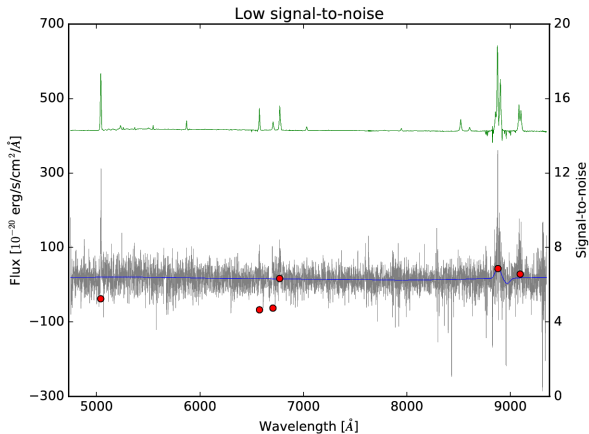


Figure 3. In low signal-to-noise ($S/N \leq 10$) cases, ringing effects can occasionally lead to signal dips with similar amplitude as the signal peaks of the reconstructed emission line. Therefore care has to be taken not to mistake these effects for absorption lines. This might be ameliorated by re-running the wavelet reconstruction with a different setup. The colors have the same meaning as in figure 2 and the reconstruction was performed with a 5σ cutoff and $\epsilon = 0.01$.

($S/N \leq 10$), care has therefore to be taken not to mistake the signal dip due to ringing effects as an absorption signal, as the ringing effect might occasionally have a similar (negative) amplitude as the signal peak of the reconstructed emission line, see figure 3. When this effect occurs in practice, it might be improved by changing the wavelet setup, e.g. by lowering the ϵ value. A lower ϵ is designed to detect a larger fraction of the signal peak and should thus increase its height. However, care has to be taken as a lower ϵ might also lead to stronger ringing effects.

The 3σ wavelet reconstruction recovered more emission lines than the 5σ cleaning, but it also produced false detections. We therefore adopted a conservative approach and used the 5σ wavelet cleaning when applying the code to real data.

Finally, we compared the noise free emission line shapes with the reconstructed ones. We find that the shape reconstruction is generally good, but the reconstructed line shape and height recovered from the noisy data can differ from the original, clean ones, in particular in low signal-to-noise scenarios. Therefore we use the wavelet cleaning only to distinguish true from false spectrum peaks, and we perform all data operations such as fitting a Gaussian to obtain the centering error on the real, noisy data.

5 APPLICATION TO MUSE DATA: MACSJ1931

We apply IFS-RedEx to our data set of the strong lensing cluster MACSJ1931 obtained with MUSE (Bacon et al. 2010) on the Very Large Telescope (VLT). We combine our data with the publicly available *Hubble Space Telescope* (HST) imaging from the Cluster Lensing And Supernova survey with Hubble (CLASH, Postman et al. 2012). The

cluster is part of the MAssive Cluster Survey (MACS), which comprises more than one hundred highly X-ray luminous clusters (Ebeling et al. 2010, 2001).

The core of MACSJ1931 ($z = 0.35$) was observed with MUSE on June 12 and July 17 2015 (ESO program 095.A-0525(A), PI: Jean-Paul Kneib). The 1×1 arcmin² field of view was pointed at $\alpha = 19:31:49.66$ and $\delta = -26:34:34.0$ (J2000) and we observed for a total exposure time of 2.44 hours, divided into 6 exposures of 1462 seconds each. We rotated the second exposure of each exposure pair by 90 degrees to allow for cosmic ray rejection and improve the overall image quality. The data were taken using the WFM-NOAO-N mode of MUSE in good seeing conditions with FWHM ≈ 0.7 arcseconds.

We reduced the data using the MUSE pipeline version 1.2.1 (Weilbacher et al. 2014, 2012), which includes bias and flat-field corrections, sky subtraction, and wavelength and flux calibrations. The six individual exposures were finally combined into a single data cube and we subtracted the remaining sky residuals with ZAP (Soto et al. 2016). The wavelength range of the data cube stretches from 4750 to 9351 Å in steps of 1.25 Å. The spatial pixel size is 0.2 arcseconds.

We used the HST data for MACSJ1931 obtained as part of the CLASH program (Zitrin et al. 2015) in the bands F105W, F475W, F625W, and F814W with a spatial sampling of 0.03 arcsec/pixel. The HST data products are publicly available on the CLASH website⁴.

We use only redshift identifications which we consider secure because we see e.g. several lines or a clear Ly α emission line shape. We extract 54 sources with redshifts ranging from 0.21 to 5.8. Among them, 29 are cluster members with $0.3419 \leq z \leq 0.3672$ and 22 are background sources with $0.4 \leq z \leq 5.8$. A table of all sources with spectroscopic redshifts is presented in the companion paper Rexroth et al. 2017 (in preparation), in which we use the data to improve the cluster lens model. Figure 4 shows a histogram of the source distribution in redshift space.

6 SUMMARY

We describe IFS-RedEx, a public spectrum and redshift extraction pipeline for integral-field spectrographs. The software supports SExtractor catalogs as well as MUSELET narrow-band detection catalogs as input. The pipeline has several features which allow a quick identification of reliable spectrum features, most notably a wavelet-based spectrum cleaning tool. The tool only reconstructs spectral features above a given significance threshold. We test it with degraded MUSE spectra and find that it can detect spectral features with $S/N \gtrsim 8$. We find no fake detections in our test. Finally, we apply IFS-RedEx to a MUSE data cube of the strong lensing cluster MACSJ1931 and extract 54 spectroscopic redshifts.

⁴ <https://archive.stsci.edu/prepds/clash/>

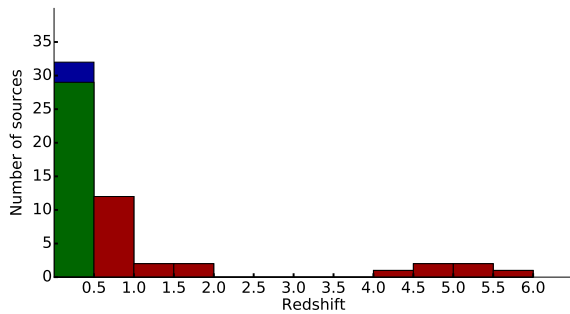


Figure 4. Source distribution in redshift space. Background sources are colored in red, cluster members in green and the remaining objects in blue.

ACKNOWLEDGEMENTS

MR thanks Timothée Delubac for verifying the spectral line identifications and Yves Revaz and the ESO user support center for their help with a non-critical issue in the MUSE pipeline. He thanks Thibault Kuntzer, Pierre North and Frédéric Vogt for fruitful discussions and Anton Koekemoer for his help with processing the Simple Imaging Polynomial (SIP) distortion information from FITS image headers. MR and JPK gratefully acknowledge support from the ERC advanced grant LIDA. RJ gratefully acknowledges support from the Swiss National Science Foundation. JR gratefully acknowledges support from the ERC starting grant 336736-CALENDS. This research made use of SAOImage DS9, numpy (van der Walt et al. 2011), scipy (Jones et al. 2001), matplotlib (Hunter 2007), PyFITS, PyRAF/IRAF (Tody 1986), Astropy (Astropy Collaboration et al. 2013), pyds9, GPL ghostscript, and TeX Live. PyRAF and PyFITS are public software created by the Space Telescope Science Institute, which is operated by AURA for NASA. This research has made use of NASA’s Astrophysics Data System.

REFERENCES

- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
 Bacon R., et al., 2010, in *Ground-based and Airborne Instrumentation for Astronomy III*. p. 773508, doi:10.1117/12.856027
 Bacon R., et al., 2015, *A&A*, **575**, A75
 Bertin E., Arnouts S., 1996, *A&AS*, **117**, 393
 Cosby P. C., Sharpee B. D., Slanger T. G., Huestis D. L., Hanuschik R. W., 2006, *Journal of Geophysical Research (Space Physics)*, **111**, A12307
 Ebeling H., Edge A. C., Henry J. P., 2001, *ApJ*, **553**, 668
 Ebeling H., Edge A. C., Mantz A., Barrett E., Henry J. P., Ma C. J., van Speybroeck L., 2010, *MNRAS*, **407**, 83
 Holschneider M., Kronland-Martinet R., Morlet J., Tchamitchian P., 1989, in *Combes J.-M., Grossmann A., Tchamitchian P., eds, Wavelets. Time-Frequency Methods and Phase Space*. p. 286
 Hunter J. D., 2007, *Computing in Science and Engineering*, **9**, 90
 Jones E., Oliphant T., Peterson P., et al., 2001, *SciPy: Open source scientific tools for Python* [Online; accessed 16.12.2016], <http://www.scipy.org/>
 Joseph R., Courbin F., Starck J.-L., 2016, *A&A*, **589**, A2

- Lanusse F., Starck J.-L., Leonard A., Pires S., 2016, preprint, (arXiv:1603.01599)
 Livermore R. C., Finkelstein S. L., Lotz J. M., 2016, preprint, (arXiv:1604.06799)
 Postman M., et al., 2012, *ApJS*, **199**, 25
 Soto K. T., Lilly S. J., Bacon R., Richard J., Conseil S., 2016, *MNRAS*, **458**, 3210
 Starck J.-L., Murtagh F., 2006, *Astronomical Image and Data Analysis*, doi:10.1007/978-3-540-33025-7.
 Starck J.-L., Fadili J., Murtagh F., 2007, *IEEE Transactions on Image Processing*, **16**, 297
 Starck J.-L., Murtagh F., Fadili J., 2016, *Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis*. Cambridge University Press
 Tody D., 1986, in *Crawford D. L., ed., Proc. SPIE Vol. 627, Instrumentation in astronomy VI*. p. 733
 Weilbacher P. M., Streicher O., Urrutia T., Jarno A., Pécontal-Rousset A., Bacon R., Böhm P., 2012, in *Software and Cyberinfrastructure for Astronomy II*. p. 84510B, doi:10.1117/12.925114
 Weilbacher P. M., Streicher O., Urrutia T., Pécontal-Rousset A., Jarno A., Bacon R., 2014, in *Manset N., Forshay P., eds, Astronomical Society of the Pacific Conference Series Vol. 485, Astronomical Data Analysis Software and Systems XXIII*. p. 451 (arXiv:1507.00034)
 Zitrin A., et al., 2015, *ApJ*, **801**, 44
 van der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science and Engineering*, **13**, 22

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.

5.4 An improved lens model for the galaxy cluster MACSJ1931

This section describes the paper Rexroth et al. (2017b, in preparation). In this publication, we improve the mass model for the galaxy cluster lens MACSJ1931 by combining archival HST imaging with MUSE spectroscopy. In addition, we present line emission maps for several galaxies.

5.4.1 Introduction

We present results for the SL cluster MACSJ1931 obtained from combining MUSE spectroscopy (PI: Jean-Paul Kneib) and CLASH HST imaging. The cluster is part of the MAssive Cluster Survey (MACS), which comprises more than one hundred highly X-ray luminous clusters (Ebeling et al., 2010, 2001).

The first strong and weak lensing modeling of the cluster was presented in Zitrin et al. (2015). They use photometric redshifts and, when available, spectroscopic ones from CLASH-VLT (Rosati et al., 2014) (see Jouvel et al. (2014) for an analysis of the photometric redshift accuracy). They find 4 multiple image systems and in addition 3 candidate multiple image families, resulting in a total of 22 multiple images which they use for the SL analysis. Their light-traces-mass and PIEMD + eNFW modeling techniques achieve an image reproduction root mean square of 2.28 arcsec and 0.77 arcsec, respectively. The cluster mass and concentration are determined in Umetsu et al. (2016) by combining the results from Zitrin et al. (2015) with the WL and magnification measurements from Umetsu et al. (2014). These parameters are also obtained by Merten et al. (2015) from combining the SL constraints presented in Zitrin et al. (2015) with HST WL observations and ground-based WL data from Umetsu et al. (2014). The cluster's lensing mass estimates are compared in Donahue et al. (2014) with those from X-ray observations.

Our MUSE observations of the central 1×1 arcmin² region of the galaxy cluster result in a substantial number of spectroscopically confirmed cluster members and faint background galaxies. The addition of spectroscopic data can improve lens mass models (see e.g. Richard et al., 2015; Mahler et al., 2017; Lagattuta et al., 2017; Monna et al., 2017; Johnson et al., 2014) and we combine our results with the archival HST imaging to construct an improved model. We find four new multiple image systems and we confirm two of the candidate multiple image families suggested in Zitrin et al. (2015), while we find that the third lies outside of the multiple image area. This leads to a total of ten multiple image systems and a mass model with an image reproduction root mean square of 0.62 arcsec.

In addition, we compute line emission maps of the Brightest Cluster Galaxy (BCG), galaxy cluster members, and a $z = 0.8$ background galaxy. This illustrates the capability of MUSE to simultaneously study galaxy clusters and distant galaxy populations (see e.g. Bina et al., 2016; Karman et al., 2015). Furthermore, it illustrates the promise of MUSE data for studying

5.4. An improved lens model for the galaxy cluster MACSJ1931

ram pressure stripping, “jellyfish” galaxies, and the feeding of SMBH, which has been recently demonstrated e.g. in [Bellhouse et al. \(2017\)](#), [Poggianti et al. \(2017b\)](#), [Sheen et al. \(2017\)](#), and [Poggianti et al. \(2017a\)](#).

5.4.2 MUSE data and reduction

The core of MACSJ1931 ($z = 0.35$) was observed with MUSE on 12 June and 17 July 2015 (ESO program 095.A-0525(A), PI: Jean-Paul Kneib). The 1×1 arcmin² field-of-view was pointed at $\alpha = 19:31:49.66$ and $\delta = -26:34:34.0$ (J2000) and we observed for a total exposure time of 2.44 hours, divided into 6 exposures of 1462 seconds each. We rotated the second exposure of each exposure pair by 90 degrees to allow cosmic ray rejection and improve the overall image quality. The data were taken using the WFM-NOAO-N mode of MUSE in good seeing conditions with Full Width at Half Maximum (FWHM) ≈ 0.7 arcseconds.

We reduce the data using the MUSE pipeline version 1.2.1 ([Weilbacher et al., 2014, 2012](#)), which includes bias and flat-fielding corrections, sky subtraction, and wavelength and flux calibrations. The six individual exposures are combined into a single datacube and we subtract the remaining sky residuals with ZAP ([Soto et al., 2016](#)). The wavelength range of the datacube stretches from 4750 to 9350 Å in steps of 1.25 Å. The spatial pixel size is 0.2 arcseconds.

We use the HST data for MACSJ1931 obtained as part of the CLASH program ([Zitrin et al., 2015](#)) in the bands F105W, F475W, F625W, and F814W with a pixel size of 0.03 arcsec in addition to the MUSE data. The HST data products are publicly available on the CLASH website¹.

5.4.3 Spectrum and redshift extraction

We create a source catalog using the F814W HST image and the associated weights data as inputs for SExtractor ([Bertin and Arnouts, 1996](#)). For the remaining steps, we use an early version of the IFS-RedEx redshift extraction software ([Rexroth et al., 2017](#)). We align the MUSE datacube and the HST image and subsequently we extract the spectra from the MUSE datacube. We extract them in a ring around the detected source with a minimum radius of 3 MUSE pixels and a maximum radius of 5 pixels to avoid blending. IFS-RedEx displays each circle for the spectrum extraction both on the stacked MUSE image and on the HST color image, which we obtain by combining the filters F814W, F625W, and F475W. We check for additional sources in each extraction radius which would contaminate our spectra and exclude these detections from our analysis. Furthermore, we exclude overlapping extraction radii, sources which are too close to the BCG, and sources which are on the image boundary, as edge effects could contaminate the spectra.

¹<https://archive.stsci.edu/prepds/clash/>

Chapter 5. Improving lens models with spectroscopic redshifts

In the next step, we create a second and a third catalog. We use narrow-band images to perform a blind search for isolated emission lines (second catalog) and continuum emission (third catalog). We use the publicly available software MUSELET for this task, which is part of MPDAF (Piqueras et al., 2017). We extract the spectra using a radius of 3 MUSE pixels. We investigate these spectra and the corresponding sources in MUSE and HST color images and we remove sources which are too close to the image border, too close to the BCG, or fake detections due to spurious pixel signals. Finally, we merge the three catalogs into a final catalog.

We fit line templates to the spectrum of every source in the catalog to find the respective redshift. We take a conservative approach and use only secure identifications, e.g. because we see several spectral features. The identification of spectral features is independently confirmed by a colleague (Timothée Delubac). We clean the spectra with the wavelet-based spectrum cleaner, so that we can easily identify reliable lines. We use the wavelet cleaning only to distinguish true from false spectrum peaks and we perform all data operations on the real, noisy data, as recommended in Rexroth et al. (2017). We compute the statistical error by fitting a Gaussian to the most prominent emission or absorption line and use the standard deviation of the center position as error estimate. We combine the error with the standard deviation of the wavelength calibration obtained from the MUSE data reduction pipeline to obtain the final redshift error estimate. We only include statistical errors and do not account for possible systematic effects on the redshift measurement.

5.4.4 Redshifts and line emission maps

We obtain 54 sources with spectroscopic redshifts $0.21 \leq z \leq 5.8$. We compare our sources with the ones in the photometric redshift catalog available on the CLASH website. We identify two sources if their respective right ascension and declination coordinates differ by 3 MUSE pixels (0.6 arcsec) or less. We visually inspect the cross-matched catalog and if one source is identified with two different objects, we delete the redundant object from the catalog. The resulting table with spectroscopic and cross-matched photometric redshifts is presented in the appendix in subsection 5.4.5. Most photometric redshifts are in good agreement with the measured spec- z , but some catastrophic photo- z errors are observed.

Among the sources with spectroscopic redshifts are 29 cluster members with $0.3419 \leq z \leq 0.3672$. In addition, we have 22 background sources with $0.4 \leq z \leq 5.8$. Figure 5.4 shows the detected Lyman- α emission lines for the six background objects with $z > 4$. The color image of the galaxy cluster in figure 5.3 indicates the locations of the cluster member galaxies and the background sources.

We compute the $H\alpha$ line emission map of the BCG in the following way. We calculate the redshifted wavelength which corresponds to the $H\alpha$ rest frame wavelength of 6562.8 \AA at the BCG's redshift of $z = 0.3526$ and we obtain $\lambda = 8876.8 \text{ \AA}$. We examine the BCG spectrum and find that the $H\alpha$ peak stretches from 8862 \AA to 8884 \AA . Subsequently, we stack the

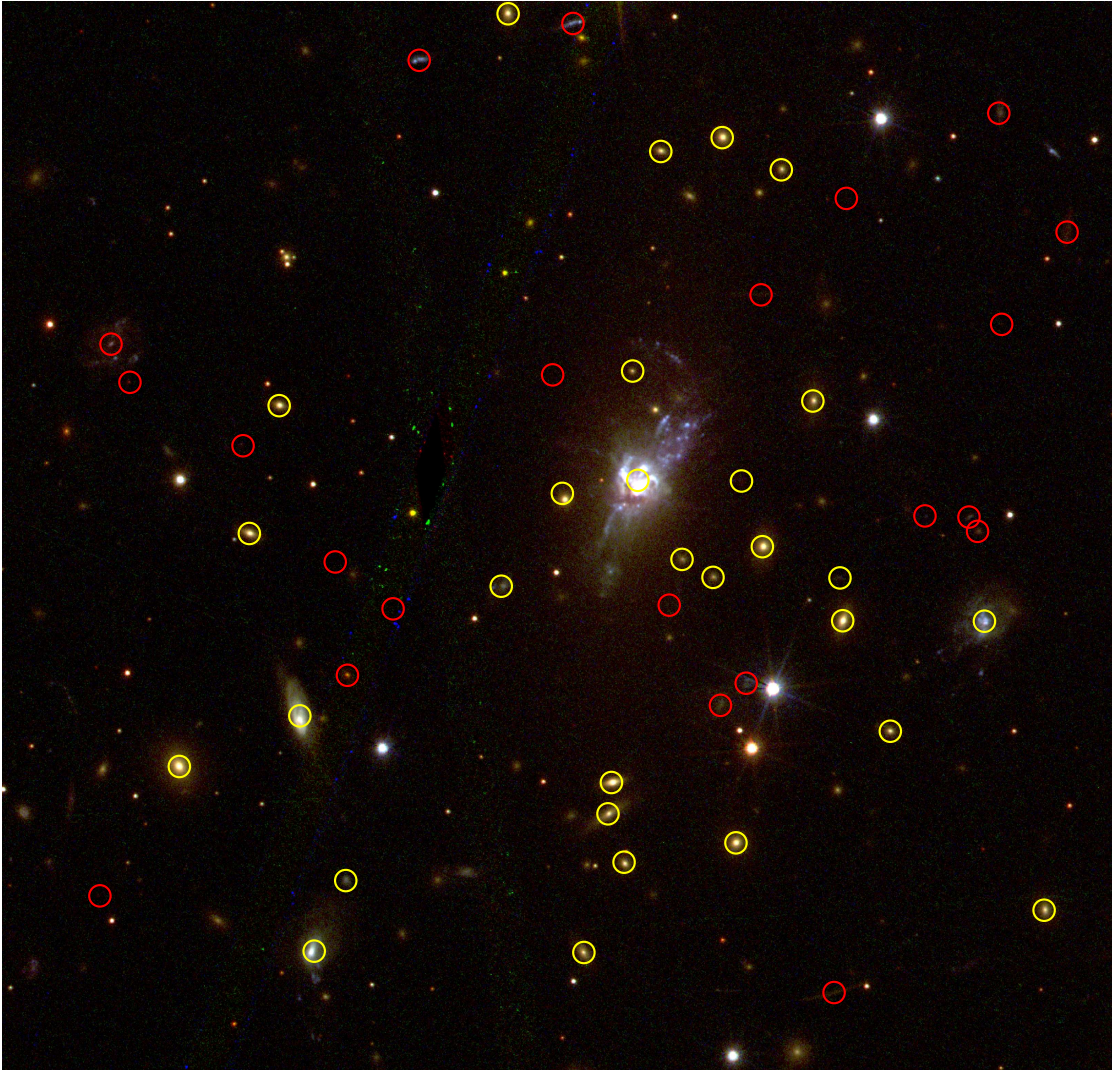


Figure 5.3 – HST color image of the galaxy cluster MACSJ1931 from combining the F814W, F625W, and F475W filters. Yellow circles mark the cluster members and red circles indicate the background sources with spectroscopic redshifts from MUSE.

MUSE signal in this wavelength range and subtract the averaged continuum signal at these two boundary wavelengths. The resulting $H\alpha$ map for the BCG is shown in figure 5.5.

In the same way we create OII emission maps for several galaxies in the MUSE field-of-view. The OII emission is a doublet with rest-frame wavelengths of 3726.0 Å and 3728.8 Å. Therefore we take 3727 Å as the reference rest-frame wavelength to calculate the redshifted center wavelength used for the stacking. For the $z = 0.8035$ galaxy shown in figure 5.6, the resulting OII center wavelength is $\lambda = 6721.6$ Å. The emission begins at 6714 Å and ends at 6731 Å. We stack the MUSE signal in this range and subtract the averaged continuum. We perform the same steps for two “jellyfish” galaxies at approximately the cluster redshift and we

show the resulting maps in figure 5.7. For the galaxy shown in the top part of the figure, we have a redshift $z = 0.3672$ and thus a wavelength range between 5078 \AA and 5102 \AA and for the one in the bottom part, the redshift is $z = 0.3442$ and thus the range is between 5005 \AA and 5019 \AA .

The resulting OII maps reveal several emission peaks in addition to the maxima close to the center of the objects. The two additional maxima were so strong for the $z = 0.8$ galaxy that we detect them as separate sources with MUSE. We remove these objects from the final source catalog presented here and we perform the same cleaning for other detected secondary peaks. The high sensitivity of MUSE leads to the detection of five additional emission peaks for the “jellyfish” galaxy shown in the top part of figure 5.7 and four for the one shown in the bottom part. These detailed emission maps illustrate the promise of MUSE data to study “jellyfish” galaxies and their tails, which permits the study of extreme ram-pressure stripping (e.g., [McPartland et al., 2016](#); [Ebeling et al., 2014](#)). In addition, the detailed $H\alpha$ map presented in figure 5.5 should allow us to study the feedback process of the SMBH at the center of the BCG (see e.g. [Kravtsov and Borgani, 2012](#)). Indeed, in a recent series of papers [Bellhouse et al. \(2017\)](#), [Poggianti et al. \(2017b\)](#), and [Sheen et al. \(2017\)](#) have used MUSE data to study gas stripping in galaxies and [Poggianti et al. \(2017a\)](#) investigated the feeding of SMBH at the center of “jellyfish” galaxies by ram pressure.

5.4.5 Lens model

The lens model which we present in this subsection is a preliminary model. We discuss the improvements which still need to be done at the end of this subsection.

Multiple images

The lens model is built following the same methodology as presented in previous papers from our team (e.g., [Limousin et al., 2007](#); [Richard et al., 2010](#); [Jauzac et al., 2014](#)). It relies on an iterative process to identify the different counter-images of each detected lensed galaxy. We start with 22 MUSE objects with a spectroscopic redshift $z > 0.4$. Among the 22 objects, 13 are found to be singly imaged. We investigate the multiplicity of the other nine images and identify their counter-images when possible.

The images 1.1 and 1.2 of the [Zitrin et al. \(2015\)](#) system #1 are detected in the MUSE data, and we measure a redshift of $z = 1.834$, which is in good agreement with the CLASH-VLT value presented in [Zitrin et al. \(2015\)](#). We also detect an image of the [Zitrin et al. \(2015\)](#) system #3, which is an elongated arc. While the positions of our multiple images for this system do not correspond exactly to the ones in [Zitrin et al. \(2015\)](#), it is however the same system. We measure a redshift of $z = 4.001$, in good agreement with the photometric redshift estimate of $z = 3.78$. We find a new multiple image system (#8), which is comprised of two background objects with spectroscopic redshift $z = 4.745$. It is difficult to identify the third

5.4. An improved lens model for the galaxy cluster MACSJ1931

counter-image of this system as it is too faint to be seen in the HST data. The new system #11 has also two multiple images detected by MUSE with spectroscopic redshifts $z = 1.1781$ and $z = 1.1783$ for 11.1 and 11.2, respectively. In total we identify four new multiple-image systems thanks to MUSE.

For every lensed galaxy detected with MUSE, we visually identify the counter-images. We base our identification on three main criteria:

1. The geometry of the system has to be in agreement with the model prediction.
2. The colours of the different multiple images of one system have to be similar.
3. A similar morphology is expected for all resolved lensed objects of a system.

Using these criteria, we identify a total of 11 multiple images for the 4 new systems. In this process we also revise the position of image 3.3 which was given by [Zitrin et al. \(2015\)](#).

In the last step, we include image systems #2, #4, #5, and #6 presented by [Zitrin et al. \(2015\)](#) into our mass model, which were not detected with MUSE. The latter two systems are outside of the field-of-view covered by our data. Systems #2 and #4 are well fitted by our model, however we revise the position of image 4.3. The lens model predicts redshifts for systems #2 and #4 of $z = 2.2 \pm 0.1$ and $z = 4.7 \pm 0.3$, respectively. This is in good agreement with the photometric measurements in [Zitrin et al. \(2015\)](#) for system #2 of $z = 2.35$, although there is a slight disagreement with system #4's estimated redshift of $z \sim 3$. Systems #5 and #6 were presented as candidates in [Zitrin et al. \(2015\)](#). For system #5, we revise the position of image 5.3, as it does not fit into our model. Using the criteria presented above, we identify a new candidate for 5.3. However, we have to remove image 5.1 from our model as the [Zitrin et al. \(2015\)](#) candidate does not fit and we cannot identify a reliable candidate. We obtain a redshift estimate for this system of $z = 6.1 \pm 0.7$. For system #6, we find that the images 6.1 and 6.2 are well fitted by the model, and we obtain a redshift estimate of $z = 7.7 \pm 0.5$. The position of image 6.3 is revised compared to [Zitrin et al. \(2015\)](#).

Finally, [Zitrin et al. \(2015\)](#) also presented a candidate multiple-image system (#7) which is found to be outside of the multiple image region with our new model. Thus it is only singly imaged. Table 5.1 presents all multiple images used for our model along with their spectroscopic or model-estimated redshifts and figure 5.8 shows their locations in the HST image.

ID	R.A.	Decl.	z
*1.1	292.96063	-26.569172	1.834
*1.2	292.95795	-26.568598	1.834
*1.3	292.94962	-26.570605	–
*2.1	292.9623	-26.568782	2.2 ± 0.1
*2.2	292.95508	-26.567931	–
*2.3	292.95103	-26.569241	–
*3.1	292.9534	-26.5837	4.001
*3.2	292.96471	-26.581836	–
3.3	292.95444	-26.583895	–
*4.1	292.95199	-26.582811	4.7 ± 0.3
*4.2	292.95544	-26.583678	–
4.3	292.96505	-26.581058	–
*5.2	292.95653	-26.589058	6.1 ± 0.7
5.3	292.96243	-26.587235	–
*6.1	292.95376	-26.586061	7.7 ± 0.5
*6.2	292.95554	-26.586314	–
6.3	292.96371	-26.585212	–
8.1	292.950480236	-26.5732870822	4.745
8.2	292.953186774	-26.5713240047	4.745
9.1	292.958304671	-26.5740740664	5.078
9.2	292.96238	-26.574929	–
9.3	292.94607	-26.576107	–
10.1	292.962093376	-26.5769896499	5.078
10.2	292.95961	-26.578445	–
10.3	292.94605	-26.577121	–
11.1	292.951816465	-26.5762723465	1.178
11.2	292.9563	-26.577666	1.178
11.3	292.96263	-26.575766	–

Table 5.1 – List of multiple image systems considered for our mass model. Images which have been previously identified by [Zitrin et al. \(2015\)](#) are marked with *. The redshifts given with an error correspond to systems that do not have a spectroscopic redshift from MUSE, thus the estimate comes from our best-fit mass model.

Mass Modeling

The mass modeling and optimization are performed using the LENSTOOL² software package (Jullo et al., 2007; Kneib et al., 1996). It uses a Bayesian estimator to sample the entire parameter space and return a best-fit model. LENSTOOL is described in more detail in chapter 6. To model the mass distribution, we use both cluster-scale halos to account for the large-scale mass distribution of the cluster and galaxy-scale halos to model the cluster galaxies. The latter can have a local lensing effect that needs to be taken into account to reconstruct the geometry of the multiple images (e.g., Harvey et al., 2016; Kneib et al., 1996).

Our best-fit mass model includes two cluster-scale halos, one centered on the BCG and one centered on a bright galaxy south of the cluster, as well as one galaxy-halo for the BCG itself and 202 galaxy-scale halos which are being optimized assuming a Faber-Jackson scaling relation (Faber and Jackson, 1976; Natarajan et al., 1998),

$$\begin{aligned} r_{\text{core}} &= r_{\text{core}}^* \left(\frac{L}{L_*} \right)^{\frac{1}{2}}, \\ r_{\text{cut}} &= r_{\text{cut}}^* \left(\frac{L}{L_*} \right)^{\frac{1}{2}}, \\ \sigma &= \sigma^* \left(\frac{L}{L_*} \right)^{\frac{1}{4}}. \end{aligned} \tag{5.1}$$

We model the halos using the dual Pseudo Isothermal Elliptical mass distribution (dPIE) profile (Elíasdóttir et al., 2007; Kassiola and Kovner, 1993).

The cluster member catalog includes both the 29 cluster galaxies spectroscopically confirmed by MUSE and the 172 cluster galaxies with a CLASH photometric redshift between 0.3 and 0.4. We use SExtractor to obtain their ellipticities and magnitudes and we exclude 10 cluster members from the CLASH catalog which could not be detected by SExtractor. The inclusion of the photometrically selected cluster members is necessary to trace the mass properly, as using only MUSE objects would restrict the field-of-view to a region smaller than the multiple image area. As shown in Harvey et al. (2016), it is important to include all cluster galaxies in the multiple image region, as they can have an impact of up to 0.5 arcsec on the global root mean square of the model. The root mean square is the difference between the observed positions of the multiple images and the positions predicted by the optimized model. It is a strong indicator of the goodness of fit of the model.

Our best-fit model is optimized using 28 multiple images (10 lensed galaxies) and reproduces the multiple image positions with a root mean square of 0.62 arcsec. The best-fit parameters are summarized in table 5.2. Figure 5.8 shows the location of the critical and caustic lines for a source redshift $z_s = 2$ and figure 5.9 presents the amplification map.

²Publicly available at <https://projets.lam.fr/projects/lenstool>

Chapter 5. Improving lens models with spectroscopic redshifts

Component	#1	#2	BCG	L* elliptical galaxy
Δ RA	$0.94^{+0.10}_{-0.09}$	$-1.33^{+0.35}_{-0.67}$	–	–
Δ DEC	$3.73^{+0.53}_{-0.49}$	$-44.90^{+0.88}_{-0.10}$	–	–
e	0.57 ± 0.02	0.62 ± 0.13	–	–
θ	86.1 ± 0.4	8 ± 56	–	–
r_{core} (kpc)	$73.6^{+3.3}_{-4.0}$	$5.3^{+4.8}_{-0.2}$	–	[0.15]
r_{cut} (kpc)	[1000]	[1000]	$161.8^{+67.9}_{-34.3}$	26.9 ± 16.3
σ (km s ⁻¹)	1073^{+11}_{-10}	316^{+28}_{-15}	125^{+32}_{-53}	104 ± 16

Table 5.2 – Best-fit dPIE parameters for the two large-scale DM halos, the BCG halo, and for the L* elliptical galaxy. The coordinates are quoted in arcseconds with respect to the BCG coordinates, $\alpha = 292.95682$ and $\delta = -26.575718$. The error bars correspond to the 1σ confidence level. The parameters in brackets are not optimized. The reference magnitude for the scaling relations is $\text{mag}_{\text{F814W}} = 19.65$.

Further work

The preliminary lens model results in an excellent fit to the data and the resulting root mean square is lower than the ones for the two [Zitrin et al. \(2015\)](#) lens models. However, we have found that the Markov Chain Monte Carlo (MCMC) engine of LENSTOOL converges for three lens model parameters to values which are exactly on the boundary of the value range allowed by the user. We have thus to re-run LENSTOOL with modified value ranges and investigate how the model changes. Once this is finished, we will use a pipeline which we have already set up to compute the cumulative mass for the cluster lens and its associated statistical error bounds. In addition, we will compute the predicted magnifications and the respective statistical errors for the multiple images. We will compare these results with the ones from the two [Zitrin et al. \(2015\)](#) models to quantify how much the addition of MUSE has improved the lens model.

Appendix

The spectroscopic redshifts measured with MUSE are listed in table 5.3. We also list the cross-matched photometric redshifts from the CLASH catalog.

5.4. An improved lens model for the galaxy cluster MACSJ1931

Right ascension	Declination	Spec-z	Δz	Photo-z	Lower limit	Upper limit
292.96246	-26.58306	0.3442	2e-05	0.259	0.21	0.292
292.95776	-26.58308	0.3419	0.00014	0.411	0.234	0.455
292.9534	-26.5837	4.0012	0.0001	3.779	3.593	3.959
292.94974	-26.58242	0.3468	0.00014	0.381	0.045	0.416
292.95706	-26.58167	0.3534	0.00022			
292.95511	-26.58137	0.3549	9e-05	0.517	0.452	0.545
292.95682	-26.57572	0.3526	1e-05	1.404	1.386	1.422
292.96191	-26.58196	0.3603	4e-05	0.392	0.259	0.418
292.96481	-26.58018	0.3581	5e-05	0.386	0.365	0.459
292.96271	-26.57939	0.3435	3e-05	0.194	0.174	0.218
292.95981	-26.58182	0.2304	1e-05	0.188	0.11	0.27
292.95734	-26.58092	0.3444	0.0001			
292.95728	-26.58043	0.3483	0.00011			
292.95242	-26.57963	0.35	0.00024	0.444	0.373	0.504
292.95078	-26.57791	0.3672	1e-05	0.355	0.322	0.387
292.95538	-26.57922	0.5223	2e-05			
292.95325	-26.57791	0.3507	8e-05	0.259	0.208	0.301
292.96188	-26.57876	0.8171	0.00014	4.623	4.563	4.699
292.95465	-26.57675	0.35	8e-05	0.341	0.286	0.375
292.9592	-26.57737	0.3582	2e-05			
292.95551	-26.57723	0.3579	0.00034			
292.95605	-26.57695	0.3537	0.00019			
292.96359	-26.57655	0.3477	9e-05	0.346	0.311	0.414
292.9509	-26.57651	0.5223	1e-05	0.719	0.212	0.769
292.95105	-26.57629	0.5228	1e-05	0.677	0.585	0.72
292.95377	-26.57448	0.3585	0.00013	0.435	0.299	0.515
292.96307	-26.57455	0.3494	0.00012	0.4	0.304	0.425
292.966	-26.5736	0.8035	1e-05	0.79	0.75	0.839
292.96567	-26.57419	0.8036	7e-05	5.597	5.472	5.711
292.94934	-26.57185	0.8229	2e-05	0.721	0.572	0.875
292.95432	-26.57088	0.3591	0.00018			
292.95535	-26.57038	0.3515	0.00014	0.323	0.257	0.362
292.9516	-26.57102	0.2126	1e-05	3.182	3.14	3.272
292.95206	-26.57081	0.2126	1e-05			

Table 5.3 – Spectroscopic redshifts for the sources detected with MUSE. We also list the best photometric redshift estimate from the CLASH catalog and their upper and lower redshift bounds if available. *Continued on the next page.*

Chapter 5. Improving lens models with spectroscopic redshifts

Right ascension	Declination	Spec-z	Δz	Photo-z	Lower limit	Upper limit
292.95642	-26.57059	0.3486	0.00013			
292.95053	-26.57	0.5192	3e-05	0.443	0.256	0.548
292.95908	-26.56845	0.3644	0.00011	0.383	0.357	0.409
292.96063	-26.56917	1.8342	9e-05	1.821	1.793	1.94
292.95795	-26.5686	1.8346	6e-05			
292.95467	-26.57283	0.6395	8e-05			
292.95691	-26.57401	0.3538	2e-05			
292.95814	-26.57592	0.3489	4e-05	0.381	0.354	0.43
292.95182	-26.57627	1.1781	6e-05	1.909	1.177	2.542
292.9533	-26.57724	0.3475	1e-05			
292.95501	-26.57573	0.3541	1e-05			
292.96108	-26.57772	0.5114	7e-05			
292.95048	-26.57329	4.7451	6e-05			
292.95319	-26.57132	4.7453	6e-05			
292.9583	-26.57407	5.0786	0.00011			
292.96209	-26.57699	5.0784	9e-05			
292.95493	-26.57888	0.5319	4e-05			
292.95627	-26.57766	1.1783	5e-05			
292.96619	-26.58219	5.8333	0.0001	5.556	5.108	5.812
292.9637	-26.57518	0.8005	1e-05	0.735	0.287	0.82

Continued from the previous page.

5.4. An improved lens model for the galaxy cluster MACSJ1931

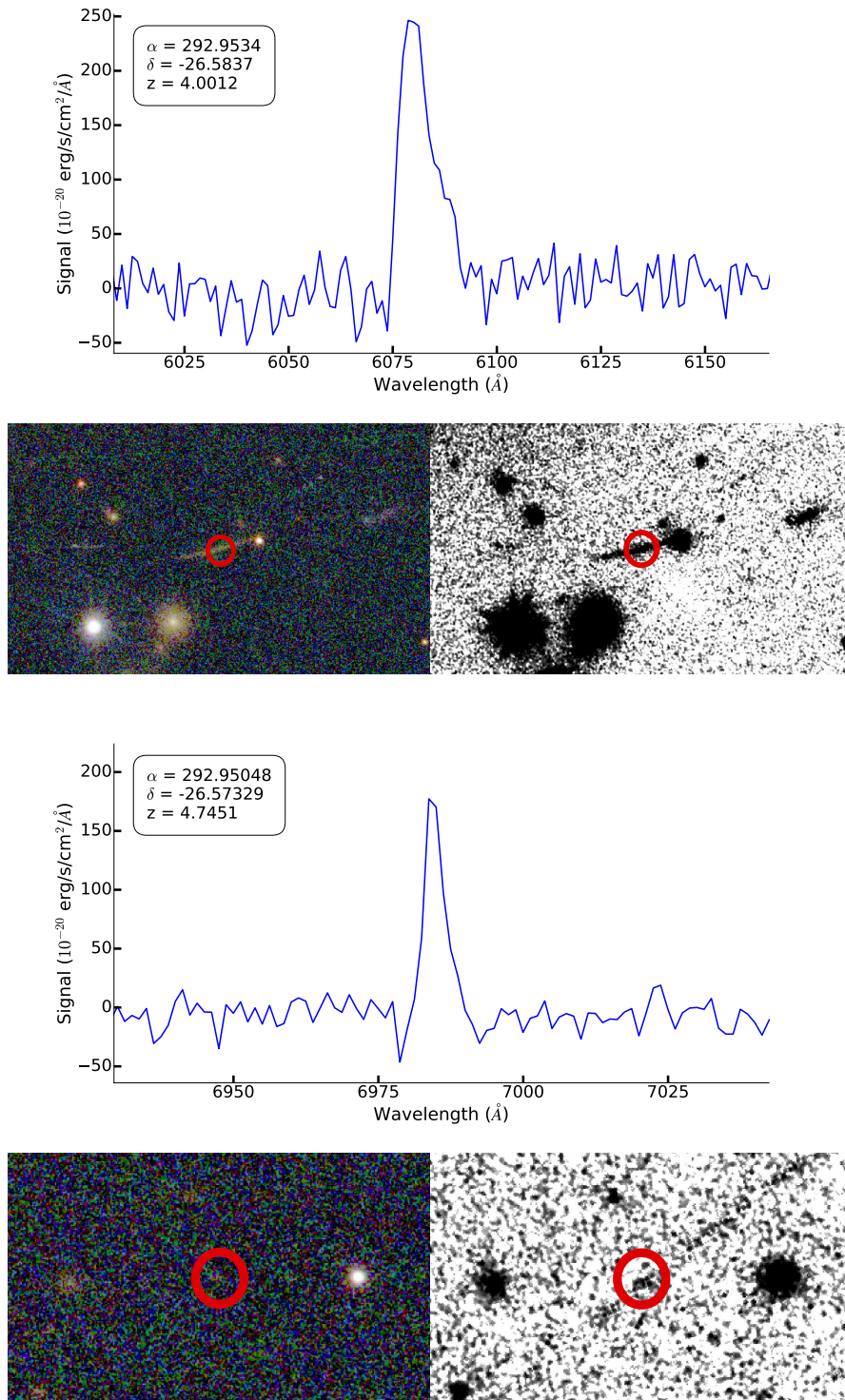
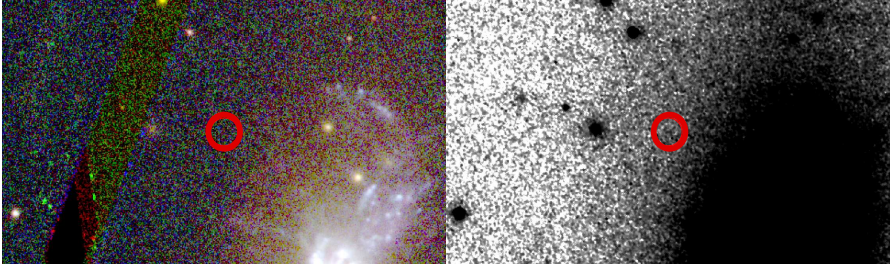
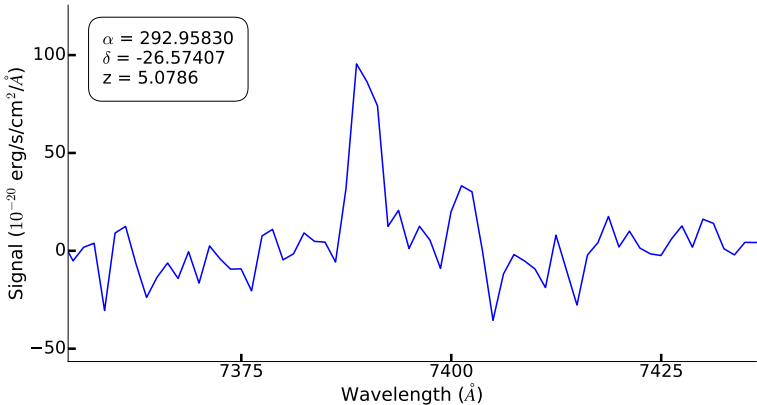
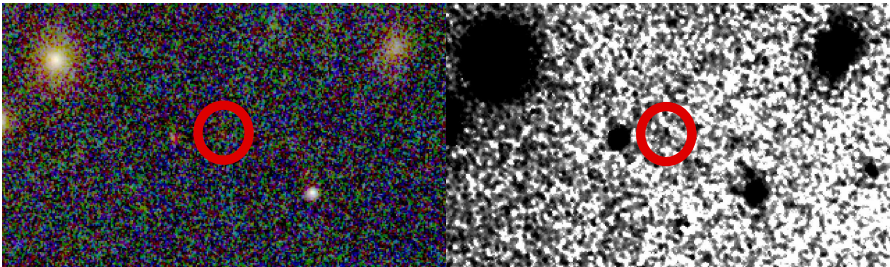
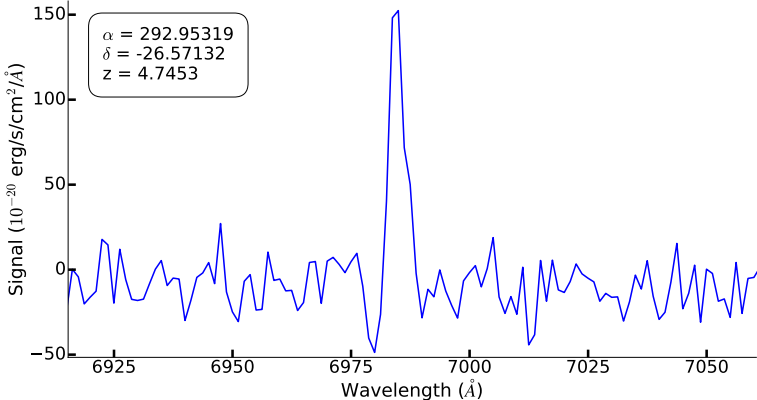


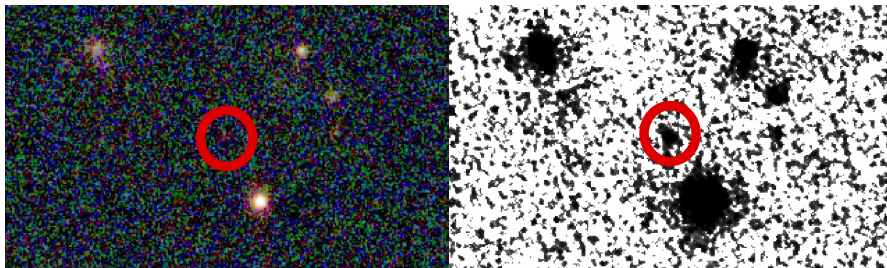
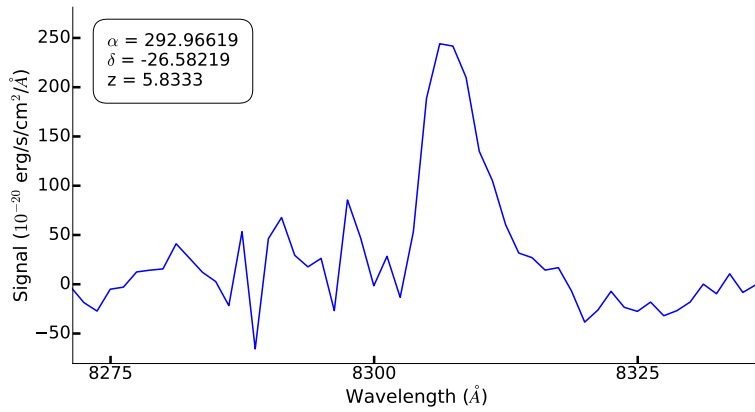
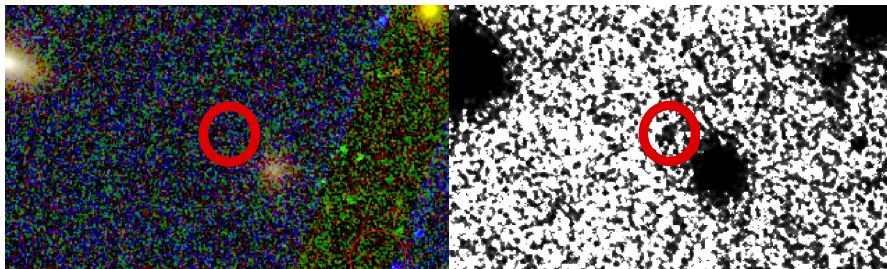
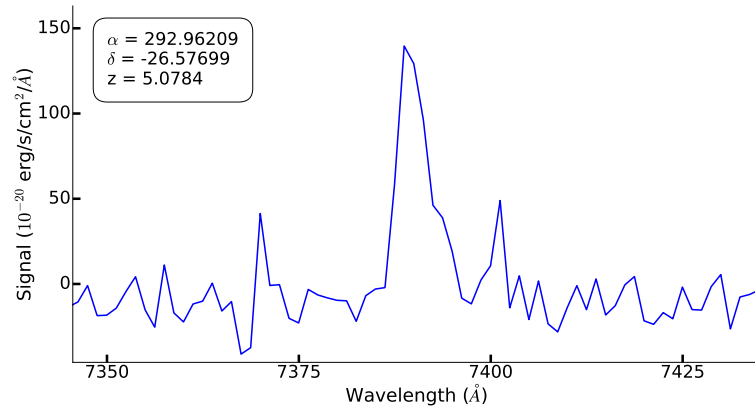
Figure 5.4 – The image locations and the spectral parts with the Lyman- α emission line are shown for the six sources with spectroscopic $z > 4$. Top: Spectral part with Lyman- α emission line. Bottom left: Location in the HST color image. Bottom right: Location in the HST F105W filter image. *Continued on the next pages.*

Chapter 5. Improving lens models with spectroscopic redshifts



Continued from the previous page.

5.4. An improved lens model for the galaxy cluster MACSJ1931



Continued from the previous page.

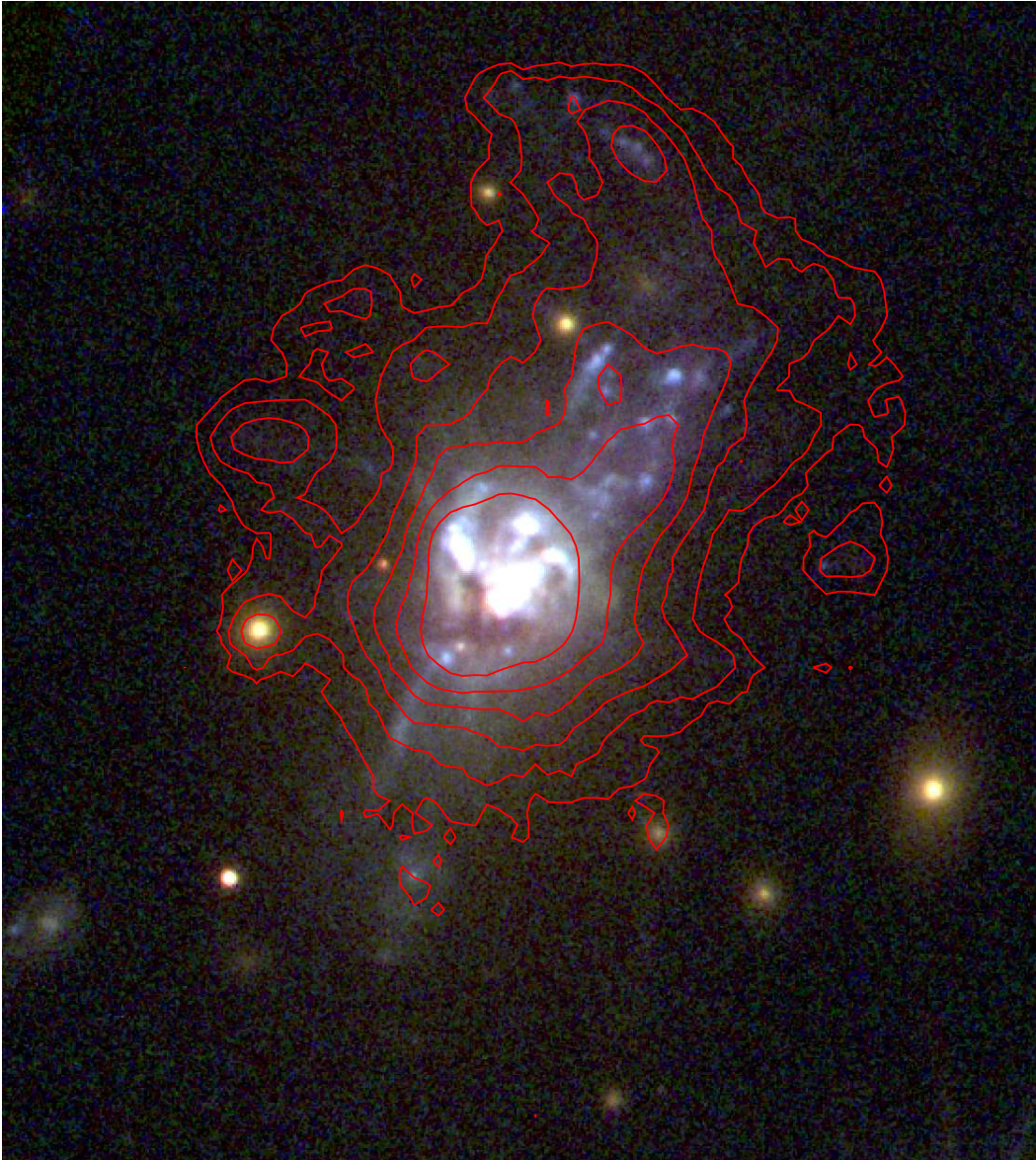


Figure 5.5 – The $H\alpha$ emission contours of the BCG are shown in red. The contours indicate an emission of 100, 200, 400, 800, 1600, and $3200 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$, respectively. Note that we show only the emission contours of the BCG and not those of close-by sources.



Figure 5.6 – The OII emission contours for a galaxy at redshift $z = 0.8035$ are shown in yellow. The outer contour indicates $50 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ and the step size is $25 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$. We show only emission contours for the galaxy and not those for close-by sources.

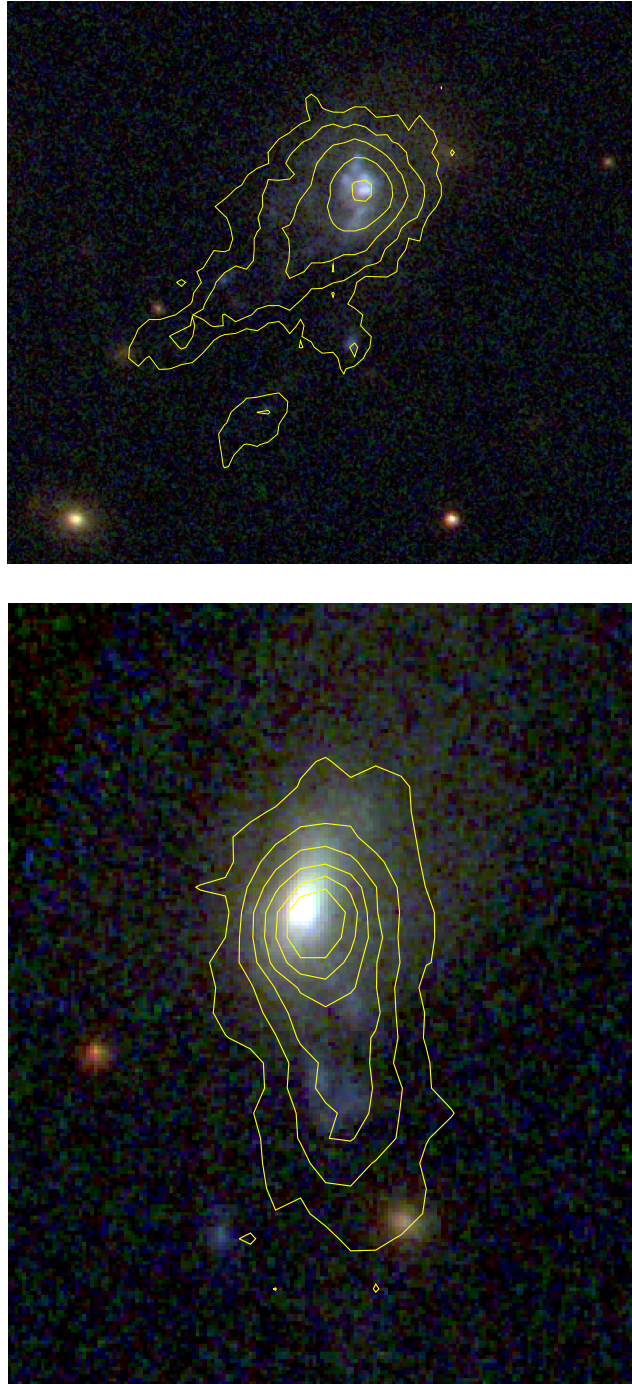


Figure 5.7 – The OII emission contours for two “jellyfish” galaxies at redshift $z \approx 0.34$ are shown in yellow. Top image: The contours indicate $50, 100, 200, 400,$ and $800 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$. Note the clear detection of the OII emission region in the lower part, which is almost invisible in the color image. Bottom image: The contours indicate $50, 100, 150, 200, 250,$ and $300 \times 10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$. In both images we show only the emission contours for the respective galaxies and not for close-by sources.

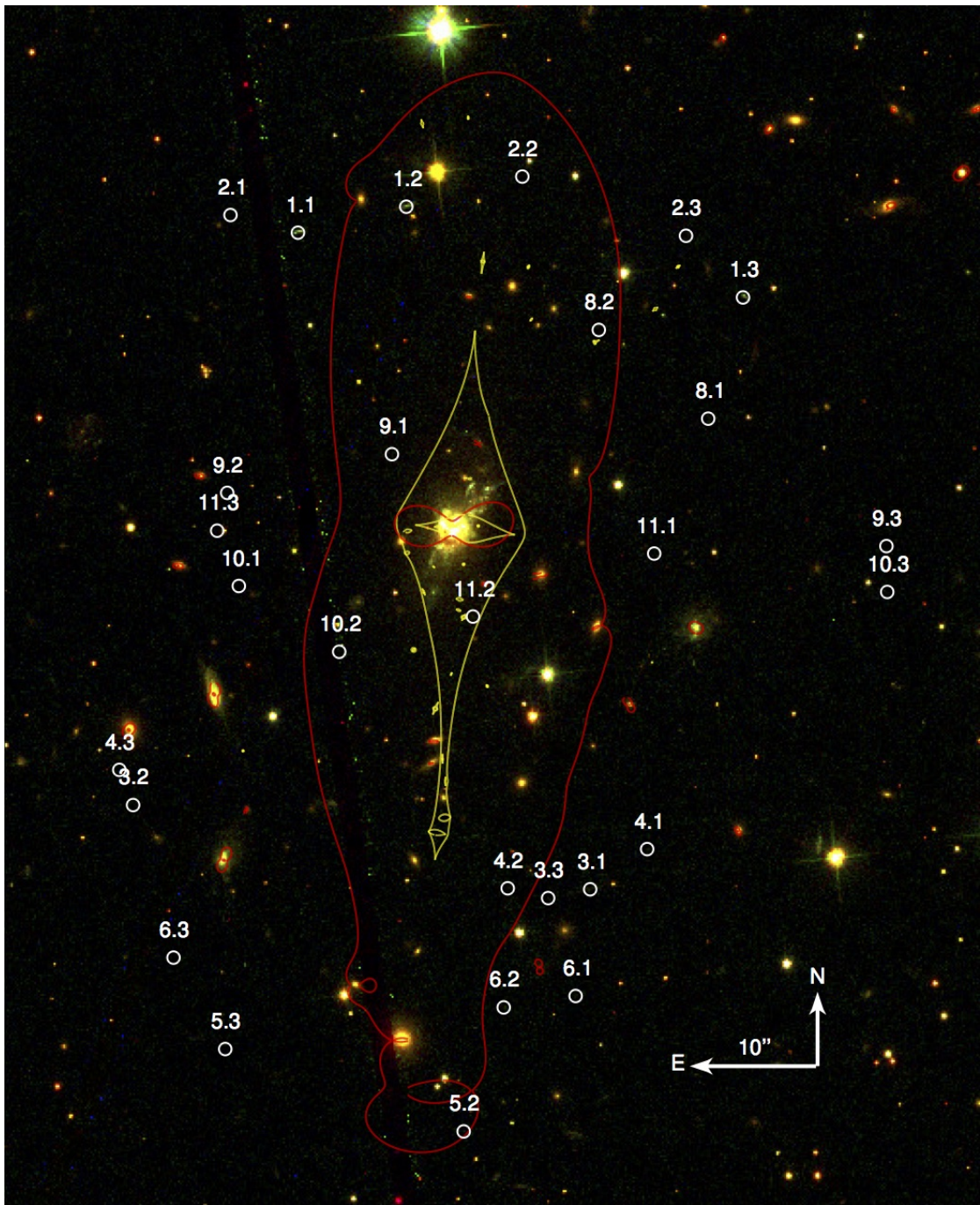


Figure 5.8 – The critical lines (red) and caustics (yellow) for the best-fit lens model are shown for a source redshift $z = 2$. The white circles indicate the multiple image systems used to build the lens model.

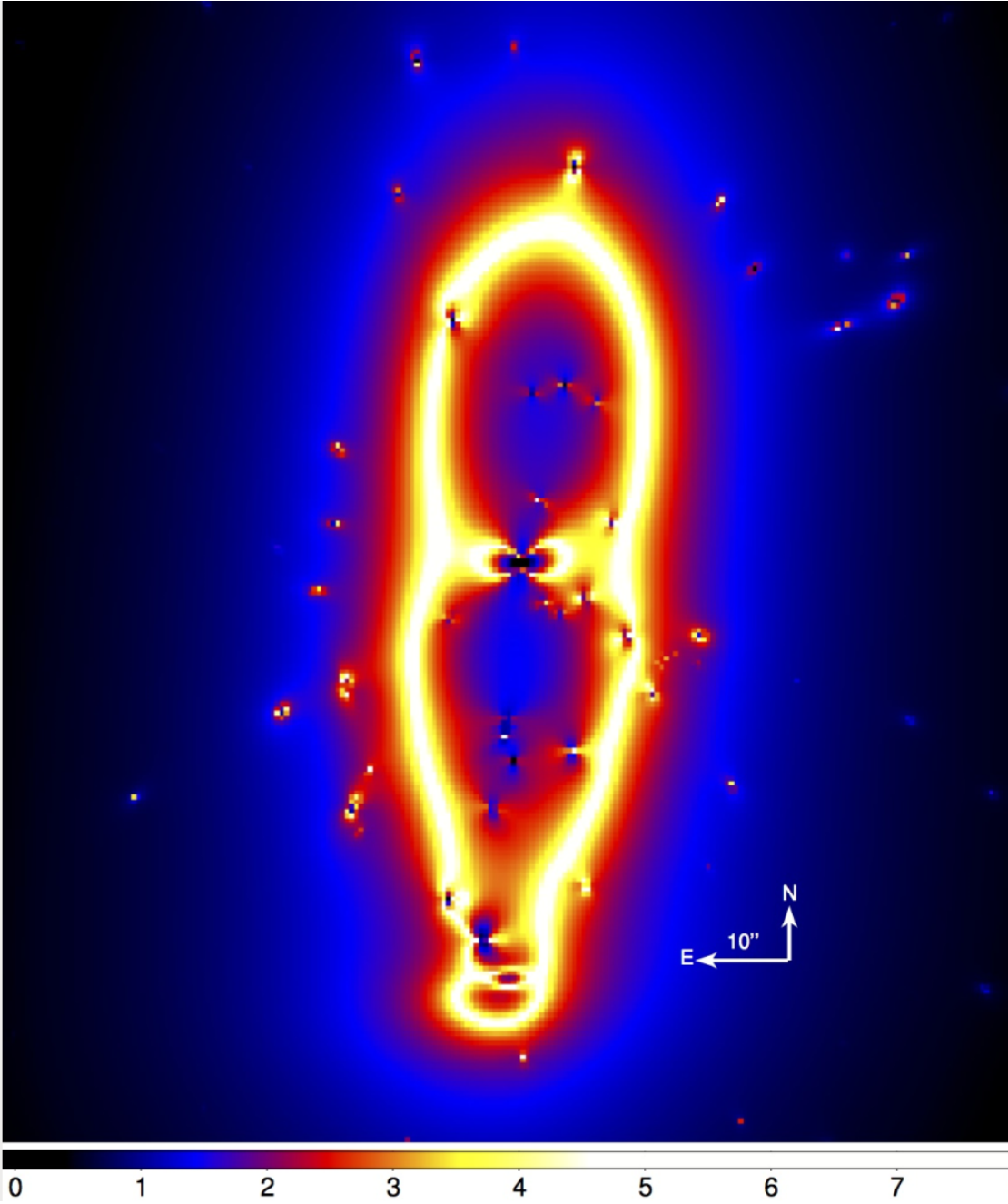


Figure 5.9 – The amplification map for the best-fit lens model for MACSJ1931 is shown for a source redshift $z = 2$.

6 High Performance Computing for cluster lensing

6.1 Wanted: A power boost for LENSTOOL

6.1.1 LENSTOOL, a successful lens modeling software

The first version of the LENSTOOL software which we use in section 5.4 to model the cluster lens was created by Jean-Paul Kneib in the early 1990s. It was successfully applied to model several lenses (e.g., [Kneib et al., 1996](#); [Natarajan et al., 1998](#)) and further developed. New lens modeling techniques were incorporated (e.g., [Natarajan et al., 1998](#)). A major evolution of the software was completed in 2007, when the Bayesian MCMC engine BayeSys¹ was added to LENSTOOL ([Jullo et al., 2007](#)). It permits the user to obtain the most likely lens model instead of simply the one which gives the best fit. In addition, the MCMC efficiently samples the high-dimensional space of allowed lens parameters. Only two years later, yet another major addition was incorporated into the software, namely multiscale cluster lens mass mapping ([Jullo and Kneib, 2009](#)).

These efforts were very fruitful as demonstrated by the plethora of lens models which have been created with the software (e.g., [Limousin et al., 2007](#); [Elíasdóttir et al., 2007](#); [Richard et al., 2010](#); [Sharon and Johnson, 2015](#); [Johnson et al., 2014](#); [Paraficz et al., 2016](#); [Limousin et al., 2016](#)). Indeed, it has become a very well-known tool for SL and it is used worldwide. The public availability of the software and the open source code have facilitated this development. Lens modeling has prospered thanks to high quality data from large surveys like CLASH and the HFF, and LENSTOOL has been a major driver for high precision lens models like the ones presented in [Jauzac et al. \(2014\)](#) and [Jauzac et al. \(2015\)](#), which provide mass determinations with percent-level precision. In a recent comparison of SL modeling software based on HFF-like simulations it has performed very well ([Meneghetti et al., 2016](#)).

This leads to the obvious question: If LENSTOOL performs so well, why should we redesign it? After all, there is truth in the proverb “never change a running system”. However, the HFF data

¹Publicly available at <http://www.inference.org.uk/bayesys/>

sets have provided a huge number of lensing constraints and this posed a serious computing challenge for LENSTOOL. It can take several weeks on a computing cluster to compute a single HFF lens model and the lens models have to be iteratively updated and modified until the final lens mass distribution is obtained. With new exciting telescopes like JWST, Euclid, and the LSST on the horizon, LENSTOOL faces a serious performance challenge.

It is not trivial to speed up the software. A part of the code has been parallelized by a software engineer, but LENSTOOL uses global variables which prevent it from being run on several nodes of a computing cluster. As a result, it can use only the ≈ 16 CPU cores available on a modern server node. In addition, it is not easy to include new features in the source code. Therefore the software has to be completely rewritten and redesigned with maximum performance and extensibility in mind. However, the lensing community naturally wants to keep the proven algorithms and to preserve backwards compatibility. As a result, the performance challenge for LENSTOOL has developed into a software development challenge for us.

The new version of the software, LENSTOOL-HPC, is designed to meet this challenge by using HPC methods. The LENSTOOL algorithms are very well suited for massive parallelism and we employ this technique to accelerate the computations. First steps towards this goal were taken with the research project of Thomas Jalabert and the Master's thesis of Christoph Schäfer ([Schäfer, 2016](#)). The fruitful collaboration with the EPFL HPC expert Gilles Fourestey has accelerated this development. We are approaching the first milestone, namely the release of a first LENSTOOL-HPC version which performs the essential LENSTOOL task, the modeling of strong gravitational lensing, in a rapid fashion. The project is led by Christoph Schäfer, now a PhD student, and Gilles Fourestey. We are developing this software with several different hardware setups in parallel to find the best solution. In particular, we are using CPUs, Graphics Processing Unit (GPU) acceleration, and Intel Xeon Phi coprocessors, as we are lucky to have access to some resources of the EPFL Scientific IT and Application Support Center and the Swiss National Supercomputing Centre.

During the development phase, we have asked ourselves the question: Can we achieve even better performance by using Single Precision (SP) instead of the commonly used Double Precision (DP) for the computations? The computing power of both CPUs and GPUs is higher for SP (see e.g. [Eijkhout et al., 2016](#); [Besl, 2013](#)), so we can expect a good performance improvement. The downside is that this might lead to an error in our results if SP is not precise enough for our computations. We have investigated this question for a performance crucial part of the LENSTOOL lensing algorithm and we will give an answer in section 6.2. Before doing so, we present the LENSTOOL SL algorithm in the next section. The results of this chapter will be presented in Rexroth et al. (2017c, in preparation).

6.1.2 The strong lensing algorithm

Overview

LENSTOOL models galaxy clusters by using parametric models of the large-scale cluster halos and the galaxy-scale halos. In a typical merging cluster, we have two large-scale halos and hundreds of galaxy halos. Depending on the chosen parametric model, we have several free parameters such as x and y position, velocity dispersion, etc. for each halo. It is possible to constrain the range of the free parameters or to reduce their number, e.g. by assuming a scaling relation like the Faber-Jackson relation (Faber and Jackson, 1976) for galaxy-scale halos (Natarajan et al., 1998). Nevertheless, the best lens model will still be hidden in a massive, high dimensional parameter space. LENSTOOL uses BayeSys3, a Bayesian MCMC software package, to sample this parameter space (see Jullo et al., 2007, for a detailed description). For each parameter combination probed by the MCMC, LENSTOOL computes the goodness of fit of the corresponding lens model given the observational data. It does this by modeling the lens with the given set of parameters and, using this model, lensing the observed multiple images into the source plane and subsequently back into the image plane, see figure 6.1. If the probed lens model is close to the true matter distribution, the re-lensed multiple image positions will be close to the observed multiple image positions and the goodness of fit parameter

$$\chi^2 = \sum_i \sum_j \frac{(x_{\text{obs},ij} - x_{ij})^2}{\sigma_{ij}^2} \quad (6.1)$$

will be small (Jullo et al., 2007). We denote the observed position of multiple image j of multiple image system i with $x_{\text{obs},ij}$, the re-lensed position with x_{ij} , and the error budget of the position with σ_{ij} . Since the parameter space probed by the MCMC is massive, it can take several weeks of computation time to find the best model for lenses with HFF-like data.

There are two ways to speed up this computation. The first is to speed up the MCMC, e.g. by parallelizing it. This will be explored in the upcoming paper by Christoph Schäfer, which will also present the first HPC version of LENSTOOL. The second way is to speed up the χ^2 computation. In this part of the thesis, we will focus on accelerating a crucial part of the χ^2 computation, the gradient calculation. Since we will have to take a very precise look at the algorithm when we compute the impact of SP and DP on its result, we will now present a detailed description.

Gradient computation in the χ^2 algorithm

Before we present the χ^2 algorithm, we reformulate the lens equation 3.14 by introducing

$$\varphi = \frac{D_{\text{os}}}{D_{\text{ls}}} \Psi \quad (6.2)$$

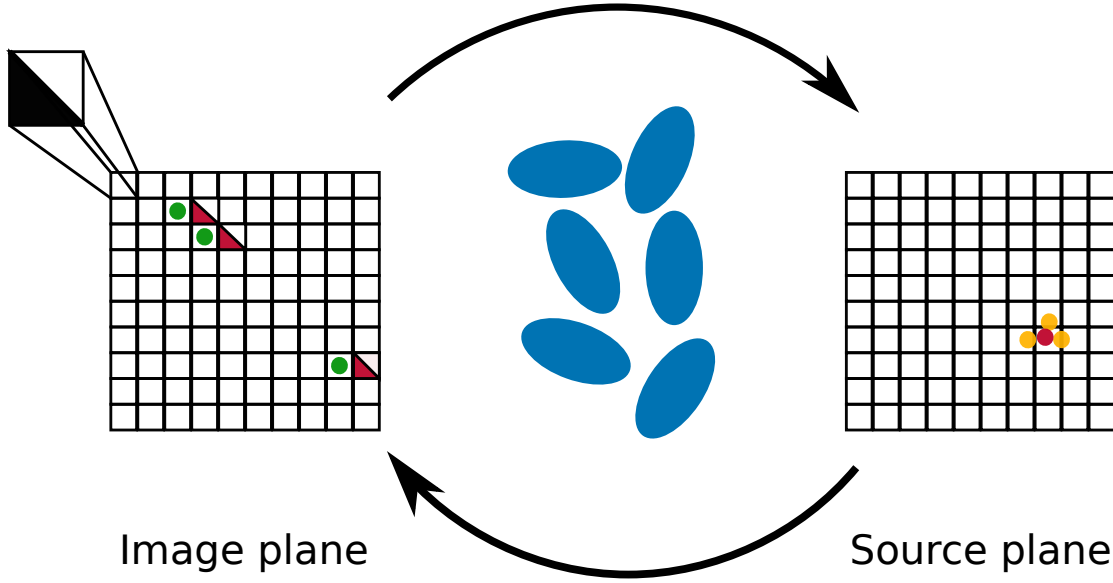


Figure 6.1 – LENSSTOOL computes the multiple image positions predicted by a lens model (red triangles, image plane). In the first step, it lenses the observed multiple images (green dots, image plane) onto their respective predicted sources (yellow dots, source plane) and computes their barycenter (red dot, source plane). In the second step, it decomposes the image plane pixels into triangles and lenses each triangle into the source plane. Every time that the source plane triangle includes the barycenter, a predicted multiple image is found. If the lens model is close to the true model, these re-lensed images will be located very close to the observed images. Note that image plane pixels lensed into the source plane will typically be distorted due to the SL effect. We do not show this effect to keep the figure simple. As a result of this distortion, squares are not always mapped onto squares and we thus have to partition the pixels into triangles (top left corner, image plane), which are always mapped onto triangles.

and making the gradient dependence explicit:

$$\begin{aligned}\beta_1 &= \theta_1 - \frac{D_{ls}}{D_{os}} (\nabla\varphi(\boldsymbol{\theta}))_1, \\ \beta_2 &= \theta_2 - \frac{D_{ls}}{D_{os}} (\nabla\varphi(\boldsymbol{\theta}))_2.\end{aligned}\tag{6.3}$$

As a result, we only have to compute the constant D_{ls}/D_{os} once instead of for every image pixel. Note that the deflection potential at position $\boldsymbol{\theta}$ is a superposition of all cluster-scale and galaxy-scale deflection potentials Ψ_{cluster} and Ψ_{galaxy} ,

$$\Psi(\boldsymbol{\theta}) = \sum \Psi_{\text{cluster}}(\boldsymbol{\theta}) + \sum \Psi_{\text{galaxy}}(\boldsymbol{\theta}),\tag{6.4}$$

(see e.g. Jullo et al., 2007) and as a result we have

$$\nabla\varphi(\boldsymbol{\theta}) = \sum \nabla\varphi_{\text{cluster}}(\boldsymbol{\theta}) + \sum \nabla\varphi_{\text{galaxy}}(\boldsymbol{\theta}).\tag{6.5}$$

We see that the lens equation is computationally cheap to evaluate once the total gradient $\nabla\varphi$ is known. The computation of $\nabla\varphi$, however, involves potentially complicated gradient calculations for hundreds of potentials and as we will soon see, it has to be computed for every pixel in our image. The HST *Advanced Camera for Surveys* produces images with 4096×4096 pixels at a pixel scale of ≈ 0.05 arcsec/pixel (Avila et al., 2017), which we can typically upsample to 0.03 arcsec/pixel (Lotz et al., 2017), so that HFF images have a total of $\approx 6730 \times 6730$ pixels ≈ 45 million pixels. This shows that the computation of $\nabla\varphi$ is computationally expensive and an excellent target for speedup with HPC parallelism methods.

The χ^2 computation is now performed as follows. We compute $\nabla\varphi$ for each pixel of the image plane. Then we loop over each multiple image j in each multiple image system i . For each multiple image, we use equation 6.3 to compute the source coordinates, $\beta_{ij,1}$ and $\beta_{ij,2}$. Subsequently, we determine the barycenter of the sources of a given multiple image system i . If we are close to the true lens model, all multiple images will be mapped onto approximately the same source location, but in general the locations of the predicted sources can differ substantially, which makes it necessary to use the barycenter. In the next step, we re-lens the barycenter back into the image plane to obtain the locations of the multiple images predicted by the lens model. However, the lens equation cannot easily be inverted, so we have to find the locations in a different way. First, we divide each pixel in the image plane into two triangles, see figure 6.1. We do this because lensing always maps triangles onto triangles, but not squares onto squares. Second, we lens each triangle into the source plane by using equation 6.3 and we check if the barycenter is inside this triangle in the source plane. If it is, a predicted multiple image location in the image plane is found. Once we have found the locations of all predicted multiple images for all multiple image systems, we compute the χ^2 according to equation 6.1.

The gradient calculations will naturally differ for different chosen parametric models. As an example, we present the gradient computation for a generalized form of the SIS, the Singular Isothermal pseudo Elliptical sphere (SIE), in algorithm 1. It is necessary to generalize the parametric model, as we want to use this algorithm to model any SIS lens configuration by simply choosing the appropriate number of lenses and parameter values. We expand our treatment of the SIS by following the procedure in Golse and Kneib (2002). We introduce the pseudo-ellipticity of the deflection potential, ϵ , and the coordinate system

$$\begin{aligned} R &= \sqrt{\theta_{1,\epsilon}^2 + \theta_{2,\epsilon}^2}, \\ \phi &= \arctan\left(\frac{\theta_{2,\epsilon}}{\theta_{1,\epsilon}}\right), \end{aligned} \tag{6.6}$$

with

$$\begin{aligned}\theta_{1,\epsilon} &= \sqrt{a_{1,\epsilon}} \theta_1, \\ \theta_{2,\epsilon} &= \sqrt{a_{2,\epsilon}} \theta_2,\end{aligned}\tag{6.7}$$

$$\begin{aligned}a_{1,\epsilon} &= 1 - \epsilon, \\ a_{2,\epsilon} &= 1 + \epsilon.\end{aligned}\tag{6.8}$$

Note that we call ϵ a pseudo-ellipticity, because the resulting elliptical shapes will only correspond to ellipses with classical ellipticity $\epsilon' = 1 - b/a$, where a and b are the semi-major and semi-minor axes of the ellipse, for small values of ϵ (Golse and Kneib, 2002). Therefore we assume in the following $\epsilon \ll 1$. The advantage of using a pseudo-elliptical parametric model is that it leads to relatively simple analytic expressions of the derived lensing quantities (Golse and Kneib, 2002). Now we can simply calculate the values of the pseudo-elliptical deflection potential Ψ_ϵ at location θ by using the relation (Golse and Kneib, 2002)

$$\Psi_\epsilon(\theta) = \Psi(R, \phi),\tag{6.9}$$

and analogous for φ_ϵ . The resulting pseudo-elliptical shape is stretched along the θ_1 -axis, so that we have $\Phi = 0$, where Φ is the counter-clockwise angle between the semi-major-axis and the θ_1 -axis. Algorithm 1 extends this approach to potentials with $\Phi \neq 0$ by using rotations. We obtain the following equations for the scaled deflection angle (Golse and Kneib, 2002),

$$\begin{aligned}\alpha_{1,\epsilon}(\theta) &= |\alpha(R, \phi)| \sqrt{a_{1,\epsilon}} \cos(\phi), \\ \alpha_{2,\epsilon}(\theta) &= |\alpha(R, \phi)| \sqrt{a_{2,\epsilon}} \sin(\phi).\end{aligned}\tag{6.10}$$

We can now derive the gradient expressions for the SIE,

$$\begin{aligned}(\nabla\varphi_\epsilon)_1 &= (1 - \epsilon) b_0 \frac{\theta_1}{R}, \\ (\nabla\varphi_\epsilon)_2 &= (1 + \epsilon) b_0 \frac{\theta_2}{R},\end{aligned}\tag{6.11}$$

where we introduced the parameter

$$b_0 = 4\pi \left(\frac{\sigma v}{c} \right)^2.\tag{6.12}$$

The presented equations for the SIE always reduce to the equations for the spherical SIS for $\epsilon = 0$.

Algorithm 1 Compute $\nabla\varphi_\epsilon$ in each image pixel for a SIE

Require: $\theta_{\text{center}}, b_0, \epsilon, \Phi \forall$ SIE lenses, image \mathbf{l}
Output: $\nabla\varphi_\epsilon \forall$ pixels $(\theta_1, \theta_2) \in \mathbf{l}$

- 1: **Procedure** GRADIENT($\mathbf{l}, \{\theta_{\text{center},i}, b_{0,i}, \epsilon_i, \Phi_i\}$):
- 2: **for all** $(\theta_1, \theta_2) \in \mathbf{l}$ **do**
- 3: **for all** SIE lenses **do**
- 4: $\Delta\theta_{1,i} \leftarrow \theta_1 - \theta_{\text{center},i,1}$
- 5: $\Delta\theta_{2,i} \leftarrow \theta_2 - \theta_{\text{center},i,2}$
- 6: $\Delta\theta_{1,i} \leftarrow \Delta\theta_{1,i} \cos(\Phi_i) + \Delta\theta_{2,i} \sin(\Phi_i)$
- 7: $\Delta\theta_{2,i} \leftarrow \Delta\theta_{2,i} \cos(\Phi_i) - \Delta\theta_{1,i} \sin(\Phi_i)$
- 8: $R_i \leftarrow \text{sqrt}((\Delta\theta_{1,i})^2(1 - \epsilon_i) + (\Delta\theta_{2,i})^2(1 + \epsilon_i))$
- 9: $(\nabla\varphi_\epsilon)_{1,i} \leftarrow (1 - \epsilon_i) b_{0,i} \Delta\theta_{1,i} / R_i$
- 10: $(\nabla\varphi_\epsilon)_{2,i} \leftarrow (1 + \epsilon_i) b_{0,i} \Delta\theta_{2,i} / R_i$
- 11: $(\nabla\varphi_\epsilon)_{1,i} \leftarrow (\nabla\varphi_\epsilon)_{1,i} \cos(-\Phi_i)$
 $\quad\quad\quad + (\nabla\varphi_\epsilon)_{2,i} \sin(-\Phi_i)$
- 12: $(\nabla\varphi_\epsilon)_{2,i} \leftarrow (\nabla\varphi_\epsilon)_{2,i} \cos(-\Phi_i)$
 $\quad\quad\quad - (\nabla\varphi_\epsilon)_{1,i} \sin(-\Phi_i)$
- 13: **end for**
- 14: $(\nabla\varphi_\epsilon)_1 \leftarrow \sum_i (\nabla\varphi_\epsilon)_{1,i}$
- 15: $(\nabla\varphi_\epsilon)_2 \leftarrow \sum_i (\nabla\varphi_\epsilon)_{2,i}$
- 16: **end for**
- 17: **return** $\{\nabla\varphi_\epsilon\}$

6.2 Single vs Double Precision and GPUs vs CPUs

6.2.1 CPU and GPU implementations

We implement a performance-optimized CPU version of the gradient computation by using the following techniques. First, we structure our data in the Structures of Arrays (SoA) format instead of the Arrays of Structures (AoS) format. To illustrate the difference, we take a look at the internal representation of five SIS potentials using SoA and AoS. In the AoS format, they are stored as an array comprised of five different data structures. Each data structure corresponds to a SIS potential and it contains the respective data of the SIS like $\theta_{\text{center},1}$, $\theta_{\text{center},2}$, and σ_v . In the SoA format, the potentials are stored in one data structure which consists of different arrays. Each array corresponds to a SIS parameter like $\theta_{\text{center},1}$ and array element number one of the $\theta_{\text{center},1}$ -array would correspond to the $\theta_{\text{center},1}$ location of SIS number one, element number two to the $\theta_{\text{center},1}$ location of SIS number two etc. As a result, the SIS parameters occupy contiguous parts of the memory, which is usually beneficial for vectorized computations (e.g., Eijkhout et al., 2016; Besl, 2013). Second, we use Advanced Vector Extensions (AVX) technology available on the latest CPU generations to harvest their built-in vectorization potential. For example, Besl (2013) obtained a significant speed-up by using SoA and AVX. CPU cores with AVX technology use registers with a width of 256 bits to process 8 SP or 4 DP values simultaneously (Besl, 2013). Note that the same operation

has to be performed for each of the simultaneously processed data values. Therefore AVX is a Single Instruction Multiple Data (SIMD) parallelism technique (Eijkhout et al., 2016). Third, we parallelize the computation using Open Multi-Processing (OpenMP)² on the outermost loop of algorithm 1. Each core of the multi-core CPU will now work on computing the total gradient for its assigned pixel and thus we compute the gradients for several pixels in parallel.

We implement the GPU version of the algorithm with CUDA³. First, we structure our data again in the SoA format. Second, we use the massively parallel architecture of GPUs to parallelize the gradient computation. Modern GPUs have many Streaming Multiprocessors, which in turn consist of many Streaming Processors, so the total amount of processor cores is computed by multiplying the two (e.g., Eijkhout et al., 2016). The number of cores available depends on the GPU model, for example the Nvidia Tesla K80 GPU Accelerator (henceforth called Nvidia K80) consists of two GPUs with 2496 cores each (Nvidia Corporation, 2015). In addition, GPUs are designed to be extremely efficient at switching between threads, where all threads in a single block of threads execute the same instruction (Eijkhout et al., 2016). Therefore we can effectively use many more threads than we have GPU cores. Different blocks of threads can be processed independently. This GPU parallelism is called Single Instruction Multiple Thread (SIMT) (Eijkhout et al., 2016). We use GPU threads to parallelize the outermost loop of algorithm 1. Each GPU thread computes the total gradient for its assigned pixel. Therefore we can compute the gradients for thousands of pixels simultaneously.

We implement both the GPU and the CPU version twice, once in SP and once in DP. The respective versions are identical up to the change in precision.

6.2.2 Single and double precision

Modern computers usually store real numbers in the IEEE 754 single precision floating-point representation (henceforth called SP) or the IEEE 754 double precision floating-point representation (henceforth called DP) (Institute of Electrical and Electronics Engineers, 2008, see e.g. Goldberg (1991) for an overview of floating-point arithmetic). A real number $x \in \mathbb{R}$ in decimal representation is thus stored in a binary format,

$$x = \sigma \times \bar{x}_2 \times 2^e, \tag{6.13}$$

where the integer e is the exponent, the sign σ equals $+1$ or -1 , and \bar{x}_2 is a binary number satisfying $(1)_2 \leq \bar{x}_2 < (10)_2$ (Institute of Electrical and Electronics Engineers, 2008). In the remainder of this thesis we will denote the binary format by using the subscript 2, so $(1)_2$ and $(10)_2$ correspond to the numbers 1 and 2 in decimal representation. For example, the

²OpenMP is an application programming interface managed by the non-profit OpenMP Architecture Review Board, public website: <http://www.openmp.org>

³CUDA is a parallel computing platform and programming model for general computing on GPUs managed by Nvidia Corporation, public website: <https://developer.nvidia.com/cuda-zone>

number 2.25 would correspond to $\sigma = +1$, $\bar{x}_2 = (1.001)_2$, and $e = (1)_2^4$. The number of digits in \bar{x}_2 is called the precision p of the representation. According to IEEE 754, SP has a precision of $p = 24$ binary digits and an exponent $-126 \leq e \leq 127$, while DP has $p = 53$ and $-1022 \leq e \leq 1023$. SP values are stored using 4 bytes (= 32 bits) and DP values using 8 bytes (= 64 bits) (Institute of Electrical and Electronics Engineers, 2008). As a result, DP can store a number x with higher accuracy than SP, but this comes at the price of increased memory consumption and usually also reduced computing performance (e.g., Besl, 2013; Eijkhout et al., 2016).

Both DP and SP have only a limited amount of memory available and thus their accuracy is limited. We define the machine epsilon ϵ as the difference between 1 and the next larger number that can be stored using the given representation (Eijkhout et al., 2016). For SP and DP we thus have respectively $\epsilon = 2^{-23} \approx 1.2 \times 10^{-7}$ and $\epsilon = 2^{-52} \approx 2.2 \times 10^{-16}$. These errors are so small that they might seem unimportant at first, but they will be magnified by the different computing operations performed in the course of an algorithm, so that they can become very large and relevant once the final result is obtained.

To illustrate this point, we now look at a hypothetical calculator⁵. For simplicity, it does not use SP or DP, but a decimal number representation with 6 digits precision and no exponent. We compute a relatively simple function, $f(x) = x \times (\sqrt{x+1} - \sqrt{x})$. For $x = 50,000$, the result from the hypothetical calculator is 100, while the true result is 111.8, so we have a relative error of more than 10%. To understand this behavior, we take a look at the different steps which the calculator has to perform. It computes $\sqrt{50,001}$ and rounds the result to 6 digits (result: 223.609) and then it repeats these steps for $\sqrt{50,000}$ (result: 223.607). Therefore we have two rounding errors, but they are very small. However, now the calculator subtracts two almost equal numbers to obtain 000.002, so only the last number of the result is a significant digit. We have lost a lot of accuracy which we cannot recover. The subsequent multiplication does not increase the error, but it propagates it into the final result. This example illustrates that even with the high precision available in modern computers, the result of a sufficiently long and complex algorithm can be significantly affected by the chosen number representation.

DP permits a much higher accuracy than SP and therefore it is tempting to simply use it for all computations. However, this accuracy comes at the price of computing performance. As shown in table 6.1, this is particularly true for GPUs. While the theoretical maximum computing performance of a modern CPU decreases by a factor of two, the peak performance of a consumer GPU like the Nvidia GeForce GTX 1080 Ti (henceforth called Nvidia GTX) drops by two orders of magnitude. This is a significant problem for GPU-accelerated scientific software, where SP is often not accurate enough. To ameliorate this problem, graphics card

⁴In practice, the leading bit of \bar{x}_2 would be implicit and e would be stored as a biased exponent, but we can ignore such intricacies here to simplify the presentation

⁵This illustration is inspired by an example in the lecture notes of Catalin Trenchea, available online at http://www.math.pitt.edu/~trenchea/math1070/MATH1070_2_Error_and_Computer_Arithmetic.pdf

Chapter 6. High Performance Computing for cluster lensing

	Intel Xeon E5-2680 12 cores, 2.50 GHz	Nvidia GTX 1080 Ti 3584 cores, 1582 MHz	Nvidia K80 4992 cores, 875 MHz	Nvidia P100 3584 cores, 1480 MHz
DP	240 GFLOPS	354 GFLOPS	2912 GFLOPS	5304 GFLOPS
SP	480 GFLOPS	11340 GFLOPS	8736 GFLOPS	10609 GFLOPS

Table 6.1 – Theoretical maximum computing performance for our used CPU and GPU models. These values can only serve as a rough indicator of expected performance, as the real application performance will depend on many parameters such as the used algorithm and its implementation. We list the base frequency for the CPU while we use the boost frequency for the GPUs, as the CPU typically reaches the boost frequency only on a few cores and not on all cores simultaneously. We compute the CPU maximum computing performance using the following formula: Two operations per cycle \times frequency \times AVX vectorization \times number of cores (Besl, 2013). Note that the AVX factor for SP is two times larger than for DP. We use the same formula, but without the AVX factor, for GPUs. The Nvidia K80’s GK210 GPU has three times less performance in DP than in SP (Nvidia Corporation, 2015, 2014), the GTX 1080 Ti’s GP102 GPU has thirty-two times less performance (e.g., Harris, 2016), and the P100’s GP100 GPU is two times slower (Nvidia Corporation, 2016). The number of cores and the frequencies are taken from Intel Corporation (2014) and Nvidia Corporation (2017, 2016, 2015).

manufacturers introduced new hardware specifically designed to improve the DP performance. The Nvidia K80 is only three times slower in DP than in SP and the recently released Nvidia Tesla P100 (henceforth called Nvidia P100) even achieves half of its SP performance when using DP. However, these special purpose GPUs are much more expensive than regular consumer GPUs, typically by an order of magnitude. Table 6.1 shows that the SP performance of a high-end consumer GPU is comparable to the SP power of the special purpose GPUs. Thus, if it is possible to use SP instead of DP in our lensing algorithm, we would not only significantly increase the code performance on both CPUs and GPUs, but we might also be able to achieve a close to optimal performance with relatively cheap hardware.

6.2.3 Computing finite precision errors for strong lensing

We will now show that using SP in algorithm 1 is accurate enough for a large fraction of the image pixels. We restrict ourselves again to the SIE model. It is possible to generalize these results to other parametric models, but the fraction of the image for which SP is accurate enough will vary and has to be computed for each model independently.

The lens equation 6.3 maps the triangular pixels in the image plane onto triangular pixels in the source plane. We assume a HFF pixel size of 0.03 arcsec and we maximize the lensing effect by using $D_{\text{ls}}/D_{\text{os}} = 1$. As a result, the lens equation is now a simple subtraction of $\nabla\varphi_{\epsilon}(\theta_1, \theta_2)$. We now look at an observed multiple image in the image plane. Note that our ability to locate the multiple image is observationally constrained by the size of the image pixels, so there is an observational error budget on the image location of half a pixel. In addition, the algorithm lenses both the triangular pixel and the image into the source plane. It

is possible that their respective errors due to machine precision have the same magnitude but the opposite sign, and therefore the error budget shrinks by another factor of two. As a result, the value of $\nabla\varphi_\epsilon$ can be considered accurate enough if the error E is smaller than a quarter of a pixel. Thus our upper limit for the gradient error is $E_i \leq 7.5 \times 10^{-3}$ arcsec, where $i = 1, 2$.

However, this error budget does not yet account for the magnification effect of strong lensing. Background sources and distance scales appear magnified when they are strongly lensed and consequently distance scales in the image plane like pixel sizes will be de-magnified when they are mapped into the source plane. The resulting error budget for $\nabla\varphi_\epsilon$ becomes thus $E_i \leq 7.5 \times 10^{-3}$ arcsec/ M_i , where M_i is the magnification along the θ_i -axis.

In 6.3 we derive an upper bound for the error of $\nabla\varphi_\epsilon$ due to finite machine precision. We assume that the lens is a strong lensing cluster modeled with two cluster-scale SIE halos and that the finite machine precision error is invariant under translation and rotation of the image plane coordinate system. The SP upper error bound along the θ_i -axis is $\Delta(\nabla\varphi_\epsilon) \leq 1.3 \times 10^{-4}$ arcsec. Consequently, SP is accurate enough for each pixel where the magnification along each axis is $M_i < 58$. In strong lensing, we typically measure the magnification μ of the area of a multiple image and not the magnification of an axis. The measured values are typically single digits (see e.g. [Jauzac et al., 2015](#), for magnification values of a HFF cluster). While these values cannot easily be converted to axis-magnifications due to the typically arc-like shape of strongly magnified images, they strongly suggest that SP will be accurate enough for a large fraction of the image. However, SP will not be accurate enough for the whole image, as SL clusters have critical lines where the magnification diverges. In the case of the SIS, this critical line is the Einstein ring. While the magnification does not become infinite in practice, it can become very large and thus SP will no longer be accurate enough. Consequently, we can use SP for a large fraction of the image, but we also need to implement a mechanism which ensures that we compute the gradients with DP whenever SP is not enough.

6.2.4 Fixing the missing accuracy close to critical lines

We add the missing accuracy close to critical lines as follows. First, we compute $\nabla\varphi_\epsilon$ for each pixel in the image using SP. Second, we compute for each pixel

$$\begin{aligned}\delta_1(\theta_1, \theta_2) &= (\nabla\varphi_\epsilon)_1(\theta_1, \theta_2) - (\nabla\varphi_\epsilon)_1(\theta_1 - \Delta x, \theta_2), \\ \delta_2(\theta_1, \theta_2) &= (\nabla\varphi_\epsilon)_2(\theta_1, \theta_2) - (\nabla\varphi_\epsilon)_2(\theta_1, \theta_2 - \Delta x),\end{aligned}\tag{6.14}$$

which is computationally cheap because we have already computed the gradient values for all pixels. Note that δ_i corresponds to the change of the pixel length along the θ_i -axis due to lensing into the source plane. Third, we recompute $\nabla\varphi_\epsilon$ in DP for all pixels where

$$|0.03 \text{ arcsec} - \delta_i(\theta_1, \theta_2)| < 0.0012 \text{ arcsec}\tag{6.15}$$

for $i = 1$ or $i = 2$, which implies that $M_i > 25$. We have conservatively chosen a magnification limit of 25 instead of 58 to account for special cases, e.g. if there are more than two cluster-scale halos or if hundreds of galaxy-scale lenses are by chance in a configuration that maximizes their error contribution. We derive this condition by computing the pixel length in the source plane Δx_{source} along the β_1 -axis,

$$\begin{aligned} |\Delta x_{\text{source},1}(\theta_1, \theta_2)| &= |\beta_1(\theta_1, \theta_2) - \beta_1(\theta_1 - \Delta x, \theta_2)| \\ &= |\Delta x - \delta_1(\theta_1, \theta_2)|. \end{aligned} \tag{6.16}$$

An analogous relation holds for the β_2 -axis. Note that taking the absolute value of Δx_{source} is necessary because lensing can change the image parity (see e.g. the review [Kneib and Natarajan, 2011](#)). For the assumed HFF pixel scale, the condition that $M_i > 25$ translates into $\Delta x_{\text{source},i} < 0.0012$ arcsec.

6.2.5 Early performance measurements and upcoming work

We want to measure the performance gain of using HPC methods in SL. For this purpose, we measure the time which the different software implementations require to compute the gradient of a HFF-like cluster lens for each pixel of a HST image. We also compute the gradients with the current LENSTOOL software, which serves as a reference. We assume a Λ CDM cosmology with $H_0 = 70$ km/(s Mpc), $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. We use an image with 6730×6730 pixels and a pixel scale of 0.03 arcsec/pixel to simulate images from the HST *Advanced Camera for Surveys*. The galaxy cluster consists of two cluster-scale and 700 galaxy-scale halos like in the HFF cluster Abell 2744 (e.g., [Jauzac et al., 2015](#)). We model the lens using SIE halos. The lens redshift is 0.3 and all sources are at the same redshift $z_{\text{source}} = 2.0$. The velocity dispersion determined by [Jauzac et al. \(2015\)](#) for one cluster-scale halo of Abell 2744 is approximately 1200 km/s and for a galaxy halo it is roughly 150 km/s. However, these values correspond to the velocity dispersion parameter of the chosen parametric lens model, which is not identical to measured line-of-sight velocity dispersions of galaxies. The exact conversion must be computed numerically, but for our purposes a rough agreement is enough, so we can use a conversion factor of 0.85 ([Elíasdóttir et al., 2007](#)). This leads to $\sigma_v \approx 1000$ km/s and $\sigma_v \approx 130$ km/s, respectively. The resulting Einstein deflection angles for the cluster-scale and galaxy-scale halos are respectively $\theta_E \approx 22$ arcsec and $\theta_E \approx 0.37$ arcsec. The pseudo-ellipticity of the potentials is set to $\epsilon = 0.05$.

Figure 6.2 shows early run-time measurements for the reference LENSTOOL software and the different hardware implementations. Note that the SP run was performed without the accuracy correction close to critical lines, as this is not yet implemented. In addition, performance optimizations might still be possible. For example, we could only use the GTX GPU with an old version of CUDA. Thus the results shown here can only serve as a first performance indicator. In addition, the accuracy fix will reduce the performance in the SP benchmarks, while the DP results will naturally be not affected. Nevertheless, the already achieved speedup

6.3. Appendix: Finite machine precision error in $\nabla\varphi_\epsilon$ computation for SIE

of both CPU and GPU implementations with respect to the current LENSTOOL is considerable. The CPU speedup for DP is a factor three and if we use SP, we gain a factor of ten. The results are even more impressive for the GPUs. They can accelerate this section of the code by up to a factor of 60. The performance impact of using SP is strongest for the Nvidia GTX, as is expected from the hardware design.

These findings demonstrate that our HPC approach to cluster lensing is very fruitful. The use of SP instead of DP seems to be also quite promising. However, we have to caution that the accuracy fix will reduce the measured performance. In addition, it is not clear if this performance benefit will exist for other parametric lens models. While it is possible to generalize this approach to other commonly used parametric models like the Navarro-Frenk-White profile (Navarro et al., 1997, 1996) or the dPIE, this will lead to much more complicated and longer gradient computation algorithms than the one for the SIE studied in this paper. As a result, we expect that the resulting error due to finite machine precision will become bigger, too. While the exact fraction of the image for which SP is accurate enough must be explicitly calculated and studied for the respective model, it seems clear that it will be considerably lower than for the SIE. Therefore the performance gain from using single precision and accuracy correction will shrink accordingly. This might be an interesting topic for a future publication. In the near future, we will implement the accuracy fix and measure the final performance of the different hardware implementations. We will also compute which fraction of the simulated lens has to be recomputed with DP and we will investigate whether the finite machine precision error is indeed invariant under translation and rotation of the image plane coordinate system as we have assumed. The updated results from this chapter will finally be presented in Rexroth et al. (2017c, in preparation).

6.3 Appendix: Finite machine precision error in $\nabla\varphi_\epsilon$ computation for SIE

We compute the error for $\nabla\varphi_\epsilon$ for a SIE parametric model due to finite machine precision. To do so, we investigate each line of the $\nabla\varphi_\epsilon$ algorithm 1, we compute the respective error due to finite machine precision, and we propagate the resulting errors into the next line of the algorithm. We use ϵ to denote the machine epsilon and we have $\epsilon \approx 10^{-7}$ and $\epsilon \approx 10^{-16}$ for SP and DP, respectively. Thus we can neglect terms of the order $\mathcal{O}(\epsilon^2)$ and higher. We make the assumption that the computer stores the result of one line of the algorithm, which typically corresponds to one line of code, in regular registers or memory. In addition, we assume that intermediate results, which occur while processing one line of the algorithm, are stored in extended precision registers, which are e.g. typically present in x87 Floating-Point Units. As a result, we can neglect error contributions due to machine precision for these intermediate results. Note that this is no longer the case if we use Streaming SIMD Extension (SSE) or AVX registers, as these do not use extended precision.

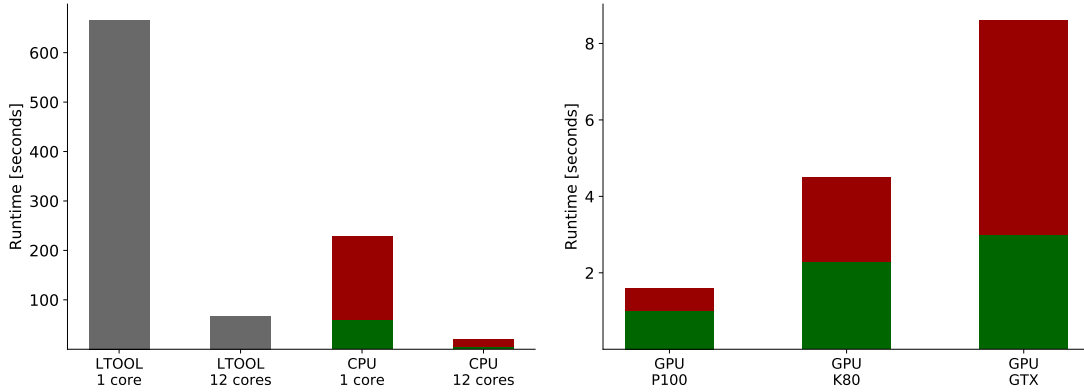


Figure 6.2 – Early performance benchmark results, which can serve as a first indicator of the expected performance. We perform the gradient computation for each pixel in the HST image of a galaxy cluster lens. The current LENSTOOL software serves as performance reference and is labeled LTOOL, the optimized CPU implementation is called CPU and the GPU implementation GPU. We also list respectively the GPU type and the number of CPU cores which we used. The green columns indicate the run-time with SP and the red columns the run-time with DP. Note that the SP implementation did not use the accuracy correction. The current LENSTOOL software only supports double precision and is thus displayed in gray.

6.3.1 Error propagation rules

We use the following error propagation rules which give upper limits on the propagated error:

Addition:

$$x \pm \epsilon a + y \pm \epsilon b = x + y \pm \epsilon(|a| + |b|) \quad (6.17)$$

Subtraction:

$$x \pm \epsilon a - y \pm \epsilon b = x - y \pm \epsilon(|a| + |b|) \quad (6.18)$$

Multiplication:

$$\begin{aligned} (x \pm \epsilon a)(y \pm \epsilon b) &= xy \pm \epsilon|ay| \pm \epsilon|xb| \pm \mathcal{O}(\epsilon^2) \\ &= xy \pm \epsilon(|ay| + |xb|) \end{aligned} \quad (6.19)$$

Division:

$$\frac{x \pm \epsilon a}{y \pm \epsilon b} = \frac{x}{y} \pm \epsilon \frac{|ay| + |bx|}{y^2 \pm \epsilon by} \quad (6.20)$$

Proof:

$$\begin{aligned} \frac{x \pm \epsilon a}{y \pm \epsilon b} - \frac{x}{y} &= \frac{(x \pm \epsilon a)y - x(y \pm \epsilon b)}{y^2 \pm \epsilon by} \\ &= \frac{xy \pm \epsilon ay - xy \pm \epsilon bx}{y^2 \pm \epsilon by} \\ &= \pm \epsilon \frac{|ay| + |bx|}{y^2 \pm \epsilon by} \end{aligned}$$

General, infinitely differentiable function $f(x)$:

We can use the Taylor expansion to first order,

$$f(x \pm \epsilon a) = f(x) \pm \epsilon a f'(x), \quad (6.21)$$

if the contribution from higher order terms is negligible:

$$\frac{f^n(x) a^n \epsilon^n}{f'(x) a \epsilon n!} \approx 0 \quad \forall n > 1,$$

where $f^n(x)$ denotes the n -th derivative.

6.3.2 Error computation

We will denote a result x stored in a regular register or memory with $\text{stored}(x)$. We want to derive an upper limit on the final error, so we will assume that each of these storage operations produces an error, $\text{stored}(x) = x \pm \epsilon x$, and we propagate these errors. Note that in practice the storing of results does not necessarily produce an error and, since the storing error is basically due to a rounding operation, errors from different storing operations can cancel each other.

We compute now the machine precision error for one SIE lens at pixel (θ_1, θ_2) :

$$\begin{aligned} \Delta\theta_1 &= \theta_1 \pm \epsilon\theta_1 - (\theta_{\text{center},1} \pm \epsilon\theta_{\text{center},1}) \\ &= \theta_1 - \theta_{\text{center},1} \pm \epsilon(|\theta_1| + |\theta_{\text{center},1}|). \end{aligned} \quad (6.22)$$

$$\begin{aligned}
 \text{stored}(\Delta\theta_1) &= \theta_1 - \theta_{\text{center},1} \pm \epsilon(|\theta_1| + |\theta_{\text{center},1}|) \pm \epsilon(|\theta_1 - \theta_{\text{center},1}|) \pm \mathcal{O}(\epsilon^2) \\
 &= \theta_1 - \theta_{\text{center},1} \pm \epsilon(|\theta_1| + |\theta_{\text{center},1}| + |\theta_1 - \theta_{\text{center},1}|) \\
 &= \Delta\theta_{1,t} \pm \epsilon A_1.
 \end{aligned} \tag{6.23}$$

In the last line, we introduced the true, error-free value of $\Delta\theta_1$, $\Delta\theta_{1,t} = \theta_1 - \theta_{\text{center},1}$. In addition, we implicitly defined the error variable A_1 , which contains all the terms which contribute to the error.

$$\begin{aligned}
 \Delta\theta_2 &= \theta_2 \pm \epsilon\theta_2 - (\theta_{\text{center},2} \pm \epsilon\theta_{\text{center},2}) \\
 &= \theta_2 - \theta_{\text{center},2} \pm \epsilon(|\theta_2| + |\theta_{\text{center},2}|).
 \end{aligned} \tag{6.24}$$

$$\begin{aligned}
 \text{stored}(\Delta\theta_2) &= \theta_2 - \theta_{\text{center},2} \pm \epsilon(|\theta_2| + |\theta_{\text{center},2}|) \pm \epsilon(|\theta_2 - \theta_{\text{center},2}|) \pm \mathcal{O}(\epsilon^2) \\
 &= \theta_2 - \theta_{\text{center},2} \pm \epsilon(|\theta_2| + |\theta_{\text{center},2}| + |\theta_2 - \theta_{\text{center},2}|) \\
 &= \Delta\theta_{2,t} \pm \epsilon A_2,
 \end{aligned} \tag{6.25}$$

where we again implicitly defined $\Delta\theta_{2,t}$ and A_2 .

$$\begin{aligned}
 \Delta\theta_1 &= (\Delta\theta_{1,t} \pm \epsilon A_1) \cos(\Phi \pm \epsilon\Phi) + (\Delta\theta_{2,t} \pm \epsilon A_2) \sin(\Phi \pm \epsilon\Phi) \\
 &= (\Delta\theta_{1,t} \pm \epsilon A_1) [\cos(\Phi) \pm \sin(\Phi)\epsilon\Phi] + (\Delta\theta_{2,t} \pm \epsilon A_2) [\sin(\Phi) \pm \cos(\Phi)\epsilon\Phi].
 \end{aligned} \tag{6.26}$$

As we have $\sin(\Phi) \leq 1$ and $\cos(\Phi) \leq 1$, we can obtain an upper bound on the error by replacing the respective sine and cosine expressions in the parts which contribute to the error with 1:

$$\begin{aligned}
 \Delta\theta_1 &= (\Delta\theta_{1,t} \pm \epsilon A_1) [\cos(\Phi) \pm \epsilon\Phi] + (\Delta\theta_{2,t} \pm \epsilon A_2) [\sin(\Phi) \pm \epsilon\Phi] \\
 &= \Delta\theta_{1,t} \cos(\Phi) \pm \epsilon(|A_1 \cos(\Phi)| + |\Delta\theta_{1,t}\Phi|) + \Delta\theta_{2,t} \sin(\Phi) \pm \epsilon(|A_2 \sin(\Phi)| + |\Delta\theta_{2,t}\Phi|) + \mathcal{O}(\epsilon^2) \\
 &= \Delta\theta_{1,t} \cos(\Phi) \pm \epsilon(|A_1| + |\Delta\theta_{1,t}\Phi|) + \Delta\theta_{2,t} \sin(\Phi) \pm \epsilon(|A_2| + |\Delta\theta_{2,t}\Phi|) \\
 &= \Delta\theta_{1,r} \pm \epsilon(|A_1| + |A_2| + |\Delta\theta_{1,t}\Phi| + |\Delta\theta_{2,t}\Phi|).
 \end{aligned} \tag{6.27}$$

In the last line we implicitly defined the true value after rotation, $\Delta\theta_{1,r}$.

6.3. Appendix: Finite machine precision error in $\nabla\varphi_\epsilon$ computation for SIE

$$\begin{aligned}
\text{stored}(\Delta\theta_1) &= \Delta\theta_{1,r} \pm \epsilon(|A_1| + |A_2| + |\Delta\theta_{1,t}\Phi| + |\Delta\theta_{2,t}\Phi| + |\Delta\theta_{1,t}\cos(\Phi)| + |\Delta\theta_{2,t}\sin(\Phi)|) + \mathcal{O}(\epsilon^2) \\
&= \Delta\theta_{1,r} \pm \epsilon(|A_1| + |A_2| + |\Delta\theta_{1,t}\Phi| + |\Delta\theta_{2,t}\Phi| + |\Delta\theta_{1,t}| + |\Delta\theta_{2,t}|) \\
&= \Delta\theta_{1,r} \pm \epsilon B,
\end{aligned} \tag{6.28}$$

where we again replaced sine and cosine with 1 and implicitly defined the error variable B .

Similarly, we obtain for $\Delta\theta_2$:

$$\begin{aligned}
\Delta\theta_2 &= (\Delta\theta_{2,t} \pm \epsilon A_2)[\cos(\Phi) \pm \epsilon\Phi] - (\Delta\theta_{1,t} \pm \epsilon A_1)[\sin(\Phi) \pm \epsilon\Phi] \\
&= \Delta\theta_{2,t}\cos(\Phi) \pm \epsilon(|A_2| + |\Delta\theta_{2,t}\Phi|) - \Delta\theta_{1,t}\sin(\Phi) \pm \epsilon(|A_1| + |\Delta\theta_{1,t}\Phi|) + \mathcal{O}(\epsilon^2) \\
&= \Delta\theta_{2,r} \pm \epsilon(|A_1| + |A_2| + |\Delta\theta_{1,t}\Phi| + |\Delta\theta_{2,t}\Phi|).
\end{aligned} \tag{6.29}$$

In the last line we implicitly defined the true value after rotation, $\Delta\theta_{2,r}$.

$$\text{stored}(\Delta\theta_2) = \Delta\theta_{2,r} \pm \epsilon B. \tag{6.30}$$

In this appendix, we denote the pseudo-ellipticity of the deflection potential with p instead of ϵ to avoid confusion with the machine epsilon. We further define $p_\star = 1 - p$, $p_\dagger = 1 + p$ and we obtain:

$$\begin{aligned}
R &= \sqrt{(\Delta\theta_{1,r} \pm \epsilon B)^2(1 - p \pm \epsilon p_\star) + (\Delta\theta_{2,r} \pm \epsilon B)^2(1 + p \pm \epsilon p_\dagger)} \\
&= \sqrt{(\Delta\theta_{1,r}^2 \pm \epsilon|2\Delta\theta_{1,r}B|)(1 - p \pm \epsilon p_\star) + (\Delta\theta_{2,r}^2 \pm \epsilon|2\Delta\theta_{2,r}B|)(1 + p \pm \epsilon p_\dagger) + \mathcal{O}(\epsilon^2)} \\
&= \sqrt{\Delta\theta_{1,r}^2(1 - p) \pm \epsilon(|\Delta\theta_{1,r}^2 p_\star| + |2\Delta\theta_{1,r}B(1 - p)|) + \Delta\theta_{2,r}^2(1 + p) \pm \epsilon(|\Delta\theta_{2,r}^2 p_\dagger| + |2\Delta\theta_{2,r}B(1 + p)|) + \mathcal{O}(\epsilon^2)} \\
&= \sqrt{\Delta\theta_{1,r}^2(1 - p) + \Delta\theta_{2,r}^2(1 + p) \pm \epsilon(|\Delta\theta_{1,r}^2 p_\star| + |\Delta\theta_{2,r}^2 p_\dagger| + |2\Delta\theta_{1,r}B(1 - p)| + |2\Delta\theta_{2,r}B(1 + p)|)}.
\end{aligned} \tag{6.31}$$

We define the true value of R ,

$$R_t = \sqrt{\Delta\theta_{1,r}^2(1 - p) + \Delta\theta_{2,r}^2(1 + p)}, \tag{6.32}$$

and use a Taylor expansion to obtain

$$R = R_t \pm \epsilon \frac{|\Delta\theta_{1,r}^2 p_\star| + |\Delta\theta_{2,r}^2 p_\dagger| + |2\Delta\theta_{1,r}B(1 - p)| + |2\Delta\theta_{2,r}B(1 + p)|}{2|R_t|}. \tag{6.33}$$

$$\begin{aligned}
 \text{stored}(R) &= R_t \pm \epsilon \left(\frac{|\Delta\theta_{1,r}^2 p_\star| + |\Delta\theta_{2,r}^2 p_\dagger| + |2\Delta\theta_{1,r} B(1-p)| + |2\Delta\theta_{2,r} B(1+p)|}{2|R_t|} + |R_t| \right) + \mathcal{O}(\epsilon^2) \\
 &= R_t \pm \epsilon C,
 \end{aligned} \tag{6.34}$$

where we implicitly defined the error variable C .

$$\begin{aligned}
 \nabla\varphi_{\epsilon,1} &= (1-p \pm \epsilon p_\star)(b_0 \pm \epsilon b_0) \frac{\Delta\theta_{1,r} \pm \epsilon B}{R_t \pm \epsilon C} \\
 &= [(1-p)b_0 \pm \epsilon(2b_0 p_\star)] \frac{\Delta\theta_{1,r} \pm \epsilon B}{R_t \pm \epsilon C} + \mathcal{O}(\epsilon^2) \\
 &= [(1-p)b_0 \pm \epsilon(2b_0 p_\star)] \left[\frac{\Delta\theta_{1,r}}{R_t} \pm \epsilon \frac{|BR_t| + |C\Delta\theta_{1,r}|}{R_t^2 \pm \epsilon|CR_t|} \right] \\
 &= (1-p)b_0 \frac{\Delta\theta_{1,r}}{R_t} \pm \epsilon \left[\left| 2b_0 p_\star \frac{\Delta\theta_{1,r}}{R_t} \right| + \left| (1-p)b_0 \frac{|BR_t| + |C\Delta\theta_{1,r}|}{R_t^2 \pm \epsilon|CR_t|} \right| \right] + \mathcal{O}(\epsilon^2) \\
 &= \nabla\varphi_{\epsilon,t,1} \pm \epsilon \left[|2\nabla\varphi_{\epsilon,t,1}| + \left| (1-p)b_0 \frac{|BR_t| + |C\Delta\theta_{1,r}|}{R_t^2 \pm \epsilon|CR_t|} \right| \right],
 \end{aligned} \tag{6.35}$$

where we implicitly defined the true value of the gradient, $\nabla\varphi_{\epsilon,t,1}$.

$$\begin{aligned}
 \text{stored}(\nabla\varphi_{\epsilon,1}) &= \nabla\varphi_{\epsilon,t,1} \pm \epsilon \left[|3\nabla\varphi_{\epsilon,t,1}| + \left| (1-p)b_0 \frac{|BR_t| + |C\Delta\theta_{1,r}|}{R_t^2 \pm \epsilon|CR_t|} \right| \right] + \mathcal{O}(\epsilon^2) \\
 &= \nabla\varphi_{\epsilon,t,1} \pm \epsilon D_1,
 \end{aligned} \tag{6.36}$$

where we implicitly defined the error variable D_1 .

6.3. Appendix: Finite machine precision error in $\nabla\varphi_\epsilon$ computation for SIE

$$\begin{aligned}
\nabla\varphi_{\epsilon,2} &= (1+p \pm \epsilon p_+)(b_0 \pm \epsilon b_0) \frac{\Delta\theta_{2,r} \pm \epsilon B}{R_t \pm \epsilon C} \\
&= [(1+p)b_0 \pm \epsilon(2b_0 p_+)] \frac{\Delta\theta_{2,r} \pm \epsilon B}{R_t \pm \epsilon C} + \mathcal{O}(\epsilon^2) \\
&= [(1+p)b_0 \pm \epsilon(2b_0 p_+)] \left[\frac{\Delta\theta_{2,r}}{R_t} \pm \epsilon \frac{|BR_t| + |C\Delta\theta_{2,r}|}{R_t^2 \pm \epsilon|CR_t|} \right] \\
&= (1+p)b_0 \frac{\Delta\theta_{2,r}}{R_t} \pm \epsilon \left[\left| 2b_0 p_+ \frac{\Delta\theta_{2,r}}{R_t} \right| + \left| (1+p)b_0 \frac{|BR_t| + |C\Delta\theta_{2,r}|}{R_t^2 \pm \epsilon|CR_t|} \right| \right] + \mathcal{O}(\epsilon^2) \\
&= \nabla\varphi_{\epsilon,t,2} \pm \epsilon \left[|2\nabla\varphi_{\epsilon,t,2}| + \left| (1+p)b_0 \frac{|BR_t| + |C\Delta\theta_{2,r}|}{R_t^2 \pm \epsilon|CR_t|} \right| \right], \tag{6.37}
\end{aligned}$$

where we implicitly defined the true value of the gradient, $\nabla\varphi_{\epsilon,t,2}$.

$$\begin{aligned}
\text{stored}(\nabla\varphi_{\epsilon,2}) &= \nabla\varphi_{\epsilon,t,2} \pm \epsilon \left[|3\nabla\varphi_{\epsilon,t,2}| + \left| (1+p)b_0 \frac{|BR_t| + |C\Delta\theta_{2,r}|}{R_t^2 \pm \epsilon|CR_t|} \right| \right] + \mathcal{O}(\epsilon^2) \\
&= \nabla\varphi_{\epsilon,t,2} \pm \epsilon D_2, \tag{6.38}
\end{aligned}$$

where we implicitly defined the error variable D_2 .

$$\begin{aligned}
\nabla\varphi_{\epsilon,1} &= (\nabla\varphi_{\epsilon,t,1} \pm \epsilon D_1) \cos(-\Phi \pm \epsilon\Phi) + (\nabla\varphi_{\epsilon,t,2} \pm \epsilon D_2) \sin(-\Phi \pm \epsilon\Phi) \\
&= (\nabla\varphi_{\epsilon,t,1} \pm \epsilon D_1) [\cos(-\Phi) \pm \sin(-\Phi)\epsilon\Phi] + (\nabla\varphi_{\epsilon,t,2} \pm \epsilon D_2) [\sin(-\Phi) \pm \cos(-\Phi)\epsilon\Phi]. \tag{6.39}
\end{aligned}$$

As we have $\sin(\Phi) \leq 1$ and $\cos(\Phi) \leq 1$, we can obtain an upper bound on the error by replacing the respective sine and cosine expressions in the parts which contribute to the error with 1:

$$\begin{aligned}
\nabla\varphi_{\epsilon,1} &= (\nabla\varphi_{\epsilon,t,1} \pm \epsilon D_1) [\cos(-\Phi) \pm \epsilon\Phi] + (\nabla\varphi_{\epsilon,t,2} \pm \epsilon D_2) [\sin(-\Phi) \pm \epsilon\Phi] \\
&= \nabla\varphi_{\epsilon,t,1} \cos(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |D_1 \cos(-\Phi)|) + \nabla\varphi_{\epsilon,t,2} \sin(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,2}\Phi| + |D_2 \sin(-\Phi)|) + \mathcal{O}(\epsilon^2) \\
&= \nabla\varphi_{\epsilon,t,1} \cos(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |D_1|) + \nabla\varphi_{\epsilon,t,2} \sin(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,2}\Phi| + |D_2|) \\
&= \nabla\varphi_{\epsilon,r,1} \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |\nabla\varphi_{\epsilon,t,2}\Phi| + |D_1| + |D_2|), \tag{6.40}
\end{aligned}$$

where we implicitly defined the true value of the first gradient component after rotation, $\nabla\varphi_{\epsilon,r,1}$.

We use the relation

$$|\nabla\varphi_{\epsilon,t,1}\cos(-\Phi) + \nabla\varphi_{\epsilon,t,2}\sin(-\Phi)| \leq |\nabla\varphi_{\epsilon,t,1}\cos(-\Phi)| + |\nabla\varphi_{\epsilon,t,2}\sin(-\Phi)| \leq |\nabla\varphi_{\epsilon,t,1}| + |\nabla\varphi_{\epsilon,t,2}| \quad (6.41)$$

to obtain:

$$\begin{aligned} \text{stored}(\nabla\varphi_{\epsilon,1}) &= \nabla\varphi_{\epsilon,r,1} \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |\nabla\varphi_{\epsilon,t,2}\Phi| + |D_1| + |D_2| + |\nabla\varphi_{\epsilon,t,1}| + |\nabla\varphi_{\epsilon,t,2}|) + \mathcal{O}(\epsilon^2) \\ &= \nabla\varphi_{\epsilon,r,1} \pm \epsilon F, \end{aligned} \quad (6.42)$$

where we implicitly defined the error variable F .

$$\begin{aligned} \nabla\varphi_{\epsilon,2} &= (\nabla\varphi_{\epsilon,t,2} \pm \epsilon D_2) \cos(-\Phi \pm \epsilon\Phi) - (\nabla\varphi_{\epsilon,t,1} \pm \epsilon D_1) \sin(-\Phi \pm \epsilon\Phi) \\ &= (\nabla\varphi_{\epsilon,t,2} \pm \epsilon D_2)[\cos(-\Phi) \pm \sin(-\Phi)\epsilon\Phi] - (\nabla\varphi_{\epsilon,t,1} \pm \epsilon D_1)[\sin(-\Phi) \pm \cos(-\Phi)\epsilon\Phi]. \end{aligned} \quad (6.43)$$

As we have $\sin(\Phi) \leq 1$ and $\cos(\Phi) \leq 1$, we can obtain an upper bound on the error by replacing the respective sine and cosine expressions in the parts which contribute to the error with 1:

$$\begin{aligned} \nabla\varphi_{\epsilon,2} &= (\nabla\varphi_{\epsilon,t,2} \pm \epsilon D_2)[\cos(-\Phi) \pm \epsilon\Phi] - (\nabla\varphi_{\epsilon,t,1} \pm \epsilon D_1)[\sin(-\Phi) \pm \epsilon\Phi] \\ &= \nabla\varphi_{\epsilon,t,2} \cos(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,2}\Phi| + |D_2 \cos(-\Phi)|) - \nabla\varphi_{\epsilon,t,1} \sin(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |D_1 \sin(-\Phi)|) + \mathcal{O}(\epsilon^2) \\ &= \nabla\varphi_{\epsilon,t,2} \cos(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,2}\Phi| + |D_2|) - \nabla\varphi_{\epsilon,t,1} \sin(-\Phi) \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |D_1|) \\ &= \nabla\varphi_{\epsilon,r,2} \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |\nabla\varphi_{\epsilon,t,2}\Phi| + |D_1| + |D_2|), \end{aligned} \quad (6.44)$$

where we implicitly defined the true value of the second gradient component after rotation, $\nabla\varphi_{\epsilon,r,2}$.

We use the relation

$$|\nabla\varphi_{\epsilon,t,2}\cos(-\Phi) - \nabla\varphi_{\epsilon,t,1}\sin(-\Phi)| \leq |\nabla\varphi_{\epsilon,t,2}\cos(-\Phi)| + |\nabla\varphi_{\epsilon,t,1}\sin(-\Phi)| \leq |\nabla\varphi_{\epsilon,t,2}| + |\nabla\varphi_{\epsilon,t,1}| \quad (6.45)$$

to obtain:

$$\begin{aligned} \text{stored}(\nabla\varphi_{\epsilon,2}) &= \nabla\varphi_{\epsilon,r,2} \pm \epsilon(|\nabla\varphi_{\epsilon,t,1}\Phi| + |\nabla\varphi_{\epsilon,t,2}\Phi| + |D_1| + |D_2| + |\nabla\varphi_{\epsilon,t,1}| + |\nabla\varphi_{\epsilon,t,2}|) + \mathcal{O}(\epsilon^2) \\ &= \nabla\varphi_{\epsilon,r,2} \pm \epsilon F. \end{aligned} \quad (6.46)$$

As a result, the total error of one computed $\nabla\varphi_{\epsilon}$ for one pixel (θ_1, θ_2) due to finite ma-

chine precision is ϵF for both gradient components.

6.3.3 Upper error bounds for cluster- and galaxy-scale SIE lenses

Let us consider a single lens at the origin of a two dimensional image plane coordinate system,

$$(\theta_{\text{center},1}, \theta_{\text{center},2}) = (0, 0). \quad (6.47)$$

We assume that the finite machine precision error is invariant under translation and rotation of the coordinate system. As a result, we can assume that the point (θ_1, θ_2) for which we compute $\nabla\varphi_\epsilon$ lies on the θ_1 -axis. We can do so without loss of generality, as this can be achieved by a simple rotation of the coordinate system. This simplifies the expression for the A terms to

$$\begin{aligned} A_1 &= 2|\theta_1|, \\ A_2 &= 0. \end{aligned} \quad (6.48)$$

We want to maximize the errors to obtain an upper bound. Therefore we maximize the angle Φ , which always appears as an error increasing factor in the error variables. Due to the symmetry of an ellipse, the largest value is $\Phi = \pi$. As a result, we have

$$B = 2|\theta_1| + \pi|\theta_1| + |\theta_1| = (3 + \pi)|\theta_1|. \quad (6.49)$$

Note that our coordinate system is now rotated by 180 degrees, so we have

$$\begin{aligned} \Delta\theta_1 &\rightarrow -\Delta\theta_1, \\ \Delta\theta_2 &\rightarrow -\Delta\theta_2. \end{aligned} \quad (6.50)$$

Next, we note that the pseudo-ellipticity p is typically small and that it appears in the error variables in connection with $\Delta\theta_1$ as a factor $1 - p$ and in connection with $\Delta\theta_2$ as a factor $1 + p$. Therefore we will minimize it and assume $p = 0$. Thus we have

$$C = \frac{|\theta_1^2| + (6 + 2\pi)|\theta_1^2|}{2|\theta_1|} + |\theta_1| = (4.5 + \pi)|\theta_1|. \quad (6.51)$$

Chapter 6. High Performance Computing for cluster lensing

The lensing effect will be maximal for a source at high redshift, so we assume $D_{os}/D_{ls} = 1$ and thus we have

$$\begin{aligned} |\nabla\varphi_{\epsilon,1}| &= |(\nabla\Psi)_1| = \theta_E, \\ |\nabla\varphi_{\epsilon,2}| &= |(\nabla\Psi)_2| = 0, \end{aligned} \tag{6.52}$$

$$b_0 = \theta_E, \tag{6.53}$$

and thus

$$\begin{aligned} D_1 &\approx 3\theta_E + \theta_E \frac{(3+\pi)|\theta_1|^2 + (4.5+\pi)|\theta_1|^2}{|\theta_1|^2} = (10.5+2\pi)\theta_E, \\ D_2 &\approx \theta_E \frac{(3+\pi)|\theta_1|^2}{|\theta_1|^2} = (3+\pi)\theta_E. \end{aligned} \tag{6.54}$$

We now rotate the coordinate system by -180 degrees,

$$\begin{aligned} \nabla\varphi_{\epsilon,1} &\rightarrow -\nabla\varphi_{\epsilon,1}, \\ \nabla\varphi_{\epsilon,2} &\rightarrow -\nabla\varphi_{\epsilon,2}, \end{aligned} \tag{6.55}$$

and we obtain

$$F = \pi\theta_E + (10.5+2\pi)\theta_E + (3+\pi)\theta_E + \theta_E \approx 27\theta_E. \tag{6.56}$$

As a result, we have for a cluster-scale halo with $\theta_E = 20$ arcsec

$$F_{\text{cluster-scale}} = 540 \text{ arcsec} \tag{6.57}$$

and for a galaxy-scale halo with $\theta_E = 0.2$ arcsec

$$F_{\text{galaxy-scale}} = 5.4 \text{ arcsec}. \tag{6.58}$$

For single and double precision, we have respectively $\epsilon_{\text{SP}} \approx 1.2 \times 10^{-7}$ and $\epsilon_{\text{DP}} \approx 2.2 \times 10^{-16}$, and thus the upper error bounds

$$\begin{aligned} \epsilon_{\text{SP}} F_{\text{cluster-scale}} &\approx 6.5 \times 10^{-5} \text{ arcsec}, \\ \epsilon_{\text{SP}} F_{\text{galaxy-scale}} &\approx 6.5 \times 10^{-7} \text{ arcsec}, \end{aligned} \tag{6.59}$$

$$\begin{aligned} \epsilon_{\text{DP}} F_{\text{cluster-scale}} &\approx 1.2 \times 10^{-13} \text{ arcsec}, \\ \epsilon_{\text{DP}} F_{\text{galaxy-scale}} &\approx 1.2 \times 10^{-15} \text{ arcsec}. \end{aligned} \tag{6.60}$$

The computed gradients for each halo are finally added up to obtain the total gradient,

$$\nabla\varphi_{\epsilon,i} = \sum_k \nabla\varphi_{\epsilon,i,k}, \tag{6.61}$$

6.3. Appendix: Finite machine precision error in $\nabla\varphi_\epsilon$ computation for SIE

and as a result, the respective errors are combined as well. However, the respective errors can have different signs and magnitudes, so we expect to see some error cancellation. We estimate the total gradient error in the following way: We neglect the contribution from the galaxy-scale halos and we add the respective upper error bounds of the cluster-scale halos. Neglecting the galaxy-scale lenses is justified, because first, their absolute errors are two orders of magnitude smaller than those of the cluster-scale halos, and second, we add many of these halos which are typically scattered throughout the image, so we expect significant error cancellation effects. We are left with typically two cluster-scale halos. The error contribution from these halos will depend on their respective parameters. To obtain an upper bound, we will add up the respective upper bounds on the gradient, so we have

$$\begin{aligned}\Delta(\nabla\varphi_{\epsilon,i})_{SP} &\approx 1.3 \times 10^{-4} \text{ arcsec}, \\ \Delta(\nabla\varphi_{\epsilon,i})_{DP} &\approx 2.4 \times 10^{-13} \text{ arcsec}.\end{aligned}\tag{6.62}$$

7 Flexion: A promising frontier

7.1 Increasing the lens model resolution with flexion

The definition of the flexions in terms of the shear derivatives in equation 3.32 and in particular the explicit formulas for \mathcal{F} and \mathcal{G} for the SIS in equation 3.44 show that flexion is much more sensitive to small-scale structure than the shear γ . Whereas we typically have

$$\gamma \propto \frac{1}{|\boldsymbol{\theta}|} \quad (7.1)$$

for the shear, we have an even quicker drop-off for the flexions,

$$\begin{aligned} \mathcal{F} &\propto \frac{1}{|\boldsymbol{\theta}|^2}, \\ \mathcal{G} &\propto \frac{1}{|\boldsymbol{\theta}|^2}. \end{aligned} \quad (7.2)$$

This theoretical advantage has been recognized and [Leonard et al. \(2009\)](#) show that this effect can be exploited in simulated lensing data. They develop an aperture mass statistic to use this potential which they further study in [Leonard and King \(2010\)](#). However, shortly thereafter [Pires and Amara \(2010\)](#) caution that flexion should never be used by itself, i.e. without adding shear information, to reconstruct mass maps, as it is less sensitive than γ to large-scale information. Indeed, flexion develops its full potential in symbiosis with shear measurements, and [Er et al. \(2010\)](#) develop a method to combine them. Nevertheless, [Bacon et al. \(2010a\)](#) demonstrate that the galaxy-galaxy flexion variance by itself can detect substructure down to $10^9 M_\odot$, but only in a statistical way which requires many lensing pairs. Recently, [Lanusse et al. \(2016\)](#) presented a very promising mass mapping algorithm for shear and flexion data. It avoids binning of the observables by treating the mass reconstruction problem as an ill-posed inverse problem which they regularize with a multi-scale wavelet sparsity prior ([Lanusse et al., 2016](#)). They show that the addition of flexion to the shear information allows the technique to recover otherwise inaccessible substructure.

However, despite all this promise, a big problem prevents most lens modelers from using flexion: The signal is almost impossible to measure.

7.2 How to measure the curvature of a few pixels?

Shortly after the introduction of flexion to the WL formalism (Goldberg and Natarajan, 2002; Irwin and Shmakova, 2003; Bacon et al., 2006), the first feasibility studies of measuring the flexion are performed (e.g. Goldberg and Bacon, 2005; Irwin et al., 2007). In the following years, three main techniques for flexion measurements are developed. One method is to use shapelets (Refregier, 2003; Refregier and Bacon, 2003) and e.g. Massey et al. (2007) propose to use polar shapelets for the measurement. In a series of papers, Okura et al. (2007), Goldberg and Leonard (2007), and Okura et al. (2008) develop the Higher Order Lensing Image's CharacteristicS (HOLICS) method, which relies on approximate equalities between certain combinations of higher multipole moments and the flexion. Finally, Cain et al. (2011) develop a fitting method to obtain the flexion.

Despite this progress, scientific results from flexion are rare. Some flexion measurements were reported (e.g. Velander et al. (2011) for galaxy-galaxy flexion and e.g. Leonard et al. (2007, 2011); Bird and Goldberg (2016) for flexion in clusters), but they can typically only detect very few signal peaks with high signal-to-noise. Using HST simulations, Rowe et al. (2013) show that the noise properties of flexion behave much worse than those for shear, which contributes to the measurement difficulties. In addition, Viola et al. (2012) show that there can be cross-talk between shear and flexion. And finally, it is simply a very challenging endeavor to precisely measure the slight curvature of a source which consists only of a few faint pixels, see figure 7.1.

7.3 Flexion measurements from an automated pipeline

Despite these difficulties, we have developed an automated pipeline to measure flexion in HST data. It is based on the HOLICS method and it includes the necessary weighting function and PSF correction terms (see Okura et al., 2008, for a discussion and derivation). The computing performance is accelerated by the HOPE just-in-time compiler (Akeret et al., 2015). In addition, we have created an extensive set of simulated flexed galaxies which we will use to calibrate our measurement pipeline, see figure 7.1. The simulation is based on the galaxy image simulation toolkit GALSIM (Rowe et al., 2015) and the flexion module GALFLEX¹. The total size of the database of simulated images is several Terabytes. Indeed, as Rowe et al. (2013) and Viola et al. (2012) show, the flexion measurement error depends on a large set of parameters, and we have varied them over the different images in our simulation. We will study the measured flexion and the associated error as a function of these parameters and use the results to calibrate our software. We will use either a principal component analysis or

¹Publicly available at <http://physics.drexel.edu/~jbird/index.html>

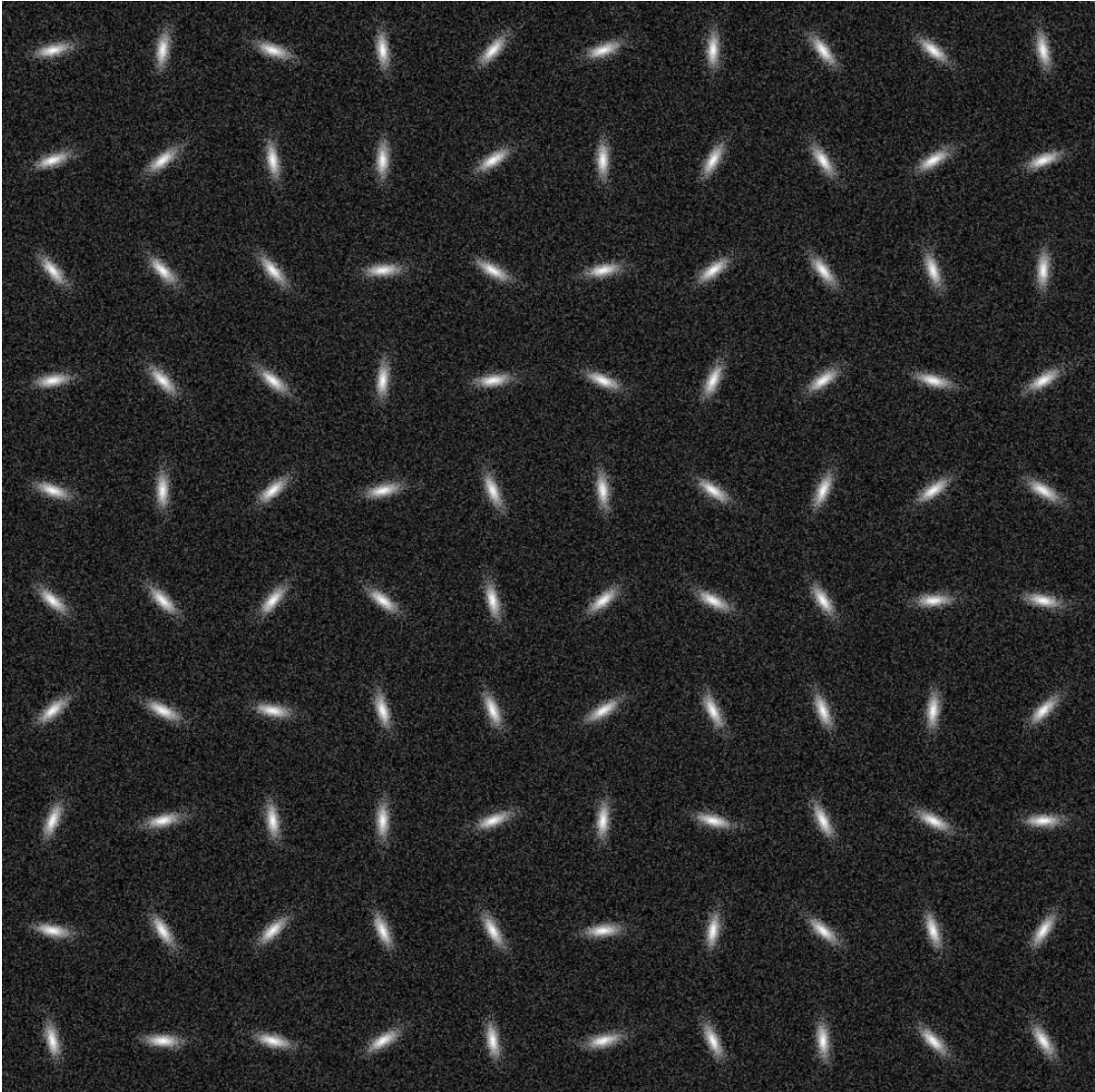


Figure 7.1 – Simulation of 100 galaxies with a strong \mathcal{F} -Flexion of 0.06 1/arcsec. The displayed galaxies have a high signal-to-noise, a half-light radius of 0.5 arcsec, and they are quite strongly flexed. Nevertheless the flexion effect is barely visible to the human eye and the large intrinsic ellipticity dominates the appearance of the galaxies. These galaxies are part of the flexion simulation which will be used to calibrate the automated measurement software.

machine learning methods for this task.

However, already an early, non-calibrated version of our software leads to very promising flexion measurements. This is shown in the publication [Rexroth \(2015\)](#), which is reprinted on the following pages with the generous permission from Cambridge University Press (E-Mail communication, 24 October 2017).

Substructure in the *Frontier Fields* from weak lensing flexion

Markus Rexroth

Laboratoire d'Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland

Abstract. Flexion is the second order weak gravitational lensing effect responsible for the arclike appearance of sources. It is highly sensitive to dark matter substructure and can greatly increase the resolution of mass maps, but it is very hard to measure. We present an automated flexion measurement pipeline for *Hubble Space Telescope* data and a preliminary application to the *Frontier Fields* cluster MACSJ0416.1-2403.

Keywords. Gravitational lensing, dark matter, clusters: individual (MACSJ0416.1-2403)

1. Introduction

In weak lensing, the unlensed 2-dimensional coordinates β_i and the lensed, observed coordinates θ_i are to first order related by $\beta_i = A_{ij}\theta_j$, where $A_{ij} = \partial\beta_i/\partial\theta_j$ is expressed in terms of the convergence κ and the shear γ . This approximation only holds if κ and γ are constant over a lensed image. Otherwise we have to expand the relation by including flexion: $\beta_i = A_{ij}\theta_j + \frac{1}{2}D_{ijk}\theta_j\theta_k$, where $D_{ijk} = \mathcal{F}_{ijk} + \mathcal{G}_{ijk}$ is the sum of the F-Flexion (spin-1) and the G-Flexion (spin-3) terms (Bacon *et al.* (2006), Goldberg & Natarajan (2002), Irwin & Shmakova (2006)). The F-Flexion shifts the centroid of a lensed source and the G-Flexion makes it triangular. The flexions are responsible for the arclets close to strong lenses. We cannot measure flexion itself in real data, but only reduced flexion, $F = \mathcal{F}/(1 - \kappa)$ and $G = \mathcal{G}/(1 - \kappa)$ (Schneider & Er (2008)). Adding flexion to weak lensing has great advantages. Typically κ and γ decline as r^{-1} , while flexion drops off as r^{-2} . Thus it is much more sensitive to small scale structure and weak lensing mass maps will have a much higher resolution (Leonard *et al.* (2009), Bacon *et al.* (2010)). Magnification maps will be more accurate. Furthermore, flexion allows us to measure signals close to the strong lensing region and thus bridges the gap between strong and weak lensing. It was demonstrated that measurements in simulations or in the strong lensing cluster Abell 1689 are in principle possible (e.g., Leonard *et al.* (2007), Okura *et al.* (2008), Rowe *et al.* (2013)). However, measurements in real data have proved to be difficult and to this day, no public measurement pipeline exists. Therefore we have developed an automated, efficient flexion pipeline for *Hubble Space Telescope* (*HST*) data.

2. An automated flexion measurement code for *HST* data

The fully automated pipeline uses the HOLICs flexion measurement technique (Okura *et al.* (2007), Okura *et al.* (2008), Goldberg & Leonard (2007)). It extends the KSB shear extraction technique (Kaiser *et al.* 1995) by including higher order image moments. In addition, our code discards overlapping sources and subtracts background noise. Flexion measurements depend on several variables, e.g. source size, signal-to-noise, and morphology (Viola *et al.* (2012), Rowe *et al.* (2013)). As a result, the measurement error is hard to estimate and several potential biases can arise. Therefore we have created simulated

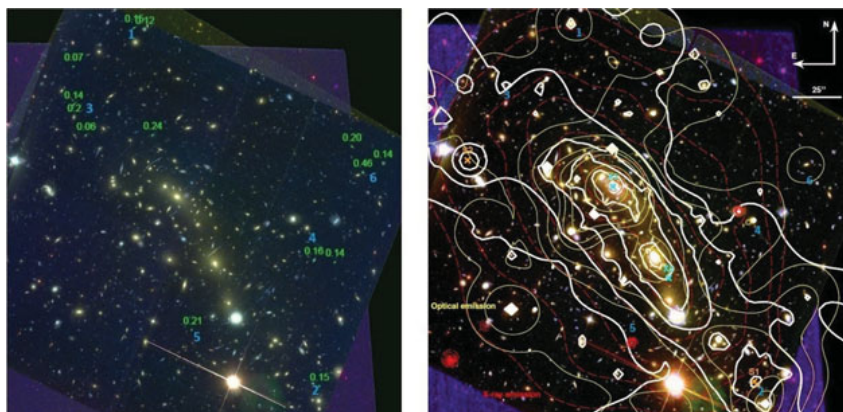


Figure 1. Preliminary F-Flexion magnitudes in the cluster MACSJ0416.1-2403 (green, left) confirm 4 substructures of the Jauzac *et al.* (2015) mass model (blue, 1 to 4) and find 2 new candidate dark matter clumps (blue, 5 and 6). White contours show the mass model, yellow lines indicate the light distribution and red contours outline the X-ray surface brightness.

images of galaxies with a wide range of different properties and flexions. We use this simulation to calibrate the pipeline, $F_{true} = m \cdot F_{meas} + c$, and analogously for G-Flexion. In addition, our code will provide a measurement error estimate.

3. Preliminary results: Application to the cluster MACSJ0416.1-2403

We applied our flexion pipeline to the *Frontier Fields* cluster MACSJ0416.1-2403. As the calibration which accounts for bias effects was not yet applied, we used only the 14 largest, most reliable sources. The measurements including the calibration of a larger sample of background galaxies will be presented in our forthcoming paper (Rexroth *et al.* (2015) in prep.). The F-Flexion confirms several substructures predicted by the high precision mass model presented in Jauzac *et al.* (2015), see Figure 1. We also find 2 new candidate dark matter clumps which the mass model could not constrain. The G-Flexion has to our knowledge never been measured in real data. We measure a G-Flexion signal that is compatible with the F-Flexion results, but has higher measurement errors. Our results show that already a small flexion sample can greatly improve mass maps.

References

- Bacon, D., Amara, A., & Read, J. 2010, *MNRAS*, 409, 389
 Bacon, D., Goldberg, D., Rowe, B., & Taylor, A. 2006, *MNRAS*, 365, 414
 Goldberg, D. & Leonard, A. 2007, *ApJ*, 660, 1003
 Goldberg, D. & Natarajan, P. 2002, *ApJ*, 564, 65
 Irwin, J. & Shmakova, M. 2006, *ApJ*, 645, 17
 Jauzac, M., Jullo, E., Eckert, D., *et al.* 2015, *MNRAS*, 446, 4132
 Kaiser, N., Squires, G., & Broadhurst, T. 1995, *ApJ*, 449, 460
 Leonard, A., Goldberg, D., Haaga, J., & Massey, R. 2007, *ApJ*, 666, 51
 Leonard, A., King, L., & Wilkins, S. 2009, *MNRAS*, 395, 1438
 Okura, Y., Umetsu, K., & Futamase, T. 2007, *ApJ*, 660, 995
 Okura, Y., Umetsu, K., & Futamase, T. 2008, *ApJ*, 680, 1
 Rowe, B., Bacon, D., Massey, R., *et al.* 2013, *MNRAS*, 435, 822
 Schneider, P. & Er, X. 2008, *A&A*, 485, 363
 Viola, M., Melchior, P., & Bartelmann, M. 2012, *MNRAS*, 419, 2215

8 Outlook and conclusion

In this thesis we have presented new developments in several areas of lensing research - observation, theory, and software development. Progress in these areas is crucial for a further advancement of lensing science and the detection of new physics, and we hope that we have contributed a small share to this development.

We have presented a novel theoretical method to break the mass-sheet degeneracy of gravitational lensing. We showed that the degeneracy parameter λ can be inferred from a simple moment comparison. In addition, we showed that we can use magnification information for this purpose without sacrificing the quality of the lens reconstruction.

We have developed a public and open source redshift extraction pipeline for the MUSE IFS. Perhaps most importantly, we have coded a wavelet-based spectrum cleaner, which is now included in the official MUSE data analysis framework. Subsequently we have used this tool to extract redshifts from observational data and we combined these redshifts with HST data to improve a cluster lens model. Even though the lens model is still preliminary and further optimizations will be studied, it already gives a better root mean square fit to the data than the previous best model.

We have investigated the potential of High Performance Computing techniques to accelerate LENSSTOOL, a widely used strong lens modeling software. We showed that it is possible to obtain a substantial speedup for a crucial part of the code. In addition, we demonstrated that further acceleration is possible if we use single precision computations. We have developed a mixed-precision algorithm which recomputes critical sections in double precision whenever necessary to obtain an accurate result, and single precision for speed otherwise. First performance measurements are promising, but the full algorithm is not yet implemented, so they can only serve as an indicator of expected computing power.

Finally, we presented an automated measurement pipeline for flexion, a higher order lensing effect which could greatly improve the resolution of mass maps. While the pipeline is not yet

Chapter 8. Outlook and conclusion

calibrated, first results for a strong lensing galaxy cluster have been obtained and they are very promising.

It is possible to combine most of these developments to create synergy effects. An exciting scenario for the future is the following. We could use our redshift extraction software to obtain as many redshifts as possible from our ongoing MUSE cluster survey. Since archival HST data is available, we could measure the flexion for these clusters. Subsequently, we use the power of LENSTOOL-HPC to simultaneously model several of the cluster lenses while incorporating the redshift and flexion information and varying the cosmological parameters. This parallelized high-resolution cluster cosmography probe could provide very interesting constraints on the Dark Matter and Dark Energy densities.

Bibliography

- B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6):061102, February 2016. doi:[10.1103/PhysRevLett.116.061102](https://doi.org/10.1103/PhysRevLett.116.061102).
- A. Acebron, E. Jullo, M. Limousin, A. Tilquin, C. Giocoli, M. Jauzac, G. Mahler, and J. Richard. Hubble Frontier Fields: systematic errors in strong lensing models of galaxy clusters - implications for cosmography. *MNRAS*, 470:1809–1825, September 2017. doi:[10.1093/mnras/stx1330](https://doi.org/10.1093/mnras/stx1330).
- G. E. Addison, Y. Huang, D. J. Watts, C. L. Bennett, M. Halpern, G. Hinshaw, and J. L. Weiland. Quantifying Discordance in the 2015 Planck CMB Spectrum. *ApJ*, 818:132, February 2016. doi:[10.3847/0004-637X/818/2/132](https://doi.org/10.3847/0004-637X/818/2/132).
- J. Akeret, L. Gamper, A. Amara, and A. Refregier. HOPE: A Python just-in-time compiler for astrophysical computations. *Astronomy and Computing*, 10:1–8, April 2015. doi:[10.1016/j.ascom.2014.12.001](https://doi.org/10.1016/j.ascom.2014.12.001).
- A. Albrecht and P. J. Steinhardt. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Physical Review Letters*, 48:1220–1223, April 1982. doi:[10.1103/PhysRevLett.48.1220](https://doi.org/10.1103/PhysRevLett.48.1220).
- C. Alcock, C. W. Akerlof, R. A. Allsman, T. S. Axelrod, D. P. Bennett, S. Chan, K. H. Cook, K. C. Freeman, K. Griest, S. L. Marshall, H.-S. Park, S. Perlmutter, B. A. Peterson, M. R. Pratt, P. J. Quinn, A. W. Rodgers, C. W. Stubbs, and W. Sutherland. Possible gravitational microlensing of a star in the Large Magellanic Cloud. *Nature*, 365:621–623, October 1993. doi:[10.1038/365621a0](https://doi.org/10.1038/365621a0).
- R. A. Alpher. A Neutron-Capture Theory of the Formation and Relative Abundance of the Elements. *Physical Review*, 74:1577–1589, December 1948. doi:[10.1103/PhysRev.74.1577](https://doi.org/10.1103/PhysRev.74.1577).
- R. A. Alpher and R. Herman. Evolution of the Universe. *Nature*, 162:774–775, November 1948. doi:[10.1038/162774b0](https://doi.org/10.1038/162774b0).
- R. A. Alpher, H. Bethe, and G. Gamow. The Origin of Chemical Elements. *Physical Review*, 73:803–804, April 1948a. doi:[10.1103/PhysRev.73.803](https://doi.org/10.1103/PhysRev.73.803).

Bibliography

- R. A. Alpher, R. Herman, and G. A. Gamow. Thermonuclear Reactions in the Expanding Universe. *Physical Review*, 74:1198–1199, November 1948b. doi:[10.1103/PhysRev.74.1198.2](https://doi.org/10.1103/PhysRev.74.1198.2).
- R. A. Alpher, J. W. Follin, and R. C. Herman. Physical Conditions in the Initial Stages of the Expanding Universe. *Physical Review*, 92:1347–1361, December 1953. doi:[10.1103/PhysRev.92.1347](https://doi.org/10.1103/PhysRev.92.1347).
- H. Atek, J. Richard, M. Jauzac, J.-P. Kneib, P. Natarajan, M. Limousin, D. Schaerer, E. Jullo, H. Ebeling, E. Egami, and B. Clement. Are Ultra-faint Galaxies at $z = 6-8$ Responsible for Cosmic Reionization? Combined Constraints from the Hubble Frontier Fields Clusters and Parallels. *ApJ*, 814:69, November 2015a. doi:[10.1088/0004-637X/814/1/69](https://doi.org/10.1088/0004-637X/814/1/69).
- H. Atek, J. Richard, J.-P. Kneib, M. Jauzac, D. Schaerer, B. Clement, M. Limousin, E. Jullo, P. Natarajan, E. Egami, and H. Ebeling. New Constraints on the Faint End of the UV Luminosity Function at $z \sim 7-8$ Using the Gravitational Lensing of the Hubble Frontier Fields Cluster A2744. *ApJ*, 800:18, February 2015b. doi:[10.1088/0004-637X/800/1/18](https://doi.org/10.1088/0004-637X/800/1/18).
- E. Aubourg, P. Bareyre, S. Bréhin, M. Gros, M. Lachièze-Rey, B. Laurent, E. Lesquoy, C. Magneville, A. Milsztajn, L. Moscoso, F. Queinnec, J. Rich, M. Spiro, L. Vigroux, S. Zylberajch, R. Ansari, F. Cavalier, M. Moniez, J.-P. Beaulieu, R. Ferlet, P. Grison, A. Vidal-Madjar, J. Guibert, O. Moreau, F. Tajahmady, E. Maurice, L. Prévôt, and C. Gry. Evidence for gravitational microlensing by dark objects in the Galactic halo. *Nature*, 365: 623–625, October 1993. doi:[10.1038/365623a0](https://doi.org/10.1038/365623a0).
- R. Avila, N. Groggin, J. Anderson, A. Bellini, R. Bohlin, D. Borncamp, M. Chiaberge, D. Coe, S. Hoffman, V. Kozhurina-Platais, R. Lucas, A. Maybhatte, M. McMaster, N. Miles, and J. Ryon. *Advanced Camera for Surveys Instrument Handbook, Version 16.0*. STScI, Baltimore, January 2017. URL <http://adsabs.harvard.edu/abs/2017acsi.book.....A>.
- W. Baade. The Period-Luminosity Relation of the Cepheids. *PASP*, 68:5, February 1956. doi:[10.1086/126870](https://doi.org/10.1086/126870).
- D. J. Bacon, A. R. Refregier, and R. S. Ellis. Detection of weak gravitational lensing by large-scale structure. *MNRAS*, 318:625–640, October 2000. doi:[10.1046/j.1365-8711.2000.03851.x](https://doi.org/10.1046/j.1365-8711.2000.03851.x).
- D. J. Bacon, D. M. Goldberg, B. T. P. Rowe, and A. N. Taylor. Weak gravitational flexion. *MNRAS*, 365:414–428, January 2006. doi:[10.1111/j.1365-2966.2005.09624.x](https://doi.org/10.1111/j.1365-2966.2005.09624.x). URL <http://adsabs.harvard.edu/abs/2006MNRAS.365..414B>.
- D. J. Bacon, A. Amara, and J. I. Read. Measuring dark matter substructure with galaxy-galaxy flexion statistics. *MNRAS*, 409:389–395, November 2010a. doi:[10.1111/j.1365-2966.2010.17316.x](https://doi.org/10.1111/j.1365-2966.2010.17316.x). URL <http://adsabs.harvard.edu/abs/2010MNRAS.409..389B>.
- R. Bacon, M. Accardo, L. Adjali, H. Anwand, S. Bauer, I. Biswas, J. Blaizot, D. Boudon, S. Brau-Nogue, J. Brinchmann, P. Caillier, L. Caponni, C. M. Carollo, T. Contini, P. Couderc,

- E. Daguisé, S. Deiries, B. Delabre, S. Dreizler, J. Dubois, M. Dupieux, C. Dupuy, E. Emselfem, T. Fechner, A. Fleischmann, M. François, G. Gallou, T. Gharsa, A. Glindemann, D. Gojak, B. Guiderdoni, G. Hansali, T. Hahn, A. Jarno, A. Kelz, C. Koehler, J. Kosmalski, F. Laurent, M. Le Floch, S. J. Lilly, J.-L. Lizon, M. Louprias, A. Manescau, C. Monstein, H. Nicklas, J.-C. Olaya, L. Pares, L. Pasquini, A. Pécontal-Rousset, R. Pelló, C. Petit, E. Popow, R. Reiss, A. Remillieux, E. Renault, M. Roth, G. Rupprecht, D. Serre, J. Schaye, G. Soucail, M. Steinmetz, O. Streicher, R. Stuik, H. Valentin, J. Vernet, P. Weilbacher, L. Wisotzki, and N. Yerle. The MUSE second-generation VLT instrument. In *Ground-based and Airborne Instrumentation for Astronomy III*, volume 7735 of *Proc. SPIE*, page 773508, July 2010b. doi:[10.1117/12.856027](https://doi.org/10.1117/12.856027).
- R. Bacon, J. Brinchmann, J. Richard, T. Contini, A. Drake, M. Franx, S. Tacchella, J. Vernet, L. Wisotzki, J. Blaizot, N. Bouché, R. Bouwens, S. Cantalupo, C. M. Carollo, D. Carton, J. Caruana, B. Clément, S. Dreizler, B. Epinat, B. Guiderdoni, C. Herenz, T.-O. Husser, S. Kamann, J. Kerutt, W. Kollatschny, D. Krajnovic, S. Lilly, T. Martinsson, L. Michel-Dansac, V. Patricio, J. Schaye, M. Shirazi, K. Soto, G. Soucail, M. Steinmetz, T. Urrutia, P. Weilbacher, and T. de Zeeuw. The MUSE 3D view of the Hubble Deep Field South. *A&A*, 575:A75, March 2015. doi:[10.1051/0004-6361/201425419](https://doi.org/10.1051/0004-6361/201425419).
- J. N. Bahcall, W. L. W. Sargent, and M. Schmidt. An Analysis of the Absorption Spectrum of 3c 191. *ApJ*, 149:L11, July 1967. doi:[10.1086/180042](https://doi.org/10.1086/180042).
- M. Bartelmann and R. Narayan. The Lens Parallax Method: Determining Redshifts of Faint Blue Galaxies through Gravitational Lensing. *ApJ*, 451:60, September 1995. doi:[10.1086/176200](https://doi.org/10.1086/176200). URL <http://adsabs.harvard.edu/abs/1995ApJ...451...60B>.
- M. Bartelmann and P. Schneider. Weak gravitational lensing. *Phys. Rep.*, 340:291–472, January 2001. doi:[10.1016/S0370-1573\(00\)00082-X](https://doi.org/10.1016/S0370-1573(00)00082-X). URL <http://adsabs.harvard.edu/abs/2001PhR...340..291B>.
- M. Bartelmann, R. Narayan, S. Seitz, and P. Schneider. Maximum-likelihood Cluster Reconstruction. *ApJ*, 464:L115, June 1996. doi:[10.1086/310114](https://doi.org/10.1086/310114). URL <http://adsabs.harvard.edu/abs/1996ApJ...464L.115B>.
- M. Bartelmann, A. Huss, J. M. Colberg, A. Jenkins, and F. R. Pearce. Arc statistics with realistic cluster potentials. IV. Clusters in different cosmologies. *A&A*, 330:1–9, February 1998.
- C. Bellhouse, Y. L. Jaffé, G. K. T. Hau, S. L. McGee, B. M. Poggianti, A. Moretti, M. Gullieszik, D. Bettoni, G. Fasano, M. D’Onofrio, J. Fritz, A. Omizzolo, Y.-K. Sheen, and B. Vulcani. GASP. II. A MUSE View of Extreme Ram-Pressure Stripping along the Line of Sight: Kinematics of the Jellyfish Galaxy JO201. *ApJ*, 844:49, July 2017. doi:[10.3847/1538-4357/aa7875](https://doi.org/10.3847/1538-4357/aa7875).
- A. Belopolsky. Die Fixsterne und extra-galaktischen Nebel. *Astronomische Nachrichten*, 236:357, October 1929. doi:[10.1002/asna.19292362204](https://doi.org/10.1002/asna.19292362204).

Bibliography

- E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *A&AS*, 117:393–404, June 1996.
- P. Besl. A case study comparing Arrays of Structures and Structures of Arrays data layouts for a compute-intensive loop run on Intel Xeon processors and Intel Xeon Phi product family coprocessors [Online; accessed September 29, 2017], 2013. URL <https://software.intel.com/sites/default/files/article/392271/aos-to-soa-optimizations-using-iterative-closest-point-mini-app.pdf>.
- H. A. Bethe. Energy Production in Stars. *Physical Review*, 55:434–456, March 1939. doi:10.1103/PhysRev.55.434.
- D. Bina, R. Pelló, J. Richard, J. Lewis, V. Patricio, S. Cantalupo, E. C. Herenz, K. Soto, P. M. Weilbacher, R. Bacon, J. D. R. Vernet, L. Wisotzki, B. Clément, J. G. Cuby, D. J. Lagattuta, G. Soucail, and A. Verhamme. MUSE observations of the lensing cluster Abell 1689. *A&A*, 590:A14, May 2016. doi:10.1051/0004-6361/201527913.
- J. P. Bird and D. M. Goldberg. Flexion in Abell 2744. *ArXiv e-prints*, November 2016.
- M. Boldrin, C. Giocoli, M. Meneghetti, L. Moscardini, G. Tormen, and A. Biviano. Cosmology through arc statistics I: sensitivity to Ω_m and σ_8 . *MNRAS*, 457:2738–2748, April 2016. doi:10.1093/mnras/stw140.
- H. Bondi and T. Gold. The Steady-State Theory of the Expanding Universe. *MNRAS*, 108:252, 1948. doi:10.1093/mnras/108.3.252.
- V. Bonvin, F. Courbin, S. H. Suyu, P. J. Marshall, C. E. Rusu, D. Sluse, M. Tewes, K. C. Wong, T. Collett, C. D. Fassnacht, T. Treu, M. W. Auger, S. Hilbert, L. V. E. Koopmans, G. Meylan, N. Rumbaugh, A. Sonnenfeld, and C. Spiniello. H0LiCOW - V. New COSMOGRAIL time delays of HE 0435-1223: H_0 to 3.8 per cent precision from strong lensing in a flat Λ CDM model. *MNRAS*, 465:4914–4930, March 2017. doi:10.1093/mnras/stw3006.
- M. Bradač, M. Lombardi, and P. Schneider. Mass-sheet degeneracy: Fundamental limit on the cluster mass reconstruction from statistical (weak) lensing. *A&A*, 424:13–22, September 2004. doi:10.1051/0004-6361:20035744. URL <http://adsabs.harvard.edu/abs/2004A%26A...424...13B>.
- M. Bradač, S. W. Allen, T. Treu, H. Ebeling, R. Massey, R. G. Morris, A. von der Linden, and D. Applegate. Revealing the Properties of Dark Matter in the Merging Cluster MACS J0025.4-1222. *ApJ*, 687:959-967, November 2008. doi:10.1086/591246.
- T. G. Brainerd, R. D. Blandford, and I. Smail. Weak Gravitational Lensing by Galaxies. *ApJ*, 466:623, August 1996. doi:10.1086/177537.
- J. Brinchmann, H. Inami, R. Bacon, T. Contini, M. Maseda, J. Chevallard, N. Bouché, L. Boogaard, M. Carollo, S. Charlot, W. Kollatschny, R. A. Marino, R. Pello, J. Richard, J. Schaye, A. Verhamme, and L. Wisotzki. The MUSE Hubble Ultra Deep Field Survey: III. Testing photometric redshifts to 30th magnitude. *ArXiv e-prints*, October 2017.

- T. J. Broadhurst, A. N. Taylor, and J. A. Peacock. Mapping cluster mass distributions via gravitational lensing of background galaxies. *ApJ*, 438:49–61, January 1995. doi:[10.1086/175053](https://doi.org/10.1086/175053). URL <http://adsabs.harvard.edu/abs/1995ApJ...438...49B>.
- V. Bromm and N. Yoshida. The First Galaxies. *ARA&A*, 49:373–407, September 2011. doi:[10.1146/annurev-astro-081710-102608](https://doi.org/10.1146/annurev-astro-081710-102608).
- E. M. Burbidge, G. R. Burbidge, W. A. Fowler, and F. Hoyle. Synthesis of the Elements in Stars. *Reviews of Modern Physics*, 29:547–650, 1957. doi:[10.1103/RevModPhys.29.547](https://doi.org/10.1103/RevModPhys.29.547).
- B. Cain, P. L. Schechter, and M. W. Bautz. Measuring Gravitational Lensing Flexion in A1689 Using an Analytic Image Model. *ApJ*, 736:43, July 2011. doi:[10.1088/0004-637X/736/1/43](https://doi.org/10.1088/0004-637X/736/1/43).
- G. B. Caminha, C. Grillo, P. Rosati, I. Balestra, W. Karman, M. Lombardi, A. Mercurio, M. Nonino, P. Tozzi, A. Zitrin, A. Biviano, M. Girardi, A. M. Koekemoer, P. Melchior, M. Meneghetti, E. Munari, S. H. Suyu, K. Umetsu, M. Annunziatella, S. Borgani, T. Broadhurst, K. I. Caputi, D. Coe, C. Delgado-Correal, S. Etori, A. Fritz, B. Frye, R. Gobat, C. Maier, A. Monna, M. Postman, B. Sartoris, S. Seitz, E. Vanzella, and B. Ziegler. CLASH-VLT: A highly precise strong lensing model of the galaxy cluster RXC J2248.7-4431 (Abell S1063) and prospects for cosmography. *A&A*, 587:A80, March 2016. doi:[10.1051/0004-6361/201527670](https://doi.org/10.1051/0004-6361/201527670).
- G. B. Caminha, C. Grillo, P. Rosati, I. Balestra, A. Mercurio, E. Vanzella, A. Biviano, K. I. Caputi, C. Delgado-Correal, W. Karman, M. Lombardi, M. Meneghetti, B. Sartoris, and P. Tozzi. A refined mass distribution of the cluster MACS J0416.1-2403 from a new large set of spectroscopic multiply lensed sources. *A&A*, 600:A90, April 2017. doi:[10.1051/0004-6361/201629297](https://doi.org/10.1051/0004-6361/201629297).
- S. Chandrasekhar and L. R. Henrich. An Attempt to Interpret the Relative Abundances of the Elements and Their Isotopes. *ApJ*, 95:288, March 1942. doi:[10.1086/144395](https://doi.org/10.1086/144395).
- K. Chang and S. Refsdal. Flux variations of QSO 0957+561 A, B and image splitting by stars near the light path. *Nature*, 282:561–564, December 1979. doi:[10.1038/282561a0](https://doi.org/10.1038/282561a0).
- O. Chwolson. Über eine mögliche Form fiktiver Doppelsterne. *Astronomische Nachrichten*, 221:329, June 1924.
- D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky. A Direct Empirical Proof of the Existence of Dark Matter. *ApJ*, 648:L109–L113, September 2006. doi:[10.1086/508162](https://doi.org/10.1086/508162).
- S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk, I. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, M. Colless, C. Collins, W. Couch, N. J. G. Cross, G. Dalton, V. R. Eke, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, K. Glazebrook, C. Jackson, A. Jenkins, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick,

Bibliography

- B. A. Peterson, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *MNRAS*, 362:505–534, September 2005. doi:[10.1111/j.1365-2966.2005.09318.x](https://doi.org/10.1111/j.1365-2966.2005.09318.x).
- S. Conseil, R. Bacon, L. Piqueras, and M. Shepherd. Advanced Data Reduction for the MUSE Deep Fields. *ArXiv e-prints*, December 2016.
- A. D’Aloisio and P. Natarajan. Cosmography with cluster strong lenses: the influence of sub-structure and line-of-sight haloes. *MNRAS*, 411:1628–1640, March 2011. doi:[10.1111/j.1365-2966.2010.17795.x](https://doi.org/10.1111/j.1365-2966.2010.17795.x).
- M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292:371–394, May 1985. doi:[10.1086/163168](https://doi.org/10.1086/163168).
- W. de Sitter. Einstein’s theory of gravitation and its astronomical consequences. Third paper. *MNRAS*, 78:3–28, November 1917. doi:[10.1093/mnras/78.1.3](https://doi.org/10.1093/mnras/78.1.3).
- T. Delubac, J. E. Bautista, N. G. Busca, J. Rich, D. Kirkby, S. Bailey, A. Font-Ribera, A. Slosar, K.-G. Lee, M. M. Pieri, J.-C. Hamilton, É. Aubourg, M. Blomqvist, J. Bovy, J. Brinkmann, W. Carithers, K. S. Dawson, D. J. Eisenstein, S. G. A. Gontcho, J.-P. Kneib, J.-M. Le Goff, D. Margala, J. Miralda-Escudé, A. D. Myers, R. C. Nichol, P. Noterdaeme, R. O’Connell, M. D. Olmstead, N. Palanque-Delabrouille, I. Pâris, P. Petitjean, N. P. Ross, G. Rossi, D. J. Schlegel, D. P. Schneider, D. H. Weinberg, C. Yèche, and D. G. York. Baryon acoustic oscillations in the Ly α forest of BOSS DR11 quasars. *A&A*, 574:A59, February 2015. doi:[10.1051/0004-6361/201423969](https://doi.org/10.1051/0004-6361/201423969).
- DES Collaboration, T. M. C. Abbott, F. B. Abdalla, A. Alarcon, J. Aleksić, S. Allam, S. Allen, A. Amara, J. Annis, J. Asorey, S. Avila, D. Bacon, E. Balbinot, M. Banerji, N. Banik, W. Barkhouse, M. Baumer, E. Baxter, K. Bechtol, M. R. Becker, A. Benoit-Lévy, B. A. Benson, G. M. Bernstein, E. Bertin, J. Blazek, S. L. Bridle, D. Brooks, D. Brout, E. Buckley-Geer, D. L. Burke, M. T. Busha, D. Capozzi, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, R. Cawthon, C. Chang, N. Chen, M. Childress, A. Choi, C. Conselice, R. Crittenden, M. Crocce, C. E. Cunha, C. B. D’Andrea, L. N. da Costa, R. Das, T. M. Davis, C. Davis, J. De Vicente, D. L. DePoy, J. DeRose, S. Desai, H. T. Diehl, J. P. Dietrich, S. Dodelson, P. Doel, A. Drlica-Wagner, T. F. Eifler, A. E. Elliott, F. Elsner, J. Elvin-Poole, J. Estrada, A. E. Evrard, Y. Fang, E. Fernandez, A. Ferté, D. A. Finley, B. Flaugher, P. Fosalba, O. Friedrich, J. Frieman, J. García-Bellido, M. Garcia-Fernandez, M. Gatti, E. Gaztanaga, D. W. Gerdes, T. Giannantonio, M. S. S. Gill, K. Glazebrook, D. A. Goldstein, D. Gruen, R. A. Gruendl, J. Gschwend, G. Gutierrez, S. Hamilton, W. G. Hartley, S. R. Hinton, K. Honscheid, B. Hoyle, D. Huterer, B. Jain, D. J. James, M. Jarvis, T. Jeltema, M. D. Johnson, M. W. G. Johnson, T. Kacprzak, S. Kent, A. G. Kim, A. King, D. Kirk, N. Kokron, A. Kovacs, E. Krause, C. Krawiec, A. Kremin, K. Kuehn, S. Kuhlmann, N. Kuropatkin, F. Lacasa, O. Lahav, T. S. Li, A. R. Liddle, C. Lidman, M. Lima, H. Lin, N. MacCrann, M. A. G. Maia, M. Makler, M. Manera, M. March, J. L. Marshall, P. Martini,

- R. G. McMahon, P. Melchior, F. Menanteau, R. Miquel, V. Miranda, D. Mudd, J. Muir, A. Möller, E. Neilsen, R. C. Nichol, B. Nord, P. Nugent, R. L. C. Ogando, A. Palmese, J. Peacock, H. V. Peiris, J. Peoples, W. J. Percival, D. Petravick, A. A. Plazas, A. Porredon, J. Prat, A. Pujol, M. M. Rau, A. Refregier, P. M. Ricker, N. Roe, R. P. Rollins, A. K. Romer, A. Roodman, R. Rosenfeld, A. J. Ross, E. Rozo, E. S. Rykoff, M. Sako, A. I. Salvador, S. Samuroff, C. Sánchez, E. Sanchez, B. Santiago, V. Scarpine, R. Schindler, D. Scolnic, L. F. Secco, S. Serrano, I. Sevilla-Noarbe, E. Sheldon, R. C. Smith, M. Smith, J. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, G. Tarle, D. Thomas, M. A. Troxel, D. L. Tucker, B. E. Tucker, S. A. Uddin, T. N. Varga, P. Vielzeuf, V. Vikram, A. K. Vivas, A. R. Walker, M. Wang, R. H. Wechsler, J. Weller, W. Wester, R. C. Wolf, B. Yanny, F. Yuan, A. Zenteno, B. Zhang, Y. Zhang, and J. Zuntz. Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing. *ArXiv e-prints*, August 2017.
- R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. *ApJ*, 142:414–419, July 1965. doi:[10.1086/148306](https://doi.org/10.1086/148306).
- M. Donahue, G. M. Voit, A. Mahdavi, K. Umetsu, S. Ettori, J. Merten, M. Postman, A. Hoffer, A. Baldi, D. Coe, N. Czakon, M. Bartelmann, N. Benitez, R. Bouwens, L. Bradley, T. Broadhurst, H. Ford, F. Gastaldello, C. Grillo, L. Infante, S. Jouvel, A. Koekemoer, D. Kelson, O. Lahav, D. Lemze, E. Medezinski, P. Melchior, M. Meneghetti, A. Molino, J. Moustakas, L. A. Moustakas, M. Nonino, P. Rosati, J. Sayers, S. Seitz, A. Van der Wel, W. Zheng, and A. Zitrin. CLASH-X: A Comparison of Lensing and X-Ray Techniques for Measuring the Mass Profiles of Galaxy Clusters. *ApJ*, 794:136, October 2014. doi:[10.1088/0004-637X/794/2/136](https://doi.org/10.1088/0004-637X/794/2/136).
- S. Dye, A. N. Taylor, T. R. Greve, Ö. E. Rögnvaldsson, E. van Kampen, P. Jakobsson, V. S. Sigmundsson, E. H. Gudmundsson, and J. Hjorth. Lens magnification by CL0024+1654 in the U and R band. *A&A*, 386:12–30, April 2002. doi:[10.1051/0004-6361:20020226](https://doi.org/10.1051/0004-6361:20020226).
- F. W. Dyson, A. S. Eddington, and C. Davidson. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London Series A*, 220:291–333, 1920. doi:[10.1098/rsta.1920.0009](https://doi.org/10.1098/rsta.1920.0009).
- H. Ebeling, A. C. Edge, and J. P. Henry. MACS: A Quest for the Most Massive Galaxy Clusters in the Universe. *ApJ*, 553:668–676, June 2001. doi:[10.1086/320958](https://doi.org/10.1086/320958).
- H. Ebeling, A. C. Edge, A. Mantz, E. Barrett, J. P. Henry, C. J. Ma, and L. van Speybroeck. The X-ray brightest clusters of galaxies from the Massive Cluster Survey. *MNRAS*, 407: 83–93, September 2010. doi:[10.1111/j.1365-2966.2010.16920.x](https://doi.org/10.1111/j.1365-2966.2010.16920.x).
- H. Ebeling, L. N. Stephenson, and A. C. Edge. Jellyfish: Evidence of Extreme Ram-pressure Stripping in Massive Galaxy Clusters. *ApJ*, 781:L40, February 2014. doi:[10.1088/2041-8205/781/2/L40](https://doi.org/10.1088/2041-8205/781/2/L40).
- G. Efstathiou and P. Lemos. Problems with KiDS. *ArXiv e-prints*, July 2017.

Bibliography

- V. Eijkhout, R. van de Geijn, and E. Chow. *Introduction to High Performance Scientific Computing*. lulu.com, 2016. ISBN 978-1257992546. doi:[10.5281/zenodo.49897](https://doi.org/10.5281/zenodo.49897).
- A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354: 769–822, 1916. doi:[10.1002/andp.19163540702](https://doi.org/10.1002/andp.19163540702).
- A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, Seite 142-152., 1917.
- A. Einstein. Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field. *Science*, 84:506–507, December 1936. doi:[10.1126/science.84.2188.506](https://doi.org/10.1126/science.84.2188.506).
- A. Einstein and W. de Sitter. On the Relation between the Expansion and the Mean Density of the Universe. *Proceedings of the National Academy of Science*, 18:213–214, March 1932. doi:[10.1073/pnas.18.3.213](https://doi.org/10.1073/pnas.18.3.213).
- D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, S. F. Anderson, J. Annis, N. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, A. Connolly, I. Csabai, M. Doi, M. Fukugita, J. A. Frieman, K. Glazebrook, J. E. Gunn, J. S. Hendry, G. Hennessy, Z. Ivezić, S. Kent, G. R. Knapp, H. Lin, Y.-S. Loh, R. H. Lupton, B. Margon, T. A. McKay, A. Meiksin, J. A. Munn, A. Pope, M. W. Richmond, D. Schlegel, D. P. Schneider, K. Shimasaku, C. Stoughton, M. A. Strauss, M. SubbaRao, A. S. Szalay, I. Szapudi, D. L. Tucker, B. Yanny, and D. G. York. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *ApJ*, 633:560–574, November 2005. doi:[10.1086/466512](https://doi.org/10.1086/466512).
- Á. Elíasdóttir, M. Limousin, J. Richard, J. Hjorth, J.-P. Kneib, P. Natarajan, K. Pedersen, E. Jullo, and D. Paraficz. Where is the matter in the Merging Cluster Abell 2218? *ArXiv e-prints*, October 2007. URL <http://adsabs.harvard.edu/abs/2007arXiv0710.5636E>.
- X. Er, G. Li, and P. Schneider. Mass reconstruction by gravitational shear and flexion. *ArXiv e-prints*, August 2010. URL <http://adsabs.harvard.edu/abs/2010arXiv1008.3088E>.
- S. M. Faber and R. E. Jackson. Velocity dispersions and mass-to-light ratios for elliptical galaxies. *ApJ*, 204:668–683, March 1976. doi:[10.1086/154215](https://doi.org/10.1086/154215).
- B. Fort, Y. Mellier, and M. Dantel-Fort. Distribution of galaxies at large redshift and cosmological parameters from the magnification bias in CL 0024+1654. *A&A*, 321:353–362, May 1997.
- A. Friedmann. Über die Krümmung des Raumes. *Zeitschrift für Physik*, 10:377–386, 1922. doi:[10.1007/BF01332580](https://doi.org/10.1007/BF01332580).
- A. Friedmann. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift für Physik*, 21:326–332, December 1924. doi:[10.1007/BF01328280](https://doi.org/10.1007/BF01328280).
- G. Gamow. Expanding Universe and the Origin of Elements. *Physical Review*, 70:572–573, October 1946. doi:[10.1103/PhysRev.70.572.2](https://doi.org/10.1103/PhysRev.70.572.2).

- G. Gamow. The Evolution of the Universe. *Nature*, 162:680–682, October 1948. doi:[10.1038/162680a0](https://doi.org/10.1038/162680a0).
- M. J. Geller and J. P. Huchra. Mapping the universe. *Science*, 246:897–903, November 1989. doi:[10.1126/science.246.4932.897](https://doi.org/10.1126/science.246.4932.897).
- J. Gilmore and P. Natarajan. Cosmography with cluster strong lensing. *MNRAS*, 396:354–364, June 2009. doi:[10.1111/j.1365-2966.2009.14612.x](https://doi.org/10.1111/j.1365-2966.2009.14612.x).
- D. M. Goldberg and D. J. Bacon. Galaxy-Galaxy Flexion: Weak Lensing to Second Order. *ApJ*, 619:741–748, February 2005. doi:[10.1086/426782](https://doi.org/10.1086/426782). URL <http://adsabs.harvard.edu/abs/2005ApJ...619..741G>.
- D. M. Goldberg and A. Leonard. Measuring Flexion. *ApJ*, 660:1003–1015, May 2007. doi:[10.1086/513137](https://doi.org/10.1086/513137). URL <http://adsabs.harvard.edu/abs/2007ApJ...660.1003G>.
- D. M. Goldberg and P. Natarajan. The Galaxy Octopole Moment as a Probe of Weak-Lensing Shear Fields. *ApJ*, 564:65–72, January 2002. doi:[10.1086/324202](https://doi.org/10.1086/324202). URL <http://adsabs.harvard.edu/abs/2002ApJ...564...65G>.
- David Goldberg. What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys*, 23(1):5–48, March 1991. ISSN 0360-0300. doi:[10.1145/103162.103163](https://doi.org/10.1145/103162.103163). URL <http://doi.acm.org/10.1145/103162.103163>.
- G. Golse and J.-P. Kneib. Pseudo elliptical lensing mass model: Application to the NFW mass distribution. *A&A*, 390:821–827, August 2002. doi:[10.1051/0004-6361:20020639](https://doi.org/10.1051/0004-6361:20020639).
- G. Golse, J.-P. Kneib, and G. Soucail. Constraining the cosmological parameters using strong lensing. *A&A*, 387:788–803, June 2002. doi:[10.1051/0004-6361:20020448](https://doi.org/10.1051/0004-6361:20020448).
- C. Grillo, S. H. Suyu, P. Rosati, A. Mercurio, I. Balestra, E. Munari, M. Nonino, G. B. Caminha, M. Lombardi, G. De Lucia, S. Borgani, R. Gobat, A. Biviano, M. Girardi, K. Umetsu, D. Coe, A. M. Koekemoer, M. Postman, A. Zitrin, A. Halkola, T. Broadhurst, B. Sartoris, V. Presotto, M. Annunziatella, C. Maier, A. Fritz, E. Vanzella, and B. Frye. CLASH-VLT: Insights on the Mass Substructures in the Frontier Fields Cluster MACS J0416.1-2403 through Accurate Strong Lens Modeling. *ApJ*, 800:38, February 2015. doi:[10.1088/0004-637X/800/1/38](https://doi.org/10.1088/0004-637X/800/1/38).
- C. Grillo, W. Karman, S. H. Suyu, P. Rosati, I. Balestra, A. Mercurio, M. Lombardi, T. Treu, G. B. Caminha, A. Halkola, S. A. Rodney, R. Gavazzi, and K. I. Caputi. The Story of Supernova Refsdal Told by Muse. *ApJ*, 822:78, May 2016. doi:[10.3847/0004-637X/822/2/78](https://doi.org/10.3847/0004-637X/822/2/78).
- A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, January 1981. doi:[10.1103/PhysRevD.23.347](https://doi.org/10.1103/PhysRevD.23.347).
- M. Harris. *Mixed-Precision Programming with CUDA 8* [Online; accessed October 2, 2017], 2016. URL <https://devblogs.nvidia.com/parallelforall/mixed-precision-programming-cuda-8/>.

Bibliography

- D. Harvey, R. Massey, T. Kitching, A. Taylor, and E. Tittley. The nongravitational interactions of dark matter in colliding galaxy clusters. *Science*, 347:1462–1465, March 2015. doi:[10.1126/science.1261381](https://doi.org/10.1126/science.1261381).
- D. Harvey, J. P. Kneib, and M. Jauzac. Systematic or signal? How dark matter misalignments can bias strong lensing models of galaxy clusters. *MNRAS*, 458:660–665, May 2016. doi:[10.1093/mnras/stw295](https://doi.org/10.1093/mnras/stw295).
- C. Hayashi. Proton-Neutron Concentration Ratio in the Expanding Universe at the Stages preceding the Formation of the Elements. *Progress of Theoretical Physics*, 5:224–235, March 1950. doi:[10.1143/ptp/5.2.224](https://doi.org/10.1143/ptp/5.2.224).
- E. Hertzsprung. Über die räumliche Verteilung der Veränderlichen vom δ Cephei-Typus. *Astronomische Nachrichten*, 196:201, November 1913.
- J. N. Hewitt, E. L. Turner, D. P. Schneider, B. F. Burke, and G. I. Langston. Unusual radio source MG1131+0456 - A possible Einstein ring. *Nature*, 333:537–540, June 1988. doi:[10.1038/333537a0](https://doi.org/10.1038/333537a0).
- H. Hildebrandt, M. Viola, C. Heymans, S. Joudaki, K. Kuijken, C. Blake, T. Erben, B. Joachimi, D. Klaes, L. Miller, C. B. Morrison, R. Nakajima, G. Verdoes Kleijn, A. Amon, A. Choi, G. Covone, J. T. A. de Jong, A. Dvornik, I. Fenech Conti, A. Grado, J. Harnois-Déraps, R. Herbonnet, H. Hoekstra, F. Köhlinger, J. McFarland, A. Mead, J. Merten, N. Napolitano, J. A. Peacock, M. Radovich, P. Schneider, P. Simon, E. A. Valentijn, J. L. van den Busch, E. van Uitert, and L. Van Waerbeke. KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. *MNRAS*, 465:1454–1498, February 2017. doi:[10.1093/mnras/stw2805](https://doi.org/10.1093/mnras/stw2805).
- F. Hoyle. A New Model for the Expanding Universe. *MNRAS*, 108:372, 1948. doi:[10.1093/mnras/108.5.372](https://doi.org/10.1093/mnras/108.5.372).
- W. Hu and S. Dodelson. Cosmic Microwave Background Anisotropies. *ARA&A*, 40:171–216, 2002. doi:[10.1146/annurev.astro.40.060401.093926](https://doi.org/10.1146/annurev.astro.40.060401.093926).
- E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929. doi:[10.1073/pnas.15.3.168](https://doi.org/10.1073/pnas.15.3.168).
- E. P. Hubble. Cepheids in spiral nebulae. *The Observatory*, 48:139–142, May 1925.
- E. P. Hubble. Extragalactic nebulae. *ApJ*, 64, December 1926. doi:[10.1086/143018](https://doi.org/10.1086/143018).
- Institute of Electrical and Electronics Engineers. *IEEE Standard 754-2008 - IEEE Standard for Floating-Point Arithmetic*. IEEE, August 2008. doi:[10.1109/IEEESTD.2008.4610935](https://doi.org/10.1109/IEEESTD.2008.4610935).
- Intel Corporation. *Intel Xeon Processor E5-2680 v3 Specifications [Online; accessed October 2, 2017]*, 2014. URL https://ark.intel.com/products/81908/Intel-Xeon-Processor-E5-2680-v3-30M-Cache-2_50-GHz.

- F. Iocco, G. Mangano, G. Miele, O. Pisanti, and P. D. Serpico. Primordial nucleosynthesis: From precision cosmology to fundamental physics. *Phys. Rep.*, 472:1–76, March 2009. doi:[10.1016/j.physrep.2009.02.002](https://doi.org/10.1016/j.physrep.2009.02.002).
- J. Irwin and M. Shmakova. Observations of cluster substructure using weakly lensed sextupole moments. *ArXiv Astrophysics e-prints*, July 2003. URL <http://adsabs.harvard.edu/abs/2003astro.ph..8007I>.
- J. Irwin, M. Shmakova, and J. Anderson. Lensing Signals in the Hubble Ultra Deep Field Using All Second-Order Shape Deformations. *ApJ*, 671:1182–1195, December 2007. doi:[10.1086/522819](https://doi.org/10.1086/522819). URL <http://adsabs.harvard.edu/abs/2007ApJ...671.1182I>.
- M. J. Irwin, R. L. Webster, P. C. Hewett, R. T. Corrigan, and R. I. Jedrzejewski. Photometric variations in the Q2237 + 0305 system - First detection of a microlensing event. *AJ*, 98:1989–1994, December 1989. doi:[10.1086/115272](https://doi.org/10.1086/115272).
- M. Ishigaki, R. Kawamata, M. Ouchi, M. Oguri, K. Shimasaku, and Y. Ono. Hubble Frontier Fields First Complete Cluster Data: Faint Galaxies at $z \sim 5-10$ for UV Luminosity Functions and Cosmic Reionization. *ApJ*, 799:12, January 2015. doi:[10.1088/0004-637X/799/1/12](https://doi.org/10.1088/0004-637X/799/1/12).
- A. H. Jaffe, P. A. Ade, A. Balbi, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, K. Coble, B. P. Crill, P. de Bernardis, P. Farese, P. G. Ferreira, K. Ganga, M. Giacometti, S. Hanany, E. Hivon, V. V. Hristov, A. Iacoangeli, A. E. Lange, A. T. Lee, L. Martinis, S. Masi, P. D. Mauskopf, A. Melchiorri, T. Montroy, C. B. Netterfield, S. Oh, E. Pascale, F. Piacentini, D. Pogosyan, S. Prunet, B. Rabii, S. Rao, P. L. Richards, G. Romeo, J. E. Ruhl, F. Scarabuzzi, D. Sforna, G. F. Smoot, R. Stompor, C. D. Winant, and J. H. Wu. Cosmology from MAXIMA-1, BOOMERANG, and COBE DMR Cosmic Microwave Background Observations. *Physical Review Letters*, 86:3475–3479, April 2001. doi:[10.1103/PhysRevLett.86.3475](https://doi.org/10.1103/PhysRevLett.86.3475).
- M. Jauzac, B. Clément, M. Limousin, J. Richard, E. Jullo, H. Ebeling, H. Atek, J.-P. Kneib, K. Knowles, P. Natarajan, D. Eckert, E. Egami, R. Massey, and M. Rexroth. Hubble Frontier Fields: a high-precision strong-lensing analysis of galaxy cluster MACSJ0416.1-2403 using ~ 200 multiple images. *MNRAS*, 443:1549–1554, September 2014. doi:[10.1093/mnras/stu1355](https://doi.org/10.1093/mnras/stu1355).
- M. Jauzac, J. Richard, E. Jullo, B. Clément, M. Limousin, J.-P. Kneib, H. Ebeling, P. Natarajan, S. Rodney, H. Atek, R. Massey, D. Eckert, E. Egami, and M. Rexroth. Hubble Frontier Fields: a high-precision strong-lensing analysis of the massive galaxy cluster Abell 2744 using ~ 180 multiple images. *MNRAS*, 452:1437–1446, September 2015. doi:[10.1093/mnras/stv1402](https://doi.org/10.1093/mnras/stv1402).
- M. Jauzac, D. Eckert, J. Schwinn, D. Harvey, C. M. Baugh, A. Robertson, S. Bose, R. Massey, M. Owers, H. Ebeling, H. Y. Shan, E. Jullo, J.-P. Kneib, J. Richard, H. Atek, B. Clément, E. Egami, H. Israel, K. Knowles, M. Limousin, P. Natarajan, M. Rexroth, P. Taylor, and C. Tchernin. The extraordinary amount of substructure in the Hubble Frontier Fields cluster Abell 2744. *MNRAS*, 463:3876–3893, December 2016a. doi:[10.1093/mnras/stw2251](https://doi.org/10.1093/mnras/stw2251).

Bibliography

- M. Jauzac, J. Richard, M. Limousin, K. Knowles, G. Mahler, G. P. Smith, J.-P. Kneib, E. Jullo, P. Natarajan, H. Ebeling, H. Atek, B. Clément, D. Eckert, E. Egami, R. Massey, and M. Rexroth. Hubble Frontier Fields: predictions for the return of SN Refsdal with the MUSE and GMOS spectrographs. *MNRAS*, 457:2029–2042, April 2016b. doi:[10.1093/mnras/stw069](https://doi.org/10.1093/mnras/stw069).
- T. L. Johnson and K. Sharon. The Systematics of Strong Lens Modeling Quantified: The Effects of Constraint Selection and Redshift Information on Magnification, Mass, and Multiple Image Predictability. *ApJ*, 832:82, November 2016. doi:[10.3847/0004-637X/832/1/82](https://doi.org/10.3847/0004-637X/832/1/82).
- T. L. Johnson, K. Sharon, M. B. Bayliss, M. D. Gladders, D. Coe, and H. Ebeling. Lens Models and Magnification Maps of the Six Hubble Frontier Fields Clusters. *ApJ*, 797:48, December 2014. doi:[10.1088/0004-637X/797/1/48](https://doi.org/10.1088/0004-637X/797/1/48).
- S. Joudaki, C. Blake, C. Heymans, A. Choi, J. Harnois-Deraps, H. Hildebrandt, B. Joachimi, A. Johnson, A. Mead, D. Parkinson, M. Viola, and L. van Waerbeke. CFHTLenS revisited: assessing concordance with Planck including astrophysical systematics. *MNRAS*, 465: 2033–2052, February 2017. doi:[10.1093/mnras/stw2665](https://doi.org/10.1093/mnras/stw2665).
- S. Jouvel, O. Host, O. Lahav, S. Seitz, A. Molino, D. Coe, M. Postman, L. Moustakas, N. Benítez, P. Rosati, I. Balestra, C. Grillo, L. Bradley, A. Fritz, D. Kelson, A. M. Koekemoer, D. Lemze, E. Medezinski, A. Mercurio, J. Moustakas, M. Nonino, M. Scodeggio, W. Zheng, A. Zitrin, M. Bartelmann, R. Bouwens, T. Broadhurst, M. Donahue, H. Ford, G. Graves, L. Infante, Y. Jimenez-Teja, R. Lazkoz, P. Melchior, M. Meneghetti, J. Merten, S. Ogaz, and K. Umetsu. CLASH: Photometric redshifts with 16 HST bands in galaxy cluster fields. *A&A*, 562:A86, February 2014. doi:[10.1051/0004-6361/201322419](https://doi.org/10.1051/0004-6361/201322419).
- E. Jullo and J.-P. Kneib. Multiscale cluster lens mass mapping - I. Strong lensing modelling. *MNRAS*, 395:1319–1332, May 2009. doi:[10.1111/j.1365-2966.2009.14654.x](https://doi.org/10.1111/j.1365-2966.2009.14654.x). URL <http://adsabs.harvard.edu/abs/2009MNRAS.395.1319J>.
- E. Jullo, J.-P. Kneib, M. Limousin, Á. Elíasdóttir, P. J. Marshall, and T. Verdugo. A Bayesian approach to strong lensing modelling of galaxy clusters. *New Journal of Physics*, 9:447, December 2007. doi:[10.1088/1367-2630/9/12/447](https://doi.org/10.1088/1367-2630/9/12/447). URL <http://adsabs.harvard.edu/abs/2007NJPh...9..447J>.
- E. Jullo, P. Natarajan, J.-P. Kneib, A. D’Aloisio, M. Limousin, J. Richard, and C. Schimd. Cosmological Constraints from Strong Gravitational Lensing in Clusters of Galaxies. *Science*, 329:924–927, August 2010. doi:[10.1126/science.1185759](https://doi.org/10.1126/science.1185759).
- N. Kaiser, G. Squires, and T. Broadhurst. A Method for Weak Lensing Observations. *ApJ*, 449:460, August 1995. doi:[10.1086/176071](https://doi.org/10.1086/176071). URL <http://adsabs.harvard.edu/abs/1995ApJ...449..460K>.
- N. Kaiser, G. Wilson, and G. A. Luppino. Large-Scale Cosmic Shear Measurements. *ArXiv Astrophysics e-prints*, March 2000.

- W. Karman, K. I. Caputi, C. Grillo, I. Balestra, P. Rosati, E. Vanzella, D. Coe, L. Christensen, A. M. Koekemoer, T. Krühler, M. Lombardi, A. Mercurio, M. Nonino, and A. van der Wel. MUSE integral-field spectroscopy towards the Frontier Fields cluster Abell S1063. I. Data products and redshift identifications. *A&A*, 574:A11, February 2015. doi:[10.1051/0004-6361/201424962](https://doi.org/10.1051/0004-6361/201424962).
- A. Kassiola and I. Kovner. Elliptic Mass Distributions versus Elliptic Potentials in Gravitational Lenses. *ApJ*, 417:450, November 1993. doi:[10.1086/173325](https://doi.org/10.1086/173325). URL <http://adsabs.harvard.edu/abs/1993ApJ...417..450K>.
- S. Y. Kim, A. H. G. Peter, and D. Wittman. In the wake of dark giants: new signatures of dark matter self-interactions in equal-mass mergers of galaxy clusters. *MNRAS*, 469: 1414–1444, August 2017. doi:[10.1093/mnras/stx896](https://doi.org/10.1093/mnras/stx896).
- B. Klein. Constraining dark matter distribution in galaxy clusters with MUSE. Master's thesis, ISAE SUPAERO/Université Toulouse III/EPFL, 2017.
- Y. G. Klimov. The Deflection of Light Rays in the Gravitational Fields of Galaxies. *Soviet Physics Doklady*, 8:119, August 1963.
- J.-P. Kneib and P. Natarajan. Cluster lenses. *A&ARv*, 19:47, November 2011. doi:[10.1007/s00159-011-0047-3](https://doi.org/10.1007/s00159-011-0047-3). URL <http://adsabs.harvard.edu/abs/2011A%26ARv..19...47K>.
- J.-P. Kneib, R. S. Ellis, I. Smail, W. J. Couch, and R. M. Sharples. Hubble Space Telescope Observations of the Lensing Cluster Abell 2218. *ApJ*, 471:643, November 1996. doi:[10.1086/177995](https://doi.org/10.1086/177995).
- J.-P. Kneib, R. S. Ellis, M. R. Santos, and J. Richard. A Probable $z \sim 7$ Galaxy Strongly Lensed by the Rich Cluster A2218: Exploring the Dark Ages. *ApJ*, 607:697–703, June 2004. doi:[10.1086/386281](https://doi.org/10.1086/386281).
- H. Kragh. Gamow's Game: The Road to the Hot Big Bang. *Centaurus*, 38:335–361, December 1996. doi:[10.1111/j.1600-0498.1996.tb00020.x](https://doi.org/10.1111/j.1600-0498.1996.tb00020.x).
- H. Kragh. Big Bang: the etymology of a name. *Astronomy and Geophysics*, 54(2):2.28–2.30, April 2013. doi:[10.1093/astrogeo/att035](https://doi.org/10.1093/astrogeo/att035).
- H. Kragh. Historical aspects of post-1850 cosmology. In *American Institute of Physics Conference Series*, volume 1632 of *American Institute of Physics Conference Series*, pages 3–26, November 2014. doi:[10.1063/1.4902842](https://doi.org/10.1063/1.4902842).
- H. Kragh. Is the Universe expanding? Fritz Zwicky and the early tired-light hypothesis. *Journal of Astronomical History and Heritage*, 20:2–12, April 2017.
- H. Kragh and R. W. Smith. Who Discovered the Expanding Universe? *History of Science*, 41: 141–162, 2003.

Bibliography

- A. V. Kravtsov and S. Borgani. Formation of Galaxy Clusters. *ARA&A*, 50:353–409, September 2012. doi:[10.1146/annurev-astro-081811-125502](https://doi.org/10.1146/annurev-astro-081811-125502).
- D. J. Lagattuta, J. Richard, B. Clément, G. Mahler, V. Patrício, R. Pelló, G. Soucail, K. B. Schmidt, L. Wisotzki, J. Martinez, and D. Bina. Lens modelling Abell 370: crowning the final frontier field with MUSE. *MNRAS*, 469:3946–3964, August 2017. doi:[10.1093/mnras/stx1079](https://doi.org/10.1093/mnras/stx1079).
- F. Lanusse, J.-L. Starck, A. Leonard, and S. Pires. High resolution weak lensing mass mapping combining shear and flexion. *A&A*, 591:A2, June 2016. doi:[10.1051/0004-6361/201628278](https://doi.org/10.1051/0004-6361/201628278).
- A. Leauthaud, R. Massey, J.-P. Kneib, J. Rhodes, D. E. Johnston, P. Capak, C. Heymans, R. S. Ellis, A. M. Koekemoer, O. Le Fèvre, Y. Mellier, A. Réfrégier, A. C. Robin, N. Scoville, L. Tasca, J. E. Taylor, and L. Van Waerbeke. Weak Gravitational Lensing with COSMOS: Galaxy Selection and Shape Measurements. *ApJS*, 172:219–238, September 2007. doi:[10.1086/516598](https://doi.org/10.1086/516598). URL <http://adsabs.harvard.edu/abs/2007ApJS..172..219L>.
- H. S. Leavitt. 1777 variables in the Magellanic Clouds. *Annals of Harvard College Observatory*, 60:87–108.3, 1908.
- H. S. Leavitt and E. C. Pickering. Periods of 25 Variable Stars in the Small Magellanic Cloud. *Harvard College Observatory Circular*, 173:1–3, March 1912.
- G. Lemaître. Note on de Sitter’s Universe. *Publications du Laboratoire d’Astronomie et de Geodesie de l’Universite de Louvain*, vol. 2, pp.37–41, 2:37–41, 1926.
- G. Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles*, 47:49–59, 1927.
- G. Lemaître. The Beginning of the World from the Point of View of Quantum Theory. *Nature*, 127:706, May 1931. doi:[10.1038/127706b0](https://doi.org/10.1038/127706b0).
- A. Leonard and L. J. King. A new tool to determine masses and mass profiles using gravitational flexion. *MNRAS*, 405:1854–1866, July 2010. doi:[10.1111/j.1365-2966.2010.16573.x](https://doi.org/10.1111/j.1365-2966.2010.16573.x). URL <http://adsabs.harvard.edu/abs/2010MNRAS.405.1854L>.
- A. Leonard, D. M. Goldberg, J. L. Haaga, and R. Massey. Gravitational Shear, Flexion, and Strong Lensing in Abell 1689. *ApJ*, 666:51–63, September 2007. doi:[10.1086/520109](https://doi.org/10.1086/520109). URL <http://adsabs.harvard.edu/abs/2007ApJ...666...51L>.
- A. Leonard, L. J. King, and S. M. Wilkins. Detecting mass substructure in galaxy clusters: an aperture mass statistic for gravitational flexion. *MNRAS*, 395:1438–1448, May 2009. doi:[10.1111/j.1365-2966.2009.14546.x](https://doi.org/10.1111/j.1365-2966.2009.14546.x). URL <http://adsabs.harvard.edu/abs/2009MNRAS.395.1438L>.

- A. Leonard, L. J. King, and D. M. Goldberg. New constraints on the complex mass substructure in Abell 1689 from gravitational flexion. *MNRAS*, 413:789–804, May 2011. doi:[10.1111/j.1365-2966.2010.18171.x](https://doi.org/10.1111/j.1365-2966.2010.18171.x). URL <http://adsabs.harvard.edu/abs/2011MNRAS.413..789L>.
- S. Liebes. Gravitational Lenses. *Physical Review*, 133:835–844, February 1964. doi:[10.1103/PhysRev.133.B835](https://doi.org/10.1103/PhysRev.133.B835).
- M. Limousin, J. Richard, E. Jullo, M. Jauzac, H. Ebeling, M. Bonamigo, A. Alavi, B. Clément, C. Giocoli, J.-P. Kneib, T. Verdugo, P. Natarajan, B. Siana, H. Atek, and M. Rexroth. Strong-lensing analysis of MACS J0717.5+3745 from Hubble Frontier Fields observations: How well can the mass distribution be constrained? *A&A*, 588:A99, April 2016. doi:[10.1051/0004-6361/201527638](https://doi.org/10.1051/0004-6361/201527638).
- Marceau Limousin, Johan Richard, Eric Jullo, Jean-Paul Kneib, Bernard Fort, Geneviève Soucail, Árdís Elíasdóttir, Priyamvada Natarajan, Richard S. Ellis, Ian Smail, Oliver Czoske, Graham P. Smith, Patrick Hudelot, Sébastien Bardeau, Harald Ebeling, Eiichi Egami, and Kirsten K. Knudsen. Combining strong and weak gravitational lensing in abell 1689. *The Astrophysical Journal*, 668(2):643, 2007. URL <http://stacks.iop.org/0004-637X/668/i=2/a=643>.
- A. Linde. Inflationary Cosmology. In M. Lemoine, J. Martin, and P. Peter, editors, *Inflationary Cosmology*, volume 738 of *Lecture Notes in Physics*, Berlin Springer Verlag, 2008. doi:[10.1007/978-3-540-74353-8](https://doi.org/10.1007/978-3-540-74353-8).
- A. D. Linde. A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108:389–393, February 1982. doi:[10.1016/0370-2693\(82\)91219-9](https://doi.org/10.1016/0370-2693(82)91219-9).
- A. D. Linde. Chaotic inflation. *Physics Letters B*, 129:177–181, September 1983. doi:[10.1016/0370-2693\(83\)90837-7](https://doi.org/10.1016/0370-2693(83)90837-7).
- O. J. Lodge. Gravitation and Light. *Nature*, 104:354, December 1919. doi:[10.1038/104354a0](https://doi.org/10.1038/104354a0).
- M. S. Longair. A brief history of cosmology. In W. L. Freedman, editor, *Measuring and Modeling the Universe*, volume 2 of *Carnegie Observatories Astrophysics Series*, pages 1–18. Cambridge University Press, November 2004.
- J. M. Lotz, A. Koekemoer, D. Coe, N. Grogin, P. Capak, J. Mack, J. Anderson, R. Avila, E. A. Barker, D. Borncamp, G. Brammer, M. Durbin, H. Gunning, B. Hilbert, H. Jenkner, H. Khandrika, Z. Levay, R. A. Lucas, J. MacKenty, S. Ogaz, B. Porterfield, N. Reid, M. Robberto, P. Royle, L. J. Smith, L. J. Storrie-Lombardi, B. Sunnquist, J. Surace, D. C. Taylor, R. Williams, J. Bullock, M. Dickinson, S. Finkelstein, P. Natarajan, J. Richard, B. Robertson, J. Tumlinson, A. Zitrin, K. Flanagan, K. Sembach, B. T. Soifer, and M. Mountain. The Frontier Fields: Survey Design and Initial Results. *ApJ*, 837:97, March 2017. doi:[10.3847/1538-4357/837/1/97](https://doi.org/10.3847/1538-4357/837/1/97).

Bibliography

- R. Lynds and V. Petrosian. Giant Luminous Arcs in Galaxy Clusters. In *Bulletin of the American Astronomical Society*, volume 18 of *BAAS*, page 1014, September 1986.
- P. Madau and M. Dickinson. Cosmic Star-Formation History. *ARA&A*, 52:415–486, August 2014. doi:[10.1146/annurev-astro-081811-125615](https://doi.org/10.1146/annurev-astro-081811-125615).
- G. Mahler, J. Richard, B. Clément, D. Lagattuta, K. Schmidt, V. Patrício, G. Soucail, R. Bacon, R. Pello, R. Bouwens, M. Maseda, J. Martinez, M. Carollo, H. Inami, F. Leclercq, and L. Wisotzki. Strong lensing analysis of Abell 2744 with MUSE and Hubble Frontier Fields images. *ArXiv e-prints*, February 2017.
- A. B. Mantz, A. von der Linden, S. W. Allen, D. E. Applegate, P. L. Kelly, R. G. Morris, D. A. Rapetti, R. W. Schmidt, S. Adhikari, M. T. Allen, P. R. Burchat, D. L. Burke, M. Cataneo, D. Donovan, H. Ebeling, S. Shandera, and A. Wright. Weighing the giants - IV. Cosmology and neutrino mass. *MNRAS*, 446:2205–2225, January 2015. doi:[10.1093/mnras/stu2096](https://doi.org/10.1093/mnras/stu2096).
- T.-X. Mao, J. Wang, C. S. Frenk, L. Gao, R. Li, and Q. Wang. Resolution of the apparent discrepancy between the number of massive subhaloes in Abell 2744 and Λ CDM. *ArXiv e-prints*, August 2017.
- M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, W. Forman, C. Jones, S. Murray, and W. Tucker. Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657-56. *ApJ*, 606:819–824, May 2004. doi:[10.1086/383178](https://doi.org/10.1086/383178).
- R. Massey, B. Rowe, A. Refregier, D. J. Bacon, and J. Bergé. Weak gravitational shear and flexion with polar shapelets. *MNRAS*, 380:229–245, September 2007. doi:[10.1111/j.1365-2966.2007.12072.x](https://doi.org/10.1111/j.1365-2966.2007.12072.x). URL <http://adsabs.harvard.edu/abs/2007MNRAS.380..229M>.
- J. C. Mather, E. S. Cheng, R. E. Eplee, Jr., R. B. Isaacman, S. S. Meyer, R. A. Shafer, R. Weiss, E. L. Wright, C. L. Bennett, N. W. Boggess, E. Dwek, S. Gulkis, M. G. Hauser, M. Janssen, T. Kelsall, P. M. Lubin, S. H. Moseley, Jr., T. L. Murdock, R. F. Silverberg, G. F. Smoot, and D. T. Wilkinson. A preliminary measurement of the cosmic microwave background spectrum by the Cosmic Background Explorer (COBE) satellite. *ApJ*, 354:L37–L40, May 1990. doi:[10.1086/185717](https://doi.org/10.1086/185717).
- J. C. Mather, E. S. Cheng, D. A. Cottingham, R. E. Eplee, Jr., D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jensen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, S. H. Moseley, Jr., R. F. Silverberg, G. F. Smoot, R. Weiss, and D. T. Wilkinson. Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument. *ApJ*, 420:439–444, January 1994. doi:[10.1086/173574](https://doi.org/10.1086/173574).
- C. McCully, C. R. Keeton, K. C. Wong, and A. I. Zabludoff. Quantifying Environmental and Line-of-sight Effects in Models of Strong Gravitational Lens Systems. *ApJ*, 836:141, February 2017. doi:[10.3847/1538-4357/836/1/141](https://doi.org/10.3847/1538-4357/836/1/141).

- C. McPartland, H. Ebeling, E. Roediger, and K. Blumenthal. Jellyfish: the origin and distribution of extreme ram-pressure stripping events in massive galaxy clusters. *MNRAS*, 455:2994–3008, January 2016. doi:[10.1093/mnras/stv2508](https://doi.org/10.1093/mnras/stv2508).
- M. Meneghetti, M. Bartelmann, H. Dahle, and M. Limousin. Arc Statistics. *Space Sci. Rev.*, 177:31–74, August 2013. doi:[10.1007/s11214-013-9981-x](https://doi.org/10.1007/s11214-013-9981-x).
- M. Meneghetti, P. Natarajan, D. Coe, E. Contini, G. De Lucia, C. Giocoli, A. Acebron, S. Borgani, M. Bradac, J. M. Diego, A. Hoag, M. Ishigaki, T. L. Johnson, E. Jullo, R. Kawamata, D. Lam, M. Limousin, J. Liesenborgs, M. Oguri, K. Sebesta, K. Sharon, L. L. R. Williams, and A. Zitrin. The Frontier Fields Lens Modeling Comparison Project. *ArXiv e-prints*, June 2016.
- J. Merten, M. Meneghetti, M. Postman, K. Umetsu, A. Zitrin, E. Medezinski, M. Nonino, A. Koekemoer, P. Melchior, D. Gruen, L. A. Moustakas, M. Bartelmann, O. Host, M. Donahue, D. Coe, A. Molino, S. Jouvel, A. Monna, S. Seitz, N. Czakon, D. Lemze, J. Sayers, I. Balestra, P. Rosati, N. Benítez, A. Biviano, R. Bouwens, L. Bradley, T. Broadhurst, M. Carrasco, H. Ford, C. Grillo, L. Infante, D. Kelson, O. Lahav, R. Massey, J. Moustakas, E. Rasia, J. Rhodes, J. Vega, and W. Zheng. CLASH: The Concentration-Mass Relation of Galaxy Clusters. *ApJ*, 806:4, June 2015. doi:[10.1088/0004-637X/806/1/4](https://doi.org/10.1088/0004-637X/806/1/4).
- E. A. Milne. World-Structure and the Expansion of the Universe. Mit 6 Abbildungen. *Z. Astrophys.*, 6:1, 1933.
- C. W. Misner. Mixmaster Universe. *Physical Review Letters*, 22:1071–1074, May 1969. doi:[10.1103/PhysRevLett.22.1071](https://doi.org/10.1103/PhysRevLett.22.1071).
- A. Molino, N. Benítez, B. Ascaso, D. Coe, M. Postman, S. Jouvel, O. Host, O. Lahav, S. Seitz, E. Medezinski, P. Rosati, W. Schoenell, A. Koekemoer, Y. Jimenez-Teja, T. Broadhurst, P. Melchior, I. Balestra, M. Bartelmann, R. Bouwens, L. Bradley, N. Czakon, M. Donahue, H. Ford, O. Graur, G. Graves, C. Grillo, L. Infante, S. W. Jha, D. Kelson, R. Lazkoz, D. Lemze, D. Maoz, A. Mercurio, M. Meneghetti, J. Merten, L. Moustakas, M. Nonino, S. Orgaz, A. Riess, S. Rodney, J. Sayers, K. Umetsu, W. Zheng, and A. Zitrin. CLASH: accurate photometric redshifts with 14 HST bands in massive galaxy cluster cores. *MNRAS*, 470:95–113, September 2017. doi:[10.1093/mnras/stx1243](https://doi.org/10.1093/mnras/stx1243).
- A. Monna, S. Seitz, I. Balestra, P. Rosati, C. Grillo, A. Halkola, S. H. Suyu, D. Coe, G. B. Caminha, B. Frye, A. Koekemoer, A. Mercurio, M. Nonino, M. Postman, and A. Zitrin. Precise strong lensing mass profile of the CLASH galaxy cluster MACS 2129. *MNRAS*, 466:4094–4106, April 2017. doi:[10.1093/mnras/stx015](https://doi.org/10.1093/mnras/stx015).
- V. F. Mukhanov and G. V. Chibisov. Quantum fluctuations and a nonsingular universe. *Soviet Journal of Experimental and Theoretical Physics Letters*, 33:532, May 1981.
- R. Narayan, R. Blandford, and R. Nityananda. Multiple imaging of quasars by galaxies and clusters. *Nature*, 310:112–115, July 1984. doi:[10.1038/310112a0](https://doi.org/10.1038/310112a0).

Bibliography

- P. Natarajan, J.-P. Kneib, I. Smail, and R. S. Ellis. The Mass-to-Light Ratio of Early-Type Galaxies: Constraints from Gravitational Lensing in the Rich Cluster AC 114. *ApJ*, 499: 600–607, May 1998.
- P. Natarajan, U. Chadayammuri, M. Jauzac, J. Richard, J.-P. Kneib, H. Ebeling, F. Jiang, F. van den Bosch, M. Limousin, E. Jullo, H. Atek, A. Pillepich, C. Popa, F. Marinacci, L. Hernquist, M. Meneghetti, and M. Vogelsberger. Mapping substructure in the HST Frontier Fields cluster lenses and in cosmological simulations. *MNRAS*, 468:1962–1980, June 2017. doi:[10.1093/mnras/stw3385](https://doi.org/10.1093/mnras/stw3385).
- J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. *ApJ*, 462:563, May 1996. doi:[10.1086/177173](https://doi.org/10.1086/177173). URL <http://adsabs.harvard.edu/abs/1996ApJ...462..563N>.
- J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *ApJ*, 490:493, December 1997. doi:[10.1086/304888](https://doi.org/10.1086/304888). URL <http://adsabs.harvard.edu/abs/1997ApJ...490..493N>.
- Nvidia Corporation. *Kepler GK110/210 Whitepaper* [Online; accessed October 2, 2017], 2014. URL <http://international.download.nvidia.com/pdf/kepler/NVIDIA-Kepler-GK110-GK210-Architecture-Whitepaper.pdf>.
- Nvidia Corporation. *Tesla K80 GPU Accelerator Board Specification* [Online; accessed September 29, 2017], January 2015. URL <http://images.nvidia.com/content/pdf/kepler/Tesla-K80-BoardSpec-07317-001-v05.pdf>.
- Nvidia Corporation. *Nvidia Tesla P100 Whitepaper* [Online; accessed October 2, 2017], 2016. URL <https://images.nvidia.com/content/pdf/tesla/whitepaper/pascal-architecture-whitepaper.pdf>.
- Nvidia Corporation. *Nvidia GeForce GTX 1080 Ti Specifications* [Online; accessed October 2, 2017], 2017. URL <https://www.nvidia.com/en-us/geforce/products/10series/geforce-gtx-1080-ti/>.
- M. Oguri. The Mass Distribution of SDSS J1004+4112 Revisited. *PASJ*, 62:1017–1024, August 2010. doi:[10.1093/pasj/62.4.1017](https://doi.org/10.1093/pasj/62.4.1017).
- M. Oguri. Predicted properties of multiple images of the strongly lensed supernova SN Refsdal. *MNRAS*, 449:L86–L89, April 2015. doi:[10.1093/mnrasl/slv025](https://doi.org/10.1093/mnrasl/slv025).
- Y. Okura, K. Umetsu, and T. Futamase. A New Measure for Weak-Lensing Flexion. *ApJ*, 660: 995–1002, May 2007. doi:[10.1086/513135](https://doi.org/10.1086/513135). URL <http://adsabs.harvard.edu/abs/2007ApJ...660..995O>.
- Y. Okura, K. Umetsu, and T. Futamase. A Method for Weak-Lensing Flexion Analysis by the HOLICs Moment Approach. *ApJ*, 680:1–16, June 2008. doi:[10.1086/587676](https://doi.org/10.1086/587676). URL <http://adsabs.harvard.edu/abs/2008ApJ...680....1O>.

- C. O’Raifeartaigh. The Contribution of V. M. Slipher to the Discovery of the Expanding Universe. In M. J. Way and D. Hunter, editors, *Origins of the Expanding Universe: 1912-1932*, volume 471 of *Astronomical Society of the Pacific Conference Series*, page 49, April 2013.
- C. O’Raifeartaigh, M. O’Keeffe, W. Nahm, and S. Mitton. Einstein’s cosmology review of 1933: a new perspective on the Einstein-de Sitter model of the cosmos. *European Physical Journal H*, 40, September 2015. doi:[10.1140/epjh/e2015-50061-y](https://doi.org/10.1140/epjh/e2015-50061-y).
- C. O’Raifeartaigh, M. O’Keeffe, W. Nahm, and S. Mitton. Einstein’s 1917 static model of the universe: a centennial review. *European Physical Journal H*, July 2017. doi:[10.1140/epjh/e2017-80002-5](https://doi.org/10.1140/epjh/e2017-80002-5).
- D. E. Osterbrock and J. B. Rogerson, Jr. The Helium and Heavy-Element Content of Gaseous-Nebulae and the Sun. *PASP*, 73:129, April 1961. doi:[10.1086/127637](https://doi.org/10.1086/127637).
- B. Paczynski. Gravitational microlensing by the galactic halo. *ApJ*, 304:1–5, May 1986. doi:[10.1086/164140](https://doi.org/10.1086/164140). URL <http://adsabs.harvard.edu/abs/1986ApJ...304....1P>.
- B. Paczynski. Giant luminous arcs discovered in two clusters of galaxies. *Nature*, 325:572–573, February 1987. doi:[10.1038/325572a0](https://doi.org/10.1038/325572a0).
- D. Paraficz, J.-P. Kneib, J. Richard, A. Morandi, M. Limousin, E. Jullo, and J. Martinez. The Bullet cluster at its best: weighing stars, gas, and dark matter. *A&A*, 594:A121, October 2016. doi:[10.1051/0004-6361/201527959](https://doi.org/10.1051/0004-6361/201527959).
- J. A. Peacock, S. Cole, P. Norberg, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. D. Cannon, M. Colless, C. Collins, W. Couch, G. Dalton, K. Deeley, R. De Propriis, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, W. J. Percival, B. A. Peterson, I. Price, W. Sutherland, and K. Taylor. A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey. *Nature*, 410:169–173, March 2001.
- P. J. E. Peebles. Primordial Helium Abundance and the Primordial Fireball. II. *ApJ*, 146:542, November 1966. doi:[10.1086/148918](https://doi.org/10.1086/148918).
- A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *ApJ*, 142:419–421, July 1965. doi:[10.1086/148307](https://doi.org/10.1086/148307).
- S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and T. S. C. Project. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *ApJ*, 517:565–586, June 1999. doi:[10.1086/307221](https://doi.org/10.1086/307221).

Bibliography

- L. Piqueras, S. Conseil, M. Shepherd, R. Bacon, F. Leclercq, and J. Richard. MPDAF - A Python package for the analysis of VLT/MUSE data. *ArXiv e-prints*, October 2017.
- S. Pires and A. Amara. Weak Lensing Mass Reconstruction: Flexion Versus Shear. *ApJ*, 723: 1507–1511, November 2010. doi:[10.1088/0004-637X/723/2/1507](https://doi.org/10.1088/0004-637X/723/2/1507).
- Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, September 2016. doi:[10.1051/0004-6361/201525830](https://doi.org/10.1051/0004-6361/201525830).
- B. M. Poggianti, Y. L. Jaffé, A. Moretti, M. Gullieuszik, M. Radovich, S. Tonnesen, J. Fritz, D. Bettoni, B. Vulcani, G. Fasano, C. Bellhouse, G. Hau, and A. Omizzolo. Ram-pressure feeding of supermassive black holes. *Nature*, 548:304–309, August 2017a. doi:[10.1038/nature23462](https://doi.org/10.1038/nature23462).
- B. M. Poggianti, A. Moretti, M. Gullieuszik, J. Fritz, Y. Jaffé, D. Bettoni, G. Fasano, C. Bellhouse, G. Hau, B. Vulcani, A. Biviano, A. Omizzolo, A. Paccagnella, M. D’Onofrio, A. Cava, Y.-K. Sheen, W. Couch, and M. Owers. GASP. I. Gas Stripping Phenomena in Galaxies with MUSE. *ApJ*, 844:48, July 2017b. doi:[10.3847/1538-4357/aa78ed](https://doi.org/10.3847/1538-4357/aa78ed).
- S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradač. Constraints on the Self-Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56. *ApJ*, 679:1173–1180, June 2008. doi:[10.1086/587859](https://doi.org/10.1086/587859).
- M. J. Rees and J. P. Ostriker. Cooling, dynamics and fragmentation of massive gas clouds - Clues to the masses and radii of galaxies and clusters. *MNRAS*, 179:541–559, June 1977. URL <http://adsabs.harvard.edu/abs/1977MNRAS.179..541R>.
- A. Refregier. Shapelets - I. A method for image analysis. *MNRAS*, 338:35–47, January 2003. doi:[10.1046/j.1365-8711.2003.05901.x](https://doi.org/10.1046/j.1365-8711.2003.05901.x). URL <http://adsabs.harvard.edu/abs/2003MNRAS.338...35R>.
- A. Refregier and A. Amara. A way forward for Cosmic Shear: Monte-Carlo Control Loops. *Physics of the Dark Universe*, 3:1–3, April 2014. doi:[10.1016/j.dark.2014.01.002](https://doi.org/10.1016/j.dark.2014.01.002).
- A. Refregier and D. Bacon. Shapelets - II. A method for weak lensing measurements. *MNRAS*, 338:48–56, January 2003. doi:[10.1046/j.1365-8711.2003.05902.x](https://doi.org/10.1046/j.1365-8711.2003.05902.x). URL <http://adsabs.harvard.edu/abs/2003MNRAS.338...48R>.
- S. Refsdal. The gravitational lens effect. *MNRAS*, 128:295, 1964a. doi:[10.1093/mnras/128.4.295](https://doi.org/10.1093/mnras/128.4.295).
- S. Refsdal. On the possibility of determining Hubble’s parameter and the masses of galaxies from the gravitational lens effect. *MNRAS*, 128:307, 1964b. doi:[10.1093/mnras/128.4.307](https://doi.org/10.1093/mnras/128.4.307).

- M. Rexroth. Substructure in the Frontier Fields from weak lensing flexion. *Proceedings of the International Astronomical Union*, 11(A29B):795, Aug 2015. doi:[10.1017/S1743921316006864](https://doi.org/10.1017/S1743921316006864).
- M. Rexroth, P. Natarajan, and J.-P. Kneib. A new method to break the mass-sheet degeneracy using aperture moments. *MNRAS*, 460:2505–2525, August 2016. doi:[10.1093/mnras/stw1017](https://doi.org/10.1093/mnras/stw1017).
- M. Rexroth, J.-P. Kneib, R. Joseph, J. Richard, and R. Her. IFS-RedEx, a redshift extraction software for integral-field spectrographs: Application to MUSE data. *ArXiv e-prints*, March 2017.
- J. Richard, L. Pei, M. Limousin, E. Jullo, and J. P. Kneib. Keck spectroscopic survey of strongly lensed galaxies in Abell 1703: further evidence of a relaxed, unimodal cluster. *A&A*, 498:37–47, April 2009. doi:[10.1051/0004-6361/200811366](https://doi.org/10.1051/0004-6361/200811366).
- J. Richard, G. P. Smith, J.-P. Kneib, R. S. Ellis, A. J. R. Sanderson, L. Pei, T. A. Targett, D. J. Sand, A. M. Swinbank, H. Dannerbauer, P. Mazzotta, M. Limousin, E. Egami, E. Jullo, V. Hamilton-Morris, and S. M. Moran. LoCuSS: first results from strong-lensing analysis of 20 massive galaxy clusters at $z = 0.2$. *MNRAS*, 404:325–349, May 2010. doi:[10.1111/j.1365-2966.2009.16274.x](https://doi.org/10.1111/j.1365-2966.2009.16274.x).
- J. Richard, V. Patricio, J. Martinez, R. Bacon, B. Clément, P. Weilbacher, K. Soto, L. Wisotzki, J. Vernet, R. Pello, J. Schaye, M. Turner, and T. Martinsson. MUSE observations of the lensing cluster SMACSJ2031.8-4036: new constraints on the mass distribution in the cluster core. *MNRAS*, 446:L16–L20, January 2015. doi:[10.1093/mnras/flu150](https://doi.org/10.1093/mnras/flu150).
- A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ*, 116:1009–1038, September 1998. doi:[10.1086/300499](https://doi.org/10.1086/300499).
- A. G. Riess, L. M. Macri, S. L. Hoffmann, D. Scolnic, S. Casertano, A. V. Filippenko, B. E. Tucker, M. J. Reid, D. O. Jones, J. M. Silverman, R. Chornock, P. Challis, W. Yuan, P. J. Brown, and R. J. Foley. A 2.4% Determination of the Local Value of the Hubble Constant. *ApJ*, 826:56, July 2016. doi:[10.3847/0004-637X/826/1/56](https://doi.org/10.3847/0004-637X/826/1/56).
- A. Riotto and M. Trodden. Recent Progress in Baryogenesis. *Annual Review of Nuclear and Particle Science*, 49:35–75, 1999. doi:[10.1146/annurev.nucl.49.1.35](https://doi.org/10.1146/annurev.nucl.49.1.35).
- H. P. Robertson. Kinematics and World-Structure. *ApJ*, 82:284, November 1935. doi:[10.1086/143681](https://doi.org/10.1086/143681).
- P. Rosati, I. Balestra, C. Grillo, A. Mercurio, M. Nonino, A. Biviano, M. Girardi, E. Vanzella, and Clash-VLT Team. CLASH-VLT: A VIMOS Large Programme to Map the Dark Matter

Bibliography

- Mass Distribution in Galaxy Clusters and Probe Distant Lensed Galaxies. *The Messenger*, 158:48–53, December 2014.
- B. Rowe, D. Bacon, R. Massey, C. Heymans, B. Häußler, A. Taylor, J. Rhodes, and Y. Mellier. Flexion measurement in simulations of Hubble Space Telescope data. *MNRAS*, 435:822–844, October 2013. doi:[10.1093/mnras/stt1353](https://doi.org/10.1093/mnras/stt1353). URL <http://adsabs.harvard.edu/abs/2013MNRAS.435..822R>.
- B. T. P. Rowe, M. Jarvis, R. Mandelbaum, G. M. Bernstein, J. Bosch, M. Simet, J. E. Meyers, T. Kacprzak, R. Nakajima, J. Zuntz, H. Miyatake, J. P. Dietrich, R. Armstrong, P. Melchior, and M. S. S. Gill. GALSIM: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150, April 2015. doi:[10.1016/j.ascom.2015.02.002](https://doi.org/10.1016/j.ascom.2015.02.002).
- V. C. Rubin and W. K. Ford, Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *ApJ*, 159:379, February 1970. doi:[10.1086/150317](https://doi.org/10.1086/150317).
- V. C. Rubin, N. Thonnard, and W. K. Ford, Jr. Extended rotation curves of high-luminosity spiral galaxies. IV - Systematic dynamical properties, SA through SC. *ApJ*, 225:L107–L111, November 1978. doi:[10.1086/182804](https://doi.org/10.1086/182804).
- V. C. Rubin, W. K. J. Ford, and N. . Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *ApJ*, 238:471–487, June 1980. doi:[10.1086/158003](https://doi.org/10.1086/158003).
- A. D. Sakharov. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Soviet Journal of Experimental and Theoretical Physics Letters*, 5:24, January 1967.
- A. Sandage. Current Problems in the Extragalactic Distance Scale. *ApJ*, 127:513, May 1958. doi:[10.1086/146483](https://doi.org/10.1086/146483).
- C. Schäfer. Lenstool-GPU. Master's thesis, EPFL, 2016.
- M. Schmidt. 3C 273 : A Star-Like Object with Large Red-Shift. *Nature*, 197:1040, March 1963. doi:[10.1038/1971040a0](https://doi.org/10.1038/1971040a0).
- P. Schneider and X. Er. Weak lensing goes bananas: what flexion really measures. *A&A*, 485: 363–376, July 2008. doi:[10.1051/0004-6361:20078631](https://doi.org/10.1051/0004-6361:20078631). URL <http://adsabs.harvard.edu/abs/2008A%26A...485..363S>.
- P. Schneider and D. Sluse. Mass-sheet degeneracy, power-law models and external convergence: Impact on the determination of the Hubble constant from gravitational lensing. *A&A*, 559: A37, November 2013. doi:[10.1051/0004-6361/201321882](https://doi.org/10.1051/0004-6361/201321882).
- P. Schneider and D. Sluse. Source-position transformation: an approximate invariance in strong gravitational lensing. *A&A*, 564:A103, April 2014. doi:[10.1051/0004-6361/201322106](https://doi.org/10.1051/0004-6361/201322106).

- P. Schneider, L. King, and T. Erben. Cluster mass profiles from weak lensing: constraints from shear and magnification information. *A&A*, 353:41–56, January 2000. URL <http://adsabs.harvard.edu/abs/2000A%26A...353...41S>.
- P. Schneider, C. S. Kochanek, and J. Wambsganss. *Gravitational Lensing: Strong, Weak and Micro. Saas-Fee Advanced Course 33*. Springer Berlin Heidelberg, 2006. ISBN 978-3-540-30309-1. doi:[10.1007/978-3-540-30310-7](https://doi.org/10.1007/978-3-540-30310-7).
- J. Schwinn, M. Jauzac, C. M. Baugh, M. Bartelmann, D. Eckert, D. Harvey, P. Natarajan, and R. Massey. Abell 2744: too much substructure for Λ CDM? *MNRAS*, 467:2913–2923, May 2017. doi:[10.1093/mnras/stx277](https://doi.org/10.1093/mnras/stx277).
- H. Shapley. Studies based on the colors and magnitudes in stellar clusters. VI. On the determination of the distances of globular clusters. *ApJ*, 48, September 1918. doi:[10.1086/142419](https://doi.org/10.1086/142419).
- K. Sharon and T. L. Johnson. Revised Lens Model for the Multiply Imaged Lensed Supernova, “SN Refsdal” in MACS J1149+2223. *ApJ*, 800:L26, February 2015. doi:[10.1088/2041-8205/800/2/L26](https://doi.org/10.1088/2041-8205/800/2/L26).
- Y.-K. Sheen, R. Smith, Y. Jaffé, M. Kim, S. K. Yi, P.-A. Duc, J. Nantais, G. Candlish, R. Demarco, and E. Treister. Discovery of Ram-pressure Stripped Gas around an Elliptical Galaxy in Abell 2670. *ApJ*, 840:L7, May 2017. doi:[10.3847/2041-8213/aa6d79](https://doi.org/10.3847/2041-8213/aa6d79).
- V. M. Slipher. The radial velocity of the Andromeda Nebula. *Lowell Observatory Bulletin*, 2: 56–57, 1913.
- V. M. Slipher. Nebulae. *Proceedings of the American Philosophical Society*, 56:403–409, 1917.
- G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. *ApJ*, 396:L1–L5, September 1992. doi:[10.1086/186504](https://doi.org/10.1086/186504).
- K. T. Soto, S. J. Lilly, R. Bacon, J. Richard, and S. Conseil. ZAP - enhanced PCA sky subtraction for integral field spectroscopy. *MNRAS*, 458:3210–3220, May 2016. doi:[10.1093/mnras/stw474](https://doi.org/10.1093/mnras/stw474).
- G. Soucail, B. Fort, Y. Mellier, and J. P. Picat. A blue ring-like structure, in the center of the A 370 cluster of galaxies. *A&A*, 172:L14–L16, January 1987.
- D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R.olta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *ApJS*, 148:175–194, September 2003. doi:[10.1086/377226](https://doi.org/10.1086/377226).

Bibliography

- D. P. Stark. Galaxies in the First Billion Years After the Big Bang. *ARA&A*, 54:761–803, September 2016. doi:[10.1146/annurev-astro-081915-023417](https://doi.org/10.1146/annurev-astro-081915-023417).
- A. A. Starobinskiĭ. Spectrum of relict gravitational radiation and the early state of the universe. *Soviet Journal of Experimental and Theoretical Physics Letters*, 30:682, December 1979.
- A. A. Starobinsky. A new type of isotropic cosmological models without singularity. *Physics Letters B*, 91:99–102, March 1980. doi:[10.1016/0370-2693\(80\)90670-X](https://doi.org/10.1016/0370-2693(80)90670-X).
- G. Steigman. Primordial Nucleosynthesis in the Precision Cosmology Era. *Annual Review of Nuclear and Particle Science*, 57:463–491, November 2007. doi:[10.1146/annurev.nucl.56.080805.140437](https://doi.org/10.1146/annurev.nucl.56.080805.140437).
- G. Steigman, D. N. Schramm, and J. E. Gunn. Cosmological limits to the number of massive leptons. *Physics Letters B*, 66:202–204, January 1977. doi:[10.1016/0370-2693\(77\)90176-9](https://doi.org/10.1016/0370-2693(77)90176-9).
- A. N. Taylor, S. Dye, T. J. Broadhurst, N. Benitez, and E. van Kampen. Gravitational Lens Magnification and the Mass of Abell 1689. *ApJ*, 501:539, July 1998. doi:[10.1086/305827](https://doi.org/10.1086/305827). URL <http://adsabs.harvard.edu/abs/1998ApJ...501..539T>.
- P. Tisserand, L. Le Guillou, C. Afonso, J. N. Albert, J. Andersen, R. Ansari, É. Aubourg, P. Bareyre, J. P. Beaulieu, X. Charlot, C. Coutures, R. Ferlet, P. Fouqué, J. F. Glicenstein, B. Goldman, A. Gould, D. Graff, M. Gros, J. Haissinski, C. Hamadache, J. de Kat, T. Lasserre, É. Lesquoy, C. Loup, C. Magneville, J. B. Marquette, É. Maurice, A. Maury, A. Milsztajn, M. Moniez, N. Palanque-Delabrouille, O. Perdureau, Y. R. Rahal, J. Rich, M. Spiro, A. Vidal-Madjar, L. Vigroux, S. Zylberajch, and EROS-2 Collaboration. Limits on the Macho content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. *A&A*, 469:387–404, July 2007. doi:[10.1051/0004-6361:20066017](https://doi.org/10.1051/0004-6361:20066017).
- E. P. Tryon. Is the Universe a Vacuum Fluctuation? *Nature*, 246:396–397, December 1973. doi:[10.1038/246396a0](https://doi.org/10.1038/246396a0).
- J. A. Tyson, F. Valdes, and R. A. Wenk. Detection of systematic gravitational lens galaxy image alignments - Mapping dark matter in galaxy clusters. *ApJ*, 349:L1–L4, January 1990. doi:[10.1086/185636](https://doi.org/10.1086/185636).
- K. Umetsu, T. Broadhurst, A. Zitrin, E. Medezinski, and L.-Y. Hsu. Cluster Mass Profiles from a Bayesian Analysis of Weak-lensing Distortion and Magnification Measurements: Applications to Subaru Data. *ApJ*, 729:127, March 2011. doi:[10.1088/0004-637X/729/2/127](https://doi.org/10.1088/0004-637X/729/2/127).
- K. Umetsu, E. Medezinski, M. Nonino, J. Merten, M. Postman, M. Meneghetti, M. Donahue, N. Czakon, A. Molino, S. Seitz, D. Gruen, D. Lemze, I. Balestra, N. Benítez, A. Biviano, T. Broadhurst, H. Ford, C. Grillo, A. Koekemoer, P. Melchior, A. Mercurio, J. Moustakas, P. Rosati, and A. Zitrin. CLASH: Weak-lensing Shear-and-magnification Analysis of 20 Galaxy Clusters. *ApJ*, 795:163, November 2014. doi:[10.1088/0004-637X/795/2/163](https://doi.org/10.1088/0004-637X/795/2/163).

- K. Umetsu, A. Zitrin, D. Gruen, J. Merten, M. Donahue, and M. Postman. CLASH: Joint Analysis of Strong-lensing, Weak-lensing Shear, and Magnification Data for 20 Galaxy Clusters. *ApJ*, 821:116, April 2016. doi:[10.3847/0004-637X/821/2/116](https://doi.org/10.3847/0004-637X/821/2/116).
- L. Van Waerbeke, Y. Mellier, T. Erben, J. C. Cuillandre, F. Bernardeau, R. Maoli, E. Bertin, H. J. McCracken, O. Le Fèvre, B. Fort, M. Dantel-Fort, B. Jain, and P. Schneider. Detection of correlated galaxy ellipticities from CFHT data: first evidence for gravitational lensing by large-scale structures. *A&A*, 358:30–44, June 2000. URL <http://adsabs.harvard.edu/abs/2000A%26A...358...30V>.
- M. Velander, K. Kuijken, and T. Schrabback. Probing galaxy dark matter haloes in COSMOS with weak lensing flexion. *MNRAS*, 412:2665–2677, April 2011. doi:[10.1111/j.1365-2966.2010.18085.x](https://doi.org/10.1111/j.1365-2966.2010.18085.x). URL <http://adsabs.harvard.edu/abs/2011MNRAS.412.2665V>.
- M. Viola, P. Melchior, and M. Bartelmann. Shear-flexion cross-talk in weak-lensing measurements. *MNRAS*, 419:2215–2225, January 2012. doi:[10.1111/j.1365-2966.2011.19872.x](https://doi.org/10.1111/j.1365-2966.2011.19872.x). URL <http://adsabs.harvard.edu/abs/2012MNRAS.419.2215V>.
- M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *MNRAS*, 444:1518–1547, October 2014. doi:[10.1093/mnras/stu1536](https://doi.org/10.1093/mnras/stu1536).
- C. Vuissoz, F. Courbin, D. Sluse, G. Meylan, V. Chantry, E. Eulaers, C. Morgan, M. E. Eyler, C. S. Kochanek, J. Coles, P. Saha, P. Magain, and E. E. Falco. COSMOGRAIL: the COSmological MONitoring of GRAVitational Lenses. VII. Time delays and the Hubble constant from WFI J2033-4723. *A&A*, 488:481–490, September 2008. doi:[10.1051/0004-6361:200809866](https://doi.org/10.1051/0004-6361:200809866).
- R. V. Wagoner, W. A. Fowler, and F. Hoyle. On the Synthesis of Elements at Very High Temperatures. *ApJ*, 148:3, April 1967. doi:[10.1086/149126](https://doi.org/10.1086/149126).
- A. G. Walker. On Riemannian spaces with spherical symmetry about a line, and the conditions for isotropy in general relativity. *The Quarterly Journal of Mathematics*, 6:81–93, 1935. doi:[10.1093/qmath/os-6.1.81](https://doi.org/10.1093/qmath/os-6.1.81).
- D. Walsh, R. F. Carswell, and R. J. Weymann. 0957 + 561 A, B - Twin quasistellar objects or gravitational lens. *Nature*, 279:381–384, May 1979. doi:[10.1038/279381a0](https://doi.org/10.1038/279381a0).
- P. M. Weilbacher, O. Streicher, T. Urrutia, A. Jarno, A. Pécontal-Rousset, R. Bacon, and P. Böhm. Design and capabilities of the MUSE data reduction software and pipeline. In *Software and Cyberinfrastructure for Astronomy II*, volume 8451 of *Proc. SPIE*, page 84510B, September 2012. doi:[10.1117/12.925114](https://doi.org/10.1117/12.925114).
- P. M. Weilbacher, O. Streicher, T. Urrutia, A. Pécontal-Rousset, A. Jarno, and R. Bacon. The MUSE Data Reduction Pipeline: Status after Preliminary Acceptance Europe. In N. Manset

Bibliography

- and P. Forshay, editors, *Astronomical Data Analysis Software and Systems XXIII*, volume 485 of *Astronomical Society of the Pacific Conference Series*, page 451, May 2014.
- C. F. v. Weizsäcker. Über Elementumwandlungen im Innern der Sterne, I. *Physikalische Zeitschrift*, 38:176–191, 1937.
- C. F. v. Weizsäcker. Über Elementumwandlungen im Innern der Sterne, II. *Physikalische Zeitschrift*, 39:633–646, 1938.
- S. D. M. White, M. Davis, G. Efstathiou, and C. S. Frenk. Galaxy distribution in a cold dark matter universe. *Nature*, 330:451–453, December 1987. doi:[10.1038/330451a0](https://doi.org/10.1038/330451a0).
- D. Wittman, N. Golovich, and W. A. Dawson. The Mismeasure of Mergers: Revised Limits on Self-interacting Dark Matter in Merging Galaxy Clusters. *ArXiv e-prints*, January 2017.
- D. M. Wittman, J. A. Tyson, D. Kirkman, I. Dell’Antonio, and G. Bernstein. Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. *Nature*, 405:143–148, May 2000. doi:[10.1038/35012001](https://doi.org/10.1038/35012001).
- A. Zitrin, T. Broadhurst, K. Umetsu, D. Coe, N. Benítez, B. Ascaso, L. Bradley, H. Ford, J. Jee, E. Medezinski, Y. Rephaeli, and W. Zheng. New multiply-lensed galaxies identified in ACS/NIC3 observations of Cl0024+1654 using an improved mass model. *MNRAS*, 396:1985–2002, July 2009. doi:[10.1111/j.1365-2966.2009.14899.x](https://doi.org/10.1111/j.1365-2966.2009.14899.x).
- A. Zitrin, A. Fabris, J. Merten, P. Melchior, M. Meneghetti, A. Koekemoer, D. Coe, M. Maturi, M. Bartelmann, M. Postman, K. Umetsu, G. Seidel, I. Sendra, T. Broadhurst, I. Balestra, A. Biviano, C. Grillo, A. Mercurio, M. Nonino, P. Rosati, L. Bradley, M. Carrasco, M. Donahue, H. Ford, B. L. Frye, and J. Moustakas. Hubble Space Telescope Combined Strong and Weak Lensing Analysis of the CLASH Sample: Mass and Magnification Models and Systematic Uncertainties. *ApJ*, 801:44, March 2015. doi:[10.1088/0004-637X/801/1/44](https://doi.org/10.1088/0004-637X/801/1/44).
- F. Zwicky. On the Red Shift of Spectral Lines through Interstellar Space. *Proceedings of the National Academy of Science*, 15:773–779, October 1929. doi:[10.1073/pnas.15.10.773](https://doi.org/10.1073/pnas.15.10.773).
- F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.
- F. Zwicky. Nebulae as Gravitational Lenses. *Physical Review*, 51:290–290, February 1937a. doi:[10.1103/PhysRev.51.290](https://doi.org/10.1103/PhysRev.51.290).
- F. Zwicky. On the Probability of Detecting Nebulae Which Act as Gravitational Lenses. *Physical Review*, 51:679–679, April 1937b. doi:[10.1103/PhysRev.51.679](https://doi.org/10.1103/PhysRev.51.679).

Acronyms and abbreviations

Λ CDM Λ Cold Dark Matter

2dF Two degree Field Galaxy Redshift Survey

AoS Arrays of Structures

AVX Advanced Vector Extensions

BAO Baryon Acoustic Oscillations

BBT Big Bang Theory

BCG Brightest Cluster Galaxy

BOOMERanG Balloon Observations Of Millimetric Extragalactic Radiation ANd Geophysics

BOSS Baryon Oscillation Spectroscopic Survey

CDM Cold Dark Matter

CFHTLenS Canada-France-Hawaii Telescope Lensing Survey

CLASH Cluster Lensing And Supernova survey with Hubble

CMB Cosmic Microwave Background

COBE Cosmic Background Explorer

CPU Central Processing Unit

DE Dark Energy

DES Dark Energy Survey

DM Dark Matter

DP Double Precision

Acronyms and abbreviations

dPIE dual Pseudo Isothermal Elliptical mass distribution

E-ELT European Extremely Large Telescope

FLRW metric Friedmann-Lemaître-Robertson-Walker metric

FWHM Full Width at Half Maximum

GPU Graphics Processing Unit

GR General Theory of Relativity

HDM Hot Dark Matter

HFF Hubble Frontier Fields

HOLICS Higher Order Lensing Image's CharacteristicS

HPC High Performance Computing

HST Hubble Space Telescope

ICM Intracluster Medium

IFS Integral-Field Spectrograph

IFS-RedEx Integral-Field Spectrograph Redshift Extractor

IGM Intergalactic Medium

JWST James Webb Space Telescope

KIDS Kilo-Degree Survey

LIGO Laser Interferometer Gravitational-Wave Observatory

LISA Laser Interferometer Space Antenna

LSS Large-Scale Structure

LSST Large Synoptic Survey Telescope

MACHO Massive Astrophysical Compact Halo Object

MACS MAssive Cluster Survey

MAXIMA Millimeter Anisotropy eXperiment IMaging Array

MCMC Markov Chain Monte Carlo

MPDAF MUSE Python Data Analysis Framework

MUSE Multi Unit Spectroscopic Explorer

OpenMP Open Multi-Processing

PSF Point Spread Function

Quasar Quasi-stellar radio source

RAM Random Access Memory

SDSS Sloan Digital Sky Survey

SIE Singular Isothermal pseudo Elliptical sphere

SIMD Single Instruction Multiple Data

SIMT Single Instruction Multiple Thread

SIS Singular Isothermal Sphere

SL Strong Gravitational Lensing

SMBH Supermassive Black Hole

SoA Structures of Arrays

SP Single Precision

VLT Very Large Telescope

WDM Warm Dark Matter

WFIRST Wide Field Infrared Survey Telescope

WIMP Weakly Interacting Massive Particle

WL Weak Gravitational Lensing

WMAP Wilkinson Microwave Anisotropy Probe

ZAP Zurich Atmosphere Purge

Symbols and mathematical notation

θ Angular lens plane coordinate

β Angular source plane coordinate

Σ_{crit} Critical surface mass density

Ω_{Λ} Dark energy density parameter at current epoch

θ_E Einstein deflection angle

G Gravitational constant

κ Gravitational lensing convergence

Ψ Gravitational lensing deflection potential

\mathcal{F} Gravitational lensing first flexion

μ Gravitational lensing magnification

F Gravitational lensing reduced first flexion

G Gravitational lensing reduced second flexion

g Gravitational lensing reduced shear

\mathcal{G} Gravitational lensing second flexion

γ Gravitational lensing shear

H Hubble parameter

H_0 Hubble constant

Ω_m Matter density parameter at current epoch

Ω_R Radiation and relativistic particles density parameter at current epoch

Symbols and mathematical notation

z Redshift

σ_8 RMS matter fluctuations in linear theory at the present epoch

α Scaled gravitational lensing deflection angle

M_\odot Solar mass

Ω_k Spatial curvature density parameter at current epoch

c Speed of light

Ω_{tot} Total density parameter at current epoch

Photographic credits

Cosmic Microwave Background Anisotropies, image credit: ESA / Planck Collaboration, source: <http://sci.esa.int/planck/51738-cosmic-microwave-background-seen-by-planck/>, permission to reproduce in this thesis granted according to ESA Image Copyright Notice (http://www.esa.int/spaceinimages/ESA_Multimedia/Copyright_Notice_Images, retrieved on October 14, 2017). Page 7

Energy density content of the universe pie chart, image credit: NASA / WMAP Science Team, source: <https://map.gsfc.nasa.gov/media/080998/index.html>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on June 15, 2017). Page 9

Evolution of CMB observations, image credit: NASA / WMAP Science Team, source: <https://wmap.gsfc.nasa.gov/media/081031/index.html>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on June 15, 2017). Page 31

Illustration of a baryon, image credit: CERN / D. Dominguez, source: <http://cds.cern.ch/record/2274060>, permission to reproduce in this thesis granted according to CERN audiovisual media usage guidelines (<http://copyright.web.cern.ch/>, retrieved on October 13, 2017). Page 5

Large Scale Structure of the universe zoom-in, image credit: Own work; Large Scale Structure: SDSS / M. Blanton, source: <http://www.sdss.org/science/orangepie/>, permission to reproduce in this thesis granted according to SDSS Image Use Policy (<http://www.sdss.org/collaboration/>, retrieved on October 14, 2017); Galaxy cluster: NASA / ESA / J. Blakeslee / K. Alamo-Martinez / H. Ford, source: <http://hubblesite.org/image/3237/gallery>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on October 14, 2017); Galaxy: NASA / ESA / K. Kuntz / F. Bresolin / J. Trauger / J. Mould / Y.-H. Chu / STScI, source: <http://hubblesite.org/image/2477/gallery>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on October 14, 2017); Planetary system: ESA / C. Carreau, source: <http://>

Photographic credits

[//www.esa.int/spaceinimages/Images/2013/03/Planck_history_of_Universe](http://www.esa.int/spaceinimages/Images/2013/03/Planck_history_of_Universe), permission to reproduce in this thesis granted according to ESA Image Copyright Notice (http://www.esa.int/spaceinimages/ESA_Multimedia/Copyright_Notice_Images, retrieved on October 14, 2017). Page 8

MUSE datacube, image credit: ESO, source: <http://www.eso.org/public/images/eso1518c/>, permission to reproduce in this thesis granted according to ESO Image Usage Guidelines (<http://www.eso.org/public/outreach/copyright/>, retrieved on October 26, 2017). Page 84

Possible spatial geometries in FLRW universe models, image credit: NASA / WMAP Science Team, source: <https://wmap.gsfc.nasa.gov/media/990006/index.html>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on June 15, 2017). Page 16

Sketch of a gravitational lens system, image credit: Own work; Hubble Space Telescope: NASA, source: https://www.nasa.gov/mission_pages/hubble/story/index.html, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on October 14, 2017); Galaxy cluster: NASA / ESA / J. Lotz and the HFF Team / STScI, source: <http://hubblesite.org/image/4024/category/15-galaxy-clusters>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on October 14, 2017); Galaxy: NASA / ESA / K. Kuntz / F. Bresolin / J. Trauger / J. Mould / Y.-H. Chu / STScI, source: <http://hubblesite.org/image/2477/gallery>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on October 14, 2017). Page 38

Spacetime curvature in General Relativity, image credit: ESA / C. Carreau, source: http://www.esa.int/spaceinimages/Images/2015/09/Spacetime_curvature, permission to reproduce in this thesis granted according to ESA Image Copyright Notice (http://www.esa.int/spaceinimages/ESA_Multimedia/Copyright_Notice_Images, retrieved on October 14, 2017). Page 11

Standard model of particle physics, image credit: CERN / D. Dominguez, source: <http://cds.cern.ch/record/2002395>, permission to reproduce in this thesis granted according to CERN audiovisual media usage guidelines (<http://copyright.web.cern.ch/>, retrieved on October 13, 2017). Page 5

Strong lensing galaxy cluster Abell 370, image credit: NASA / ESA / J. Lotz and the HFF Team / STScI, source: <http://hubblesite.org/image/4024/category/15-galaxy-clusters>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines

(<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on October 14, 2017). Page 45

The MUSE instrument on the VLT, image credit: VLT: ESO / José Francisco Salgado (www.josefrancisco.org), source: <http://www.eso.org/public/images/potw1051a/>, permission to reproduce in this thesis granted according to ESO Image Usage Guidelines (<http://www.eso.org/public/outreach/copyright/>, retrieved on October 26, 2017); MUSE instrument: Eric Le Roux / Service Communication / UCBL / MUSE, source: <http://www.eso.org/public/images/ann13071c/>, permission to reproduce in this thesis granted according to ESO Image Usage Guidelines (<http://www.eso.org/public/outreach/copyright/>, retrieved on October 26, 2017); MUSE on VLT: Roland Bacon / ESO, source: <http://www.eso.org/public/images/eso1724d/>, permission to reproduce in this thesis granted according to ESO Image Usage Guidelines (<http://www.eso.org/public/outreach/copyright/>, retrieved on October 26, 2017). Page 83

Timeline of the universe, image credit: NASA / WMAP Science Team, source: <https://map.gsfc.nasa.gov/media/060915/index.html>, permission to reproduce in this thesis granted according to NASA Media Usage Guidelines (<https://www.nasa.gov/multimedia/guidelines/index.html>, retrieved on June 15, 2017). Page 4

List of publications

Publications presented in this thesis

1. Rexroth, Natarajan, Kneib, *A new method to break the mass sheet degeneracy using aperture moments*, 2016, MNRAS, 460, 2505
2. Rexroth, Kneib, Joseph, Richard, Her, *IFS-RedEx, a redshift extraction software for integral-field spectrographs*, 2017, arXiv:1703.09239, submitted to MNRAS
3. Rexroth, *Substructure in the Frontier Fields from weak lensing flexion*, 2015, Proceedings of the International Astronomical Union, 11, 795
4. Rexroth, Schäfer, Fourestey, Kneib, *High Performance Computing for gravitational lens modeling: single vs double precision on GPUs and CPUs*, in preparation
5. Rexroth, Kneib, Richard, Jauzac, Joseph, Comparat, *MUSE meets CLASH: Observations of the lensing cluster MACSJ1931.8-2635*, in preparation

Publications not presented in this thesis

1. Jauzac, Eckert, Schwinn, Harvey, Baugh, Robertson, Bose, Massey, Owers, Ebeling, Shan, Jullo, Kneib, Richard, Atek, Clément, Egami, Israel, Knowles, Limousin, Natarajan, Rexroth, Taylor, Tchernin, *The extraordinary amount of substructure in the Hubble Frontier Fields cluster Abell 2744*, 2016, MNRAS, 463, 3876
2. Jauzac, Richard, Limousin, Knowles, Mahler, Smith, Kneib, Jullo, Natarajan, Ebeling, Atek, Clément, Eckert, Egami, Massey, Rexroth, *Hubble Frontier Fields: predictions for the return of SN Refsdal with the MUSE and GMOS spectrographs*, 2016, MNRAS, 457, 2029
3. Limousin, Richard, Jullo, Jauzac, Ebeling, Bonamigo, Alavi, Clément, Giocoli, Kneib, Verdugo, Natarajan, Siana, Atek, Rexroth, *Strong-lensing analysis of MACS J0717.5+3745 from Hubble Frontier Fields observations: How well can the mass distribution be constrained?*, 2016, A&A, 588, 99

List of publications

4. Jauzac, Richard, Jullo, Clément, Limousin, Kneib, Ebeling, Natarajan, Rodney, Atek, Massey, Eckert, Egami, Rexroth, *Hubble Frontier Fields: A high-precision strong-lensing analysis of the massive galaxy cluster Abell 2744 using ~180 multiple images*, 2015, MNRAS, 452, 1437
5. Jauzac, Jullo, Eckert, Ebeling, Richard, Limousin, Atek, Kneib, Clément, Egami, Harvey, Knowles, Massey, Natarajan, Neichel, Rexroth, *Hubble Frontier Fields: The geometry and dynamics of the massive galaxy cluster merger MACSJ0416.1-2403*, 2015, MNRAS, 446, 4132
6. Jauzac, Clément, Limousin, Richard, Jullo, Ebeling, Atek, Kneib, Knowles, Natarajan, Eckert, Egami, Massey, Rexroth, *Hubble Frontier Fields: A high-precision strong-lensing analysis of galaxy cluster MACSJ0416.1-2403 using ~200 multiple images*, 2014, MNRAS, 443, 1549

Markus Rexroth

Observatoire de Sauverny – 1290 Versoix – Switzerland

☎ +41 22 379 24 71 • ✉ markus.rexroth@epfl.ch
🌐 people.epfl.ch/markus.rexroth

Education

Ph.D. in Physics	Ecole Polytechnique Fédérale de Lausanne (EPFL) Ph.D. advisor: Jean-Paul Kneib	12/2013 - present
M.Sc. in Physics	Eidgenössische Technische Hochschule Zürich (ETH) Master's thesis at Yale University & Drexel University Advisors: Priyamvada Natarajan, David Goldberg, Alexandre Réfrégier	09/2011 - 08/2013 01/2013 - 05/2013
B.Sc. in Physics	Eidgenössische Technische Hochschule Zürich (ETH) Semester thesis at CERN Advisor: Werner Luster	09/2008 - 03/2012 01/2011 - 02/2011

Awards & Grants

Preparatory project on the Piz Daint supercomputer	2016
Scholarship of the Studienstiftung des deutschen Volkes (German National Academic Foundation)	11/2008 - 06/2013

Research interests

Cosmology from strong & weak gravitational lensing
Dark matter constraints from galaxy clusters
High Performance Computing
The high redshift universe

Technical skills

Programming languages: Python, C++, CUDA (Graphics card acceleration), MPI, OpenMP

Software development, e.g. a pipeline for redshift extraction from integral-field spectrographs, graphics card accelerated gravitational lens modeling

Data reduction for MUSE on the VLT

Webmaster for EPFL Astrophysics and Swiss participation in the Euclid space telescope websites

Student supervision & teaching experience

Co-advisor for the Master thesis on MUSE spectroscopy and lens modeling by Baptiste Klein (visiting from ISAE-SUPAERO)	2017
Co-advisor for the Master thesis on graphics card accelerated gravitational lens modeling by Christoph Schäfer (EPFL)	2016
Advisor for summer project on MUSE spectroscopy by Romaric Her (visiting from ISAE-SUPAERO)	2016
Advisor for semester thesis on the mass-sheet degeneracy in gravitational lensing by Olivier Kauffmann (EPFL)	2016
Co-advisor for the summer project on graphics card accelerated gravitational lens modeling software by Thomas Jalabert (visiting from ENS Cachan)	2015
Teaching assistant for "Astrophysique I" at EPFL	2014, 2015
Teaching assistant for "Astrophysique II" at EPFL	2016

Telescope time

12 nights on the Swiss Leonhard Euler Telescope in La Silla, Chile	2017
13 nights on the Swiss Leonhard Euler Telescope in La Silla, Chile	2015

Professional training

EPFL staff training courses

<i>Computing on GPUs</i>	2015
<i>MPI, Advanced Parallel Programming</i>	2015
<i>MPI, an introduction to parallel programming</i>	2015

Summer & winter schools

<i>The Caltech-JPL summer school on Big Data Analytics</i>	2016
Virtual summer school on Coursera	
<i>Cosmology with Wide-Field Surveys</i>	2014
Saas-Fee winter school	
<i>Introduction to sub-mm interferometry and science with ALMA</i>	2014
Dark Cosmology Center, Copenhagen	

Professional affiliations

Member of the <i>Swiss Society for Astrophysics and Astronomy</i>	2014 - present
---	----------------

Public outreach

Astrophysics talks for the general public and tours of the Geneva observatory	2014 - present
---	----------------

Languages

English (fluent)
German (native)
French (good command)
Latin