Weak lensing magnification & baryon acoustic oscillations in galaxy-galaxy lensing

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Sandra Unruh

aus

Potsdam

Bonn, 15. Juni 2020

1. Gutachter:	Prof. Dr. Peter Schneider
2. Gutachter:	Prof. Dr. Cristiano Porciani
Tag der Promotion:	06.11.2020
Erscheinunsgjahr	2021

ABSTRACT

Cosmology is the science that aims to explain the Universe in its entirety. While the standard cosmological model has had tremendous success in explaining independent astrophysical observations, we still lack understanding of the very nature of its main constituents, namely dark matter and dark energy. To better understand their origin, we can map matter structures throughout the Universe and their evolution using the distortion of light rays as they travel through the inhomogeneous Universe. This approach is called weak gravitational lensing, and this thesis focusses on galaxy-galaxy lensing (GGL), which directly reveals the relation of the visible 'normal' matter to underlying dark matter structures. Typically, GGL is measured in terms of tangential shear, i.e., the distortion of the observed galaxy shapes with respect to foreground galaxy positions.

Shear estimates from weak lensing surveys will soon enable us to determine cosmological parameters with sub-percent accuracy. The necessary analyses require excellent control over detector systematics, a sound theoretical model, and capable numerical tools. Therefore, we first developed an open-source numerical tool to extract GGL signals efficiently and then used it to pursue the following science cases.

Weak lensing magnification describes the change of a galaxy's observed flux. It consequently changes the observed number density of galaxies on the sky, which in return affects the observed tangential shear. In this thesis, we provide leading-order analytical descriptions for the magnification effects. Further, we present numerical methods to select samples of foreground (lens) and background (source) galaxies that are unbiased by magnification. Currently, the combination of the three surveys KiDS+VIKING+GAMA provides one of the best constraints on cosmological parameters. We analysed the impact which neglecting magnification effects has on such a survey and find that, for lens galaxies at redshift $z_\mathrm{d} = 0.36$ and source galaxies with mean redshift $\bar{z}_\mathrm{s} = 0.79$, the shear profile is changed by 2% and the mass of the lens is biased by 8%. We conclude, magnification effects by source and lens

galaxies must be carefully taken into account even for ongoing surveys, while the statistical power of future weak lensing surveys certainly warrants correction for this effect.

The shear-ratio test (SRT) is a null-test that probes for systematics in galaxy shape and redshift estimates simultaneously. It is a purely geometrical probe that relies on shear and distance measurements for one foreground and two background galaxy populations. In this thesis, we show that the test is heavily biased if weak lensing magnification is not accounted for. The bias is stronger for increasing redshift of lenses and therefore, affects future surveys more severely. Using simulations, we find that an SRT with flux-limited lens galaxies at redshift $z_{\rm d} = 0.8$ deviates up to 9σ from zero. To retain the useful properties of the SRT, we provide a mitigation strategy that solely relies on already present observational data. The mitigation reduces the bias by a factor of ~ 100 and, at the same time, reduces the total uncertainties. This results in a deviation of typically $<1\sigma$.

Lastly, we explore the influence of baryon acoustic oscillations (BAO) in the GGL signal. BAO are frozen-in density fluctuations in the large-scale structure that were generated by sound waves in the early Universe. The detection of the BAO signal as a function of redshift is an excellent probe for the time evolution of dark energy. Upcoming surveys will enable us to constrain the BAO signal from GGL measurements for the first time. In this thesis, the BAO signal is first modelled analytically. Then, we aimed to compare the model to the signal estimated from weak lensing simulations. However, various problems with the weak lensing simulations were discovered that prohibit detecting BAO. Nonetheless, the analysis pipeline has been set-up successfully and improved data catalogues can be analysed on the time-scale of an hour.

CONTENTS

CHAPTER 1

INTRODUCTION

"Astronomy? Impossible to understand and madness to investigate."
– Sophocles, 496-406 BC

The Universe is simple yet bizarre. It is simple because we only need six parameters to describe and predict a multitude of astrophysical observations. These observations cover a wide range of space and time. Within the framework of our standard model, we understand the dynamics of stars in galaxies, and also the motion of galaxies in galaxy clusters, the most massive bound structures in the Universe. The cosmic microwave background (CMB) was emitted 380 000 years after the Big Bang, and the six parameters of the standard model describe the CMB-data within $1-2\%$ uncertainty. The same model also provides an excellent fit to data from the local Universe, more than 13 billion years later. We obtain the same parameters when measuring the primordial composition of the interstellar gas, as well as observations that follow the bending of light around mass. Yet, it is a strange Universe because the standard model suggests that only 5% of the energy density in the Universe is made up of atoms. Dark matter that does not interact with light fills 26% of our Universe while dark energy, an energy associated with the empty space, contributes 69%. Neither of these two is understood; dark matter candidates escape any direct detection in the laboratory, and dark energy even lacks insight into its physical mechanisms.

Cosmology studies the Universe as a whole and its temporal evolution using our best understanding of its physical laws and its composition. Starting with the formulation of Einstein's general relativity (GR), this field of science has been rapidly advancing ever since. Technological breakthroughs gift us with an unprecedented sharp view deep into the Universe. With dedicated surveys, astronomers all over the world try to shed light on the nature of dark matter and dark energy. Wide-field and deep imaging surveys for almost all wavelengths of the electromagnetic

spectrum have finished recently, are currently underway, or are planned for the near future. With the next generation of surveys, we will map about ten billion galaxies a few billion years back in time. Figure 1.1 illustrates the optical sky, populated with thousands of galaxies in a field of view smaller than the angular diameter of the moon. We see some of the galaxies as they looked when the Universe was a tenth of

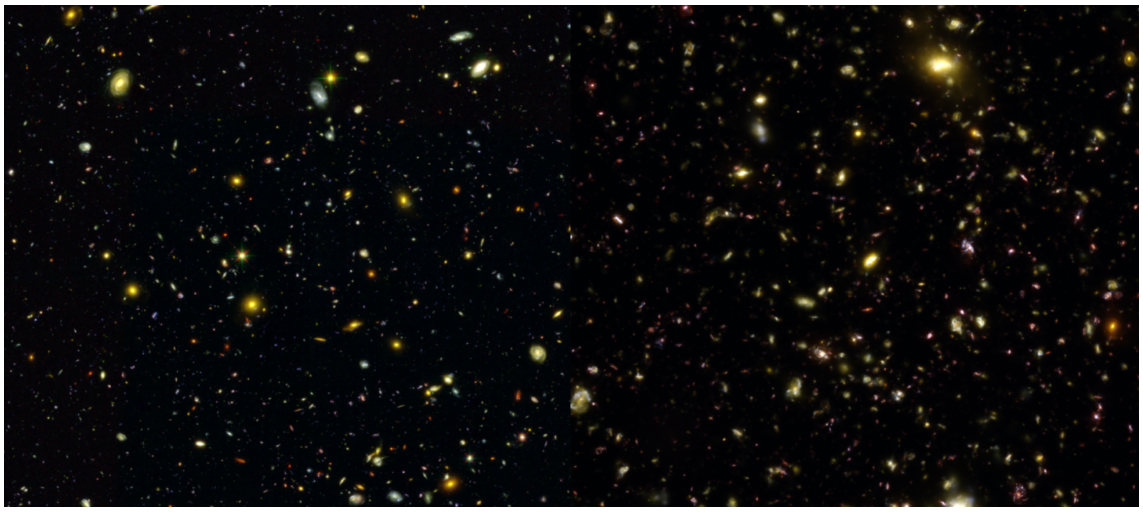


Figure 1.1. – The left-hand side is the Hubble eXtreme Deep Field which reveals several thousand galaxies in a field of view that is only a small fraction of the full moon's angular diameter. As of yet, it is the deepest image ever taken. It displays a sky that is densely populated with galaxies of various colours and shapes. The image was taken by NASA's Hubble Space Telescope from Illingworth et al. (2013). On the right, a re-creation of such a field by the Illustris simulation (Vogelsberger et al., 2014) is shown. The two sub-figures span ≈ 3 arcmin on each side.

its current age. Galaxies have colours ranging from blue to deep red and countless different shapes, which has been dubbed the 'galaxy zoo'[1]. Using information from other wavebands, especially the infrared and X-ray wavelengths, reveals different properties of the galaxies. For cosmology, we are most interested in the galaxies' distances, measured in terms of redshifts, as well as their masses.

Planned ground-based surveys are the Large Synoptic Survey Telescope (LSST, `www.lsst.org`) in the optical wavelength and the Square Kilometre Array Observatory Square Kilometre Array Observatory (SKA, `www.skatelescope.org`) for the radio sky. Near-future space-based surveys encompass the *Euclid* satellite (`www.euclid-ec.org`) in the optical and near-infrared wavebands, the James Webb Space Telescope (JWST, `www.jwst.nasa.gov`) as well as the Wide-Field Infrared Survey Telescope (WFIRST, `wfirst.gsfc.nasa.gov`) for the infrared sky, and the Advanced Telescope for High Energy Astrophysics (Athena, `www.`

[1]Galaxy Zoo is also the name of a remarkable crowd-sourced astronomy project. It invites volunteers all over the world to help classify the morphology of galaxies in recent surveys, see also, `www.zooniverse.org/projects/zookeeper/galaxy-zoo/`.

the-athena-x-ray-observatory.eu) for the X-ray wavelengths. Wide-field surveys that are currently collecting data are the Dark Energy Survey (DES, www.darkenergysurvey.org) in the optical, the Hyper Surprime-Cam Subaru Strategic Program (HSC SSP, hsc.mtk.nao.ac.jp/ssp/) for the optical and near-infrared sky, and the extended Roentgen Survey with an Imaging Telescope Array (eROSITA, www.mpe.mpg.de/eROSITA). Recently finished photometric surveys with not fully analysed data sets are the Kilo Degree Survey (KiDS, kids.strw.leidenuniv.nl) in the optical wavelength, and its partner survey the VISTA[2] Kilo-Degree Infrared Galaxy Survey (VIKING, casu.ast.cam.ac.uk/vistasp/viking). Surveys that rely on spectroscopy rather than photometry are the extended Baryon Oscillation Spectroscopic Survey (eBOSS, a component of the Sloan Digital Sky Survey – SDSS, www.sdss.org/surveys/eboss/), and the Galaxy And Mass Assembly (GAMA, www.gama-survey.org). This list of surveys, especially for the spectroscopic surveys, is not exhaustive.

In this thesis, gravitational lensing is the tool of choice to study cosmology. Lensing describes the distortion of light as it travels through the matter structures of the Universe. To tightly link observations to the underlying cosmological model, we use simulations of the Universe. As an example, simulations can be used to model the limitations of telescopes to observe the Universe. We can also test the consequences of our limited understanding of the analytical framework. The simulation that currently models our Universe best is the Illustris simulation (www.illustris-project.org) which follows the structure evolution of 'normal' matter and dark matter from very early times to today. Figure 1.1 shows on the one side Hubble's eXtreme Deep Field and on the other side a mock image from the Illustris simulation. The similarity is striking; unfortunately, the size of the simulated cube is not yet big enough to study the whole visible Universe or even its largest structures. Computationally less expensive are dark matter-only simulations of which we make extensive use in this thesis. They are accompanied by semi-analytical models for galaxy formation. While keeping the numerical accuracy in mind, we confirm and expand our analytical understanding of the statistics of matter structures in the Universe. Ultimately, this brings us a step closer to refine our standard cosmological model or even detect deviations from it.

In the following, I introduce a self-developed numerical tool to efficiently analyse two-point correlation functions which are used to study lensing magnification in-depth. Magnification or de-magnification of the incoming flux is a direct consequence of the distortion of light bundles as they travel through the Universe. However, in contrast to the change in the galaxy's apparent shape, magnification effects are less understood and more challenging to measure. If effects from lensing magnification are ignored, the inferred cosmological parameters are bound to be biased. By providing analytical insights and numerical tools, that are made publicly

[2]Visible and Infrared Survey Telescope for Astronomy

available, I hope to improve analysis pipelines, especially for the upcoming surveys. The same numerical tool is then used to study the clustering signal from baryon acoustic oscillations (BAO). BAO are frozen-in density waves that are imprinted on the CMB as well as on the distribution of matter in the local Universe approximately 13 billions years later. As of yet, they have not been detected in galaxy-shear correlation functions, which are generally referred to as galaxy-galaxy lensing (GGL). By using mock data from simulations, I try to give first insights into the detection significance for planned surveys.

The outline of the thesis is as follows. In Chapter 2, we briefly review the most critical aspects of the cosmological framework. We introduce the theoretical foundations of the standard model of our Universe and its predictions. The focus lies on the large-scale structure (LSS) of the Universe and its statistical properties. Chapter 3 introduces weak gravitational lensing as a tool to observationally study the LSS and its temporal evolution. After a general introduction, we focus on GGL, which correlates the position of galaxies to distortions of the shape from background galaxies. Chapter 4 focuses on the numerical aspects of this thesis. Simulations of matter structures in the Universe and weak lensing simulations are introduced. This is followed by a detailed account of a code developed for this work to extract shear profiles in the era of large data sets. The code is then used to study the effects of magnification in weak gravitational lensing in Chapter 5 and 6. Chapter 5 focuses on magnification bias in the shear-ratio test (SRT) and its mitigation strategies. In Chapter 6 magnification effects are discussed in general, using analytical approximations and numerical methods. Chapter 7 looks at BAO and their footprint in simulated weak lensing data. Lastly, we conclude with a summary in Chapter 8.

CHAPTER 2

COSMOLOGICAL FRAMEWORK

Humankind has always tried to describe and explain the world we live in. In the last two centuries technology advanced rapidly. We now have a multitude of observations at our disposal to understand the nature of the Universe. The concurrent development of scientific theories opened up the possibility to also explain these observations. Together, they led to the concordance or standard model of cosmology, that is referred to as the Λ *cold dark matter model (*Λ*CDM).*

The ΛCDM model is based on Einstein's well-tested GR and the observationally motivated cosmological principle. This principle states that our Universe is homogeneous and isotropic. The well-disposed reader might notice that we live in a rather inhomogeneous environment and, indeed, the average density of a human being is roughly 10^{31}-times the average density of the Universe. Cosmology, however, tries to explain the origin and the evolution of the Universe as a whole and is valid on large scales, i.e. scales that are much larger than the size of our galaxy, but still are much smaller than the size of the observable Universe. Nowadays, tensions between cosmological observation and theory exist only on a minor level and extensions to the standard theory of gravity and the cosmological principle are currently investigated. However, this thesis mainly focuses on the conventional framework of ΛCDM cosmology.

In the following, I motivate the several concepts of cosmological distances that are rather different from the well-known Euclidean space. I then explain how we expect structures in the Universe to grow depending on its constituents. For an extensive and thorough theoretical treatment the reader is referred to Peebles (1993), Dodelson (2003), and Schneider (2015).

2.1. Distance measures in cosmology

In contrast to our daily life, measuring distances in cosmology is by no means trivial. In an expanding Universe with a curved spacetime, the separation between two objects is constantly changing. Moreover, as we observe the distant Universe, we look back in time. This section provides a description of homogeneous world models and their cosmological distances.

2.1.1. Robertson-Walker metric

Since the Universe as a whole is neutral, the only fundamental force that acts on cosmological scales is gravity. In 1916, Albert Einstein formulated the theory of GR, which successfully describes gravity. In GR, space and time are expressed as a four-dimensional manifold which is shaped by its matter and energy content. This is described by Einstein's field equation

$$G_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu} - \Lambda g_{\mu\nu} \ , \tag{2.1}$$

that relates the geometry of the Universe to its energy content. The former is expressed by the Einstein tensor $G_{\mu\nu}$ and the latter by stress-energy tensor $T_{\mu\nu}$, the cosmological constant Λ, and the metric tensor $g_{\mu\nu}$. The Greek indices run over the four components '0' for time and '1, 2, 3' for space. Finally, G is the gravitational constant and c the speed of light in vacuum.

We can solve Eq. (2.1) if the cosmological principle holds. This principle is strongly supported by observations of the homogeneous and isotropic CMB (*Planck* Collaboration et al., 2018b) and the distribution of galaxies (Blanton et al., 2017). In the framework of GR, we assume the existence of fundamental observers whom all experience the same history of the Universe and observe the Universe's properties as being isotropic. In this special case, Howard P. Robertson and Arthur G. Walker showed independently that a rather simple solution of Einstein's field equations exists (Robertson, 1935; Walker, 1937),

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left[\mathrm{d}\chi^2 + f_K^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)\right] \ , \tag{2.2}$$

where t is the cosmic time, $a(t)$ the scale factor, χ the radial comoving coordinate, and θ and φ are the angular coordinates on a unit sphere. $f_K(\chi)$ is the comoving transverse distance which is a function of the Universe's curvature K

$$f_K(\chi) = \begin{cases} K^{-1/2}\sin[K^{1/2}\chi] & \text{for} \quad K > 0 \ , \\ \chi & \text{for} \quad K = 0 \ , \\ (-K)^{-1/2}\sinh[(-K)^{1/2}\chi] & \text{for} \quad K < 0 \ . \end{cases} \tag{2.3}$$

Observers on constant comoving coordinates (χ, θ, φ) are called comoving observers and coincide in the Robertson-Walker metric with fundamental observers.

A cosmological model that is based on the Robertson-Walker metric forms either an expanding or contracting spacetime. Vesto Slipher was the first person to obtain observational evidence that we live in an expanding Universe; he found that most galaxies show a spectrum shifted to lower photon energies. Subsequently, Edwin Hubble interpreted this 'redshifting' as a receding motion from us. In general, in a homogeneously expanding (or contracting) sphere with uniformly distributed density $\rho(t)$, a particle is chosen to be at spatial position $\boldsymbol{x}$ at initial time t_0 and is at position

$$\boldsymbol{r}(t) = a(t)\boldsymbol{x} \ , \tag{2.4}$$

at time t. The scale factor is conveniently chosen as $a(t_0) = a_0 = 1$ today and $\boldsymbol{x}$ is the time-independent comoving coordinate. Hence, the world-line $(\boldsymbol{r}, t)$ of any comoving particle is completely characterised by $(\boldsymbol{x}, a(t))$. In other words, if we determine the behaviour of the scale factor with time, we understand the dynamics of the Universe.

2.1.2. The Friedmann equations

When inserting the metric tensor (2.2) into Einstein's field equations (2.1), the resulting stress-energy tensor $T_{\mu\nu}$ represents a perfect fluid with density $\rho(t)$ and pressure $p(t)$. The field equations reduce to two differential equations, the so-called Friedmann equations, that describe the dynamics of the scale factor

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda c^2}{3} \ , \tag{2.5}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \ . \tag{2.6}$$

The cosmological constant Λ, as introduced by Einstein, can also be written as the vacuum or dark energy density $\rho_\Lambda = \Lambda c^2/(8\pi G)$ with negative pressure $p_\Lambda = -\rho_\Lambda c^2$. When the two Friedmann equations are combined, we obtain the cosmological equivalent of the first law of thermodynamics. The adiabatic equation reads

$$\frac{\mathrm{d}(a^3\rho c^2)}{\mathrm{d}t} = -p\frac{\mathrm{d}a^3}{\mathrm{d}t} \ . \tag{2.7}$$

It states that the temporal change of energy ρc^2 contained in a comoving volume is equal to the expansion/contraction work in this volume.

To understand the time evolution of the scale factor, we still need a relation between energy density and pressure, which we call the equation-of-state (EOS). In general, the EOS is

$$\frac{p}{\rho c^2} = w \ , \tag{2.8}$$

where all components can be time-dependent in principle. Inserting the equation above into Eq. (2.7) yields for any component 'x' of the Universe

$$\rho_{\mathrm{x}}(a) = \rho_{\mathrm{x},0} \exp\left[-3 \int_1^a \mathrm{d}a' \frac{1 + w_{\mathrm{x}}(a')}{a'}\right] . \tag{2.9}$$

The three main components of the Universe and their EOS parameter are

$$\begin{aligned} \text{matter:} \quad & w_{\mathrm{m}} = 0 \quad \dashrightarrow \quad \rho_{\mathrm{m}} \propto a^{-3} , \\ \text{radiation:} \quad & w_{\mathrm{r}} = 1/3 \quad \dashrightarrow \quad \rho_{\mathrm{r}} \propto a^{-4} , \\ \text{dark energy:} \quad & w_{\Lambda} = -1 \quad \dashrightarrow \quad \rho_{\Lambda} = \text{ const } , \end{aligned} \tag{2.10}$$

where the dependence of the scale factor on time is obtained by solving the first Friedmann equation (2.5). Since all components scale differently with scale factor, the energy density of matter, radiation and dark energy dominate the Universe energy budget at different times. The constituents of the Universe are introduced more thoroughly in Sect. 2.2.

To express the first Friedmann equation in a form that gives us a more intuitive insight into the expansion history of the Universe, it is useful to introduce the Hubble parameter $H(a)$. The parameter $H(a)$ describes the expansion rate of the Universe and is defined as $H(a) = \dot{a}/a$; the relation between scale factor and cosmic time is then expressed as

$$t(a) = \int_0^a \frac{\mathrm{d}a'}{a' \, H(a')} . \tag{2.11}$$

The Hubble parameter that we measure today in the local Universe is usually referred to as the Hubble constant H_0, which is one of the fundamental parameters of cosmology and, therefore, a quantity of great interest. Historically, it was (and is still) useful to parametrise our ignorance of the true value with $H_0 = 100h\,\mathrm{km/s/Mpc}$. An accurate estimate of H_0 is difficult to obtain and our current best estimates indicate $h \approx 0.7$ (Wong et al., 2019; Hotokezaka et al., 2019; Domínguez et al., 2019; Riess et al., 2019; Ryan et al., 2019; Macaulay et al., 2019; *Planck* Collaboration et al., 2018b), where all the references cited make use of a different branch of astronomy, namely: strong gravitational lensing, gravitational waves, gamma-ray attenuation, cosmological distance ladder, BAO, Supernova Type Ia, and the CMB.

Using the Hubble parameter and (2.5), we can calculate the expression for the current density of a flat Universe. This quantity is called the critical density and reads $\rho_{\mathrm{crit}} = 3H_0^2/(8\pi G)$. We can use it to rewrite the first Friedmann equation to

$$\left(\frac{H(a)}{H_0}\right)^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{\Omega_{\mathrm{r},0}}{a^4} + \frac{\Omega_{\mathrm{m},0}}{a^3} + \frac{1 - \Omega_{0,0}}{a^2} + \Omega_{\Lambda,0} . \tag{2.12}$$

Here, we introduced today's unitless density parameter $\Omega_{\mathrm{x},0} = \rho_{\mathrm{x},0}/\rho_{\mathrm{crit},0}$. If not noted otherwise, the density parameters are all set to the current epoch without the explicit '0'-index. The total energy density parameter Ω_0 is defined as $\Omega_0 = -Kc^2/H_0^2 = \Omega_\mathrm{r} + \Omega_\mathrm{m} + \Omega_\Lambda$. Then, the history of cosmic expansion (2.12) is fully determined by H_0 and the density parameters Ω today.

2.1.3. The concept of distance

In Euclidean space it is, with sufficient information, possible to obtain a unique distance. Along a backward light cone in a curved spacetime, however, we have to define carefully what type of distance we are investigating. Commonly used distances relate the diameter of an object to its apparent angle on the sky or compare the flux of an object with intrinsic luminosity. Alternatively, the distance can characterise the photon's light travel-time. In general, two distance measures differ in a curved, non-steady spacetime. In an expanding or contracting universe distances also constantly change over time. Hence, it is quite appealing to define a comoving distance independent of the expansion.

Comoving distance along the line-of-sight

The (radial) comoving distance of two objects without peculiar velocities is a central quantity in cosmological investigations which can be related to all other distance measures. A photon heading towards an observer follows a null geodesic, i.e. $\mathrm{d}s^2 = 0$. Hence, for a radial light ray with $\mathrm{d}\theta = \mathrm{d}\varphi = 0$, Eq. (2.2) reduces to

$$c\mathrm{d}t = -a(t)\mathrm{d}\chi \,, \tag{2.13}$$

and we observe photons from a source at time t today at comoving distance

$$\chi(t) = \int_t^{t_0} \frac{c\,\mathrm{d}t'}{a(t')} = \int_{a(t)}^{a(t_0)} \frac{c\,\mathrm{d}a}{a^2 H(a)} \,, \tag{2.14}$$

where in the second equality we have inserted the definition of the Hubble parameter and, therefore, highlighted the dependence of the comoving distance on the cosmological parameters. The largest comoving distance that allows two comoving observers to have been in causal contact is called the horizon today and obtained in the limit of $a \to 0$.

Cosmological redshift

An effect of the expanding Universe is that photons constantly lose energy as they travel to us, this phenomenon is called redshift. Redshift is defined as

$$1 + z = \frac{\lambda_\mathrm{obs}}{\lambda_\mathrm{em}} = \frac{1}{a} \,, \tag{2.15}$$

where $\lambda_{\rm obs}$ is the observed wavelength and $\lambda_{\rm em}$ the wavelength emitted in the galaxy's rest-frame. Locally, redshift can be interpreted as the source's recession velocity v using the Hubble Law

$$v = H_0 D = cz \ , \tag{2.16}$$

where the source is located at distance D.

Making use of the definition of redshift, Eq. (2.14) is rewritten as

$$\chi(z_1, z_2) = \int_{z_1}^{z_2} \frac{c\,\mathrm{d}z'}{H(z')} \ , \tag{2.17}$$

which is generally known as the distance-redshift relation. This relation allows us to infer cosmological parameters if we are able to obtain the redshift and the comoving distance independently from each other.

Comoving separation in the transverse direction

The transverse comoving distance $f_K(\chi)$ as defined in (2.3). Then the comoving separation of two comoving objects at the same redshift and observed under an angle $\Delta\theta$ is $R = f_K(\chi)\,\Delta\theta$, where we made use of the small-angle approximation.

Angular diameter distance

The angular diameter distance relates the physical transverse size d of an object to its observed angular size θ on the sky. Thus, it is closely related to the transverse comoving distance $f_K(\chi)$ by

$$D_{\rm ang}(a) = a f_K(\chi) = \frac{d}{\theta} \ . \tag{2.18}$$

It is also noteworthy that the angular diameter distance is not a monotonously increasing function with redshift, but it shows a maximum value at $z \approx 1.7$ or equivalently $a \approx 0.4$ for our Universe. Thus, objects with the same physical size appear larger at $a = 0.3$ than those nearer to us at $a = 0.4$.

In general, the transverse comoving and the angular diameter distance do not follow an additive relation for two objects at redshifts z_1 and z_2, e.g., $D_{\rm ang}(z_1, z_2) = a(z_2)\, f_K[\chi(z_1, z_2)] \neq D_{\rm ang}(z_2) - D_{\rm ang}(z_1)$. Only in the case of vanishing curvature follows the transverse comoving distance such a relation. However, in the limit of small distances, the additive relation can be used as a good approximation.

Luminosity distance

An isotropic light source with bolometric luminosity L is observed with a bolometric flux S that decreases with the inverse square of its distance. So, we can define the

luminosity distance as

$$D_{\mathrm{lum}} = \sqrt{\frac{L}{4\pi S}} \,. \tag{2.19}$$

However, the Universe is expanding while photons travel to us, which additionally dilutes the measured flux compared to what we would expect in a Euclidean metric. Hence, luminosity and angular diameter distance are not the same in general. Etherington's reciprocity theorem (1933) states that for bolometric quantities the distances obey the relation

$$a^2 D_{\mathrm{lum}} = a f_K(\chi) = D_{\mathrm{ang}} \,. \tag{2.20}$$

In Fig. 2.1 the angular diameter distance, the luminosity distance, the comoving distance, and the Hubble law are shown as a function of redshift.

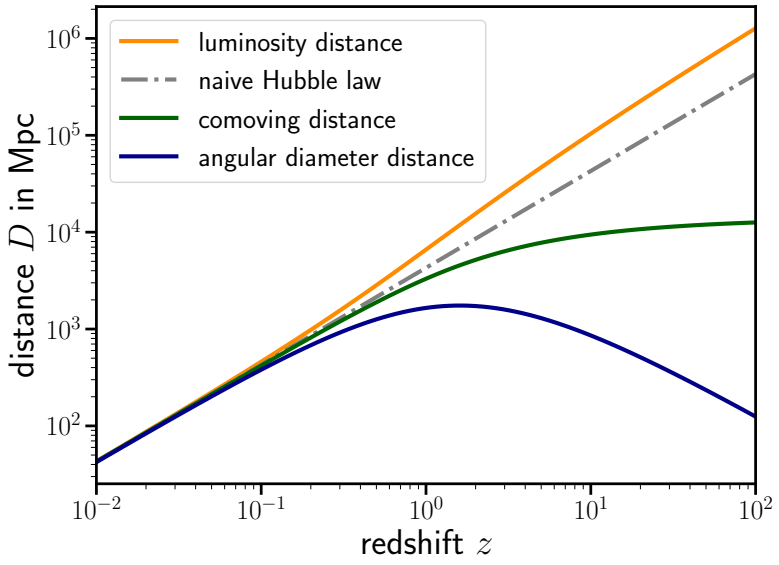


Figure 2.1. – Different cosmological distance measures are compared as a function of redshift z. The distances are computed for a flat ΛCDM model with $h = 0.7$, $\Omega_{\mathrm{m}} = 0.3$, and $\Omega_\Lambda = 0.7$; for the naive Hubble law $H(z) = H_0$ is assumed. In the local Universe, i.e. for $z \ll 1$, all distance measures agree, for higher redshifts they diverge with $D_{\mathrm{ang}} < \chi < D_{\mathrm{hubble}} < D_{\mathrm{lum}}$. For very high redshifts, the luminosity distance diverges, while the comoving distance converges towards the size of the observable Universe, and the angular diameter distance decreases after a maximum at $z \approx 1.7$.

2.2. The Universe's composition

According to the Friedmann equations, the content of the Universe governs its expansion history. Major effort has been done to develop tools to measure the properties

of the Universe's constituents and infer the corresponding cosmology. Identifying particles that govern the mass-energy budget of the Universe has not been completed yet. In the following, a summary is given, what we do know about the constituents of the Universe, what we assume to be true, and what we still have to figure out.

2.2.1. Baryonic matter

We start our account with the matter component that we have studied for the longest time. Baryons include neutrons, protons, and the atoms they form. To the dismay of particle physicists, cosmologists also include the leptonic electrons and muons to the baryonic content[1]. By electromagnetic interaction mediated through photons, they form all the luminous matter in the Universe. Observations from the CMB indicate that today's unitless baryon density parameter is $\Omega_\mathrm{b} = 0.0489 \pm 0.0003$ (*Planck* Collaboration et al., 2018b).

All the stars and galaxies that we see with our telescopes only make up approximately 10% (Fukugita and Peebles, 2004) of the baryonic matter and therefore only $\approx 0.5\%$ of the total energy budget of the Universe. Further $50-60\%$ of the baryonic matter is detected as cold clouds of gas between galaxies, which leave $30-40\%$ of the baryons unaccounted for. The missing baryons are postulated to reside as a diffuse warm-hot gas that is bound to galaxy groups but never collapsed to form a galaxy. Postulated temperatures are $10^5 - 10^6\,\mathrm{K}$ with corresponding peak emission in the far-ultraviolet and soft-X-ray radiation. This makes the warm baryonic gas very hard to detect since the interstellar medium in our galaxy is opaque for these wavelengths. Last year, several groups (Nicastro et al., 2018; Tanimura et al., 2019; de Graaff et al., 2019) found the first observational evidence of the existence of a diffuse warm-hot gas between galaxies with just the right amount to solve the missing baryon problem. Follow-up observations using quasars as well as the so-called thermal Sunyaev-Zel'dovich effect are currently in progress.

2.2.2. Dark matter

In 1937, Fritz Zwicky compared the dynamical mass of the Coma galaxy cluster to the mass expected from its luminosity and concluded the dominant amount of matter in the cluster must be non-luminous or 'dark'. Today, observations of galactic dynamics, CMB, primordial nucleosynthesis, X-ray gas analysis, and masses from gravitational lensing confirm the existence of dark matter on all scales out to the highest observable redshifts (for a review see, e.g., Freese, 2009). Dark matter behaves, like baryonic matter, as a pressureless fluid with an EOS parameter of $w = 0$, but is roughly five times more abundant than baryonic matter,

[1] to not only offend one branch of science, they commonly refer to any atom that is heavier than Helium as a metal.

i.e. $\Omega_\mathrm{m} = \Omega_\mathrm{b} + \Omega_\mathrm{dm} = 0.3111 \pm 0.0056$ (*Planck* Collaboration et al., 2018b). Following the Friedmann equations (2.12), matter becomes the dominant energy source in the Universe at redshift $z_\mathrm{eq} \approx 3400$, i.e. approximately 50 000 years after the big bang. Since a redshift of $z \approx 0.3$ or roughly 3.5 Gyr ago, dark energy governs the energy content in the Universe. Hence, matter and particularly dark matter governs most of the structure formation in the Universe. Its influence on the structure formation is determined by the weight of the dark matter particles. Light particles, referred to as hot dark matter, are free-streaming, which means all small-scale perturbations are erased, and the largest structures form first. On the other hand, a cold dark matter scenario of massive dark matter particles lets small structures, like galaxies, form first and large structures, like galaxy clusters, only emerge later. This is called a bottom-up or hierarchical scenario. A mix from dominantly cold and a small fraction of hot dark matter agrees well with observations of our Universe.

While the existence of cold dark matter is observationally strongly favoured and therefore widely accepted, we have not yet identified its physical origin. Astrophysical candidates like brown dwarfs or primordial black holes are referred to as massive compact halo objects (MACHOs) and have been excluded by microlensing studies. At most, they account for a small fraction of the total amount of dark matter (Alcock et al., 2000; Niikura et al., 2019). There exist many, many speculative possibilities on the nature of dark matter particles. Nowadays, dark matter is widely believed to be a massive particle that interacts weakly (WIMP). Bertone et al. (2005) gives an overview of possible particle candidates; examples for possible candidates are the neutralino, the least massive particle predicted by supersymmetry theories, thermal relics from the big bang, or the very massive WIMPzillas. X-ray and gamma-ray surveys are currently conducted to measure a possible spectral signature of dark matter self-annihilation, while several long-term experiments aim for direct detection of these elusive particles. Using high-energy proton-proton collisions, the High-Luminosity Large Hadron Collider (HL-LHC) will attempt to generate dark matter particles by 2026. Besides, with underground experiments, that are shielded from the ambient nuclear radiation as well as cosmic rays, several groups try to observe rare collisions of dark matter particles with regular atoms, where the reader is referred to Schumann (2019) for an extensive overview of direct detection experiments.

2.2.3. Radiation

Radiation includes all particles with zero rest mass as well as massive particles at relativistic energies; it is described as a pressurised fluid with $w = 1/3$. As the Universe keeps cooling, the number of relativistic particle species declines and radiation only dominates the energy content of the Universe for redshifts $z > z_\mathrm{eq}$. Today's best estimate is $\Omega_\mathrm{r} \approx 10^{-5}$ (*Planck* Collaboration et al., 2018b). Photons

dominate the radiation energy density, but also, depending on their rest mass, the lightest neutrino family might still be relativistic in the current epoch.

Neutrinos are produced in the early Universe in large quantities and they form a homogeneous and isotropic neutrino background with an average temperature that is slightly lower than the one of the CMB photons. The *Planck* Collaboration et al. (2018b) estimates the number of neutrino families to be $N_{\mathrm{eff}} = 3.0 \pm 0.3$, consistent with predictions from the standard model of particle physics. As an extension to the standard model, observations of neutrino oscillations require the neutrinos to have a small, but non-zero rest mass (Fukuda et al., 1998; Ahmad et al., 2001), which actually makes neutrinos the only proven-to-exist dark, weakly interacting, and massive particles. Since neutrinos form hot dark matter, they cannot account for the structure evolution in the Universe that we observe. However, they at least pose a fraction of dark matter and must be included in theoretical and numerical considerations. The exact contribution cannot be quantified yet since the individual neutrino masses are not known. However, neutrino oscillation experiments (Capozzi et al., 2017) and cosmological observations (*Planck* Collaboration et al., 2018b) constrain the sum of the neutrino masses to be $0.06\,\mathrm{eV} < \Sigma m_\nu < 0.12\,\mathrm{eV}$; oscillation experiments further give precise constraints to the mixing angles and squared mass differences with ambiguity on the sign (Gerbino and Lattanzi, 2017). In the current epoch, at least two neutrino families are non-relativistic.

2.2.4. Dark energy

The cosmic expansion is accelerated. This observation was independently made in 1998 by Adam Riess et al. and in 1999 by Saul Perlmutter et al. They investigated the cosmic expansion history by analysing the distance-redshift relation of Supernovae Type Ia and found that the Universe's expansion is not decelerating as would be expected from a Universe that is filled with matter and radiation. The term 'dark energy' was invented to denote a mysterious and unknown constituent of the Universe. Dark energy dominates the total energy content of the Universe in the current epoch with $\Omega_\Lambda = 0.6889 \pm 0.0056$ (*Planck* Collaboration et al., 2018b).

According to the second Friedmann equation (2.6), an accelerated Universe with $\ddot{a} > 0$ is only possible if the dominating energy component has negative pressure, which implies $w < -1/3$. In the special case of an EOS parameter of $w = -1$, the energy density ρ_Λ is constant in time (although Ω_Λ is still a function of time through the definition of the critical energy density) and can be related to the cosmological constant Λ as seen in the Friedmann equations (2.5) and (2.6). In this case, today's value of the cosmological constant is $\Omega_\Lambda = \Lambda c^2/3H_0^2$. A ΛCDM Universe implies that dark energy is indeed consistent with the cosmological constant and a wide range of cosmological observations support this assumption (see Fig. 2.2's upper panel). The best current estimate is $w = -1.01 \pm 0.06$ (Alam et al., 2017). The true EOS parameter of dark energy is one of the main science drivers for several planned

experiments like the Large Synoptic Survey Telescope (LSST) and *Euclid*. These experiments will also unveil the time dependence of the EOS parameter, which is usually parametrised as

$$w(a) = w_0 + w_a\,(1-a) + \mathcal{O}(a^2)\,, \tag{2.21}$$

with $w_0 = -1$ and $w_a = 0$ for a cosmological constant. Recent results can be seen in the lower panel of Fig. 2.2. While w_0 and w_a are likely to be determined soon, the physical mechanism behind the acceleration remains a mystery. It may even pose the biggest outstanding problem in physics today, and the reader is kindly referred to a review of Huterer and Shafer (2018) for the current state of the art.

2.2.5. Total energy budget

The total energy density parameter Ω_0 is the sum of all energy densities and it relates to the Universe's curvature K. Summing up the contributions of all the components mentioned above leads to $\Omega_0 \approx 1$ or more exact $1 - \Omega_0 = 0.0007 \pm 0.0019$ (*Planck* Collaboration et al., 2018b). Thus, the Universe is flat, and since it can be shown that the sign of K is preserved in cosmic evolution, a flat Universe has always been and will always remain flat. It has an average density that coincides with its critical density, which is $\bar{\rho}(t_0) = \rho_{\mathrm{crit}}(t_0) = 8.6 \times 10^{-27}\,\mathrm{kg/m^3}$ or approximately ten hydrogen atoms per cubic meter.

Since the Universe is flat, the term proportional to $1 - \Omega_0$ vanishes in the first Friedmann equation (2.12). This fact and the negligible contribution of the radiation energy density Ω_r makes it reasonable to define the cosmic evolution function

$$E^2(z) = \Omega_\mathrm{m}(1+z)^3 + \Omega_\Lambda\,. \tag{2.22}$$

The first Friedmann equation is thus sometimes written as $H(z) = H_0 E(z)$ and in this limit, the cosmic expansion in the flat ΛCDM model is fully determined by the three parameters H_0, Ω_m, and Ω_Λ.

2.3. Structures in the Universe

Fortunately, the Universe is not homogeneous at all scales. Scales below $200\,h^{-1}$Mpc comoving show rich structure, which has been revealed by observations (Sloan Digital Sky Survey (SDSS), Blanton et al., 2017). Great effort has been made to understand the structure formation in the Universe theoretically and numerically. Observations of the CMB (*Planck* Collaboration et al., 2018b) reveal a very smoothly distributed matter at redshift $z = 1100$, where deviations from the average density are of the order 10^{-5}. Starting from these tiny fluctuations in the earliest phases of the Universe, structures gravitationally collapsed, then baryonic and dark matter created the first galaxies and eventually the Earth with its inhabitants. The

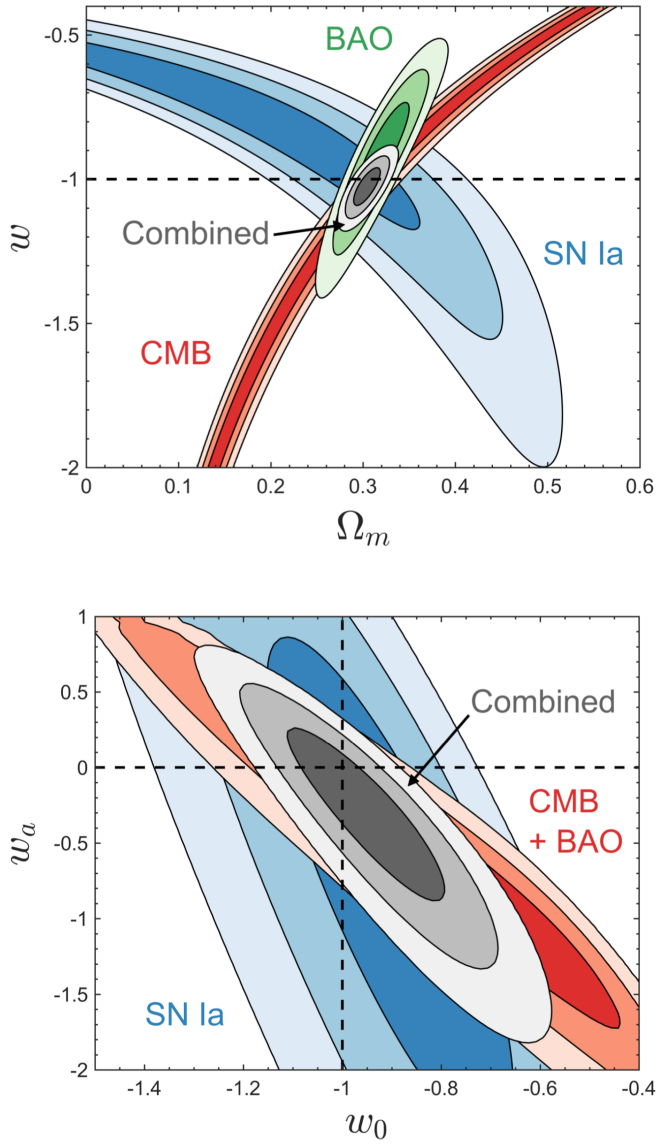


Figure 2.2. – Shown are observational constraints in the $\Omega_m - w$ parameter space for a dark energy parameter that is constant in time (upper panel). Constraints on w_0 and w_a for time-dependent dark energy are shown in the lower panel. The data are provided by Supernova Type Ia observations (SN Ia by Joint Light-curve Analysis, Betoule et al., 2014, in blue), observations of BAO (by Baryon Oscillation Spectroscopic Survey – BOSS, Alam et al., 2017, in green) and CMB observations (by the *Planck* survey, *Planck* Collaboration et al., 2016, in red). Contours indicate model likelihoods of 68.3%, 95.4%, and 99.7%. The three surveys probe different underlying physical concepts, and they also probe a wide range in redshift. Thus, it is even more amazing how well all constraints coincide, forming the ΛCDM concordance model. Both figures are adopted from Huterer and Shafer (2018).

time scale on which the structure formation happens is heavily influenced by the amount of relativistic and non-relativistic particles and the amount of dark energy in the Universe. In the following, the theoretical basis of structure formation in the Universe is outlined.

2.3.1. An analytic approach to density fluctuations

When the Robertson-Walker metric (2.2) is inserted into Einstein's field equations (2.1), the energy-momentum tensor takes the form of a perfect fluid. As long as matter is dominating the energy content in the Universe, we can focus our analytical treatment on a pressureless fluid. As a further simplification, we apply a Newtonian description of gravity instead of GR. Since in the limit of small scales GR is consistent with Newton's law, we can hope that for scales much smaller than the horizon at any given epoch, a Newtonian description is sufficient. In this framework, we can write the fluid equations for vanishing pressure as follows

$$\frac{\partial \rho}{\partial t} + \nabla_r \cdot (\rho \boldsymbol{u}) = 0 \,, \tag{2.23}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla_r)\boldsymbol{u} = -\nabla_r \phi \,, \tag{2.24}$$

$$\nabla_r^2 \phi = 4\pi G \rho - \Lambda \,, \tag{2.25}$$

where $\boldsymbol{u}$ is the velocity field of the fluid and ϕ the gravitational potential. Equation (2.23) is called the continuity equation and describes matter conservation. The Euler equation (2.24) expresses the motion of the fluid in the presence of the conservation of momentum; it describes the change of the fluid velocity as measured by an observer that follows the flow and it is solely influenced by the derivative of the gravitational potential. The third equation (2.25) is the modified Poisson equation with an additional term that allows for the presence of a cosmological constant.

A valid solution to this set of equations is a homogeneous and isotropic Universe. The solution agrees with the Friedmann equations in the absence of pressure. We now define the density contrast as relative deviations from the mean density of the Universe $\delta(\boldsymbol{r}, t) = [\rho(\boldsymbol{r}, t) - \bar{\rho}(t)]/\bar{\rho}(t)$, where $\bar{\rho}$ denotes the mean density of the Universe. To go beyond the trivial solution, we now consider small perturbations in δ and non-vanishing peculiar velocities $\boldsymbol{v} = \boldsymbol{u} - \dot{a}/a\,\boldsymbol{r}$. We further convert the fluid equations to comoving coordinates and expand to first order in density contrast and peculiar velocity. The resulting set of equations is

$$\frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla_x \cdot \boldsymbol{v} = 0 \,, \tag{2.26}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\,\boldsymbol{v} = -\frac{1}{a}\,\nabla_x \Phi \,, \tag{2.27}$$

$$\nabla_x^2 \Phi = \frac{3 H_0^2 \Omega_\mathrm{m}}{2a}\,\delta \,, \tag{2.28}$$

where $\Phi(\boldsymbol{x}, t) = \phi(a\boldsymbol{x}, t) + \ddot{a}a/2\,|\boldsymbol{x}|^2$ is the comoving gravitational potential. Equations (2.26) to (2.28) show no explicit dependence on the cosmological constant Λ or in other words, a uniform energy background field does not influence the formation of structure directly. However, as was shown in Sect. 2.1.2, the evolution of the scale factor depends on Λ and therefore, it indirectly affects the time scales on which the density contrast changes.

If we combine the linearised continuity, Euler, and Poisson equation, we obtain a linear homogeneous, second-order differential equation independent of $\boldsymbol{v}$ that describes the temporal evolution of the density contrast

$$\ddot{\delta} + \frac{2\dot{a}}{a}\,\dot{\delta} - \frac{3H_0^2\Omega_\mathrm{m}}{2a^3}\,\delta = 0\ . \tag{2.29}$$

The density contrast is still a function of space $\boldsymbol{x}$ and time t, but neither does $\boldsymbol{x}$ appear explicitly nor does (2.29) contain a derivative with respect to $\boldsymbol{x}$. Thus, we can conveniently factorise the solution to

$$\delta(\boldsymbol{x}, t) = D_+(t)\Delta_+(\boldsymbol{x}) + D_-(t)\Delta_-(\boldsymbol{x})\ , \tag{2.30}$$

where D_+ and D_- are two linearly independent solutions that satisfy (2.29) and $\Delta_{+,-}$ are comoving density fluctuation fields. As a general property, D_+ is growing with scale factor a, while D_- is decaying and, therefore, D_- is irrelevant for the structure formation of the current Universe. Hence, we can reduce the solution to

$$\delta(\boldsymbol{x}, a) = D_+(a)\delta(\boldsymbol{x}, a_0)\ . \tag{2.31}$$

D_+ is also known as the growth factor and can be computed explicitly for different cosmologies. It is normalised to $D_+(a_0) = 1$.

2.3.2. Density fluctuations and transfer function

Without showing the explicit derivations, we discuss linear structure formation for scales much larger than the horizon scale and also expand the discussion to the early Universe when radiation dominated its energy content.

Radiation-dominated phase

The considerations that led to Eq. (2.29) are only valid in the matter-dominated Universe. Since in the radiation-dominated phase the expansion rate of the Universe is different, the 'friction' term $\dot{a}/a$ in (2.29) changes. Using the first Friedmann equation with negligible curvature and dark energy density, we find for the friction term

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\left(\rho_\mathrm{r} + \rho_\mathrm{m}\right)\ . \tag{2.32}$$

Inserting this leads to a new homogeneous second-order differential equation that depends solely on the ratio of the energy densities $\rho_\mathrm{m}/\rho_\mathrm{r}$. The solution of this equation is $\delta \sim$ const for the radiation-dominated era and recovers $\delta \propto a$ in the matter-dominated era. Hence, matter density perturbations in the radiation-dominated phase do not grow.

Super-horizon perturbations

Studying the growth of super-horizon perturbations is crucial since we can always find a time where a perturbation of a given scale is still super-horizon. As mentioned before, we need to apply the framework of GR to investigate super-horizon perturbations. As a rough outline, one makes use of the decoupled spatial and temporal dependence of the density fluctuations and considers a homogeneous spherical perturbation. The spherical perturbation follows the same expansion equation as a homogeneous Universe. It can be shown that super-horizon perturbations still grow in the matter-dominated and radiation-dominated phase with $\delta \propto a$ and $\delta \propto a^2$, respectively.

Transfer function

The transfer function accounts for deviations in the structure growth for the different phases and scales in the Universe. Since the horizon scale grows proportional to the scale factor, any perturbation on a certain scale is super-horizon at some time. The density contrast in super-horizon perturbations grows proportional to a^2 during the radiation-dominated phase. If the perturbation enters the horizon in this phase, i.e. the horizon scale becomes bigger than the fixed perturbation scale, its growth stops and only resumes growing $\propto a$ when matter dominates the energy content of the Universe. Any perturbation that enters the horizon in the matter-dominated phase does not experience a stop of growth. The transition between a Universe that is dominated by radiation and a Universe dominated by matter is marked by $z_\mathrm{eq} = \Omega_\mathrm{m}/\Omega_\mathrm{r} - 1 = 3387 \pm 21$ (*Planck* Collaboration et al., 2018b). Thus, small-scale perturbations are suppressed compared to large-scale ones up to z_eq.

The difference in growth is quantified in the transfer function T_k, which is defined with respect to an early epoch a_i where all perturbations of interest are well above the horizon scale. We now consider wavenumbers $k = 2\pi/\lambda$, where λ is the comoving wavelength of a perturbation, and choose a wavenumber k_s that enters the horizon during the matter-dominated phase. Then, we can compare all density fluctuations to a non-suppressed one

$$\frac{\delta(k, a=1)}{\delta(k_\mathrm{s}, a=1)} = T_k \frac{\delta(k, a=a_\mathrm{i})}{\delta(k_\mathrm{s}, a=a_\mathrm{i})} \,. \tag{2.33}$$

The exact form of T_k also depends on the amount of massive particles, that erase their own density perturbations by free-streaming. For a cold dark matter only

Universe that contains no baryons, the transfer function has the following form (Bardeen et al., 1986)

$$T(q) = \frac{\ln(1+2.34q)}{2.34q}\left[1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4\right]^{-1/4}, \qquad (2.34)$$

where $q = k\,\mathrm{Mpc}/(\Omega_\mathrm{m}h^2)$. This function approaches unity in the limit of very large scales and is $\propto k^{-2}$ for scales much smaller than the horizon.

Baryons

In the derivations so far, we implicitly assumed that matter only interacts gravitationally. This assumption is justified since dark matter is five times more abundant in the Universe than baryonic matter and, therefore, dominates the structure formation process. However, due to the electromagnetic interaction of photons and baryons, the history of baryonic perturbation growth is quite different from that of dark matter.

The early Universe was extremely hot and has been cooling down ever since. Due to the high temperatures, electrons were not confined in neutral atoms until the phase of recombination, which happened roughly 380 000 years after the big bang. Before recombination, photons and electrons were tightly coupled via Thomson scattering, while Coulomb scattering coupled electrons and baryons. Solving the fluid equations for this baryon-photon fluid shows that its perturbations are oscillating. The gravity of the dominating dark matter causes the baryons to fall into their potential well, while the photon's pressure drives the baryons outward again, which results in an oscillating motion, the so-called baryon acoustic oscillations. The largest distance an acoustic density wave could travel until the time of recombination was 150 Mpc comoving. Shortly after recombination, atoms become neutral, the pressure of the photon-baryon fluid drops to zero and the acoustic density waves freeze out. Driven by gravity, most baryons quickly fall into the already formed potential wells of dark matter halos. Later, both matter components keep growing at the same rate.

The remnants of the density waves can still be observed today in the LSS (Ross et al., 2017). Observing the angular BAO scale as a function of redshift is a major science goal of upcoming surveys. The angular BAO scale is related to the comoving scale by the distance-redshift relation (2.17). Using the condition that the resulting distance must match the comoving scale at recombination, cosmological parameters can be inferred. Therefore, BAO are known as the Universe's standard ruler.

2.4. A statistical description of the Universe

2.4.1. Two-point statistics and power spectra

In the following, we introduce some mathematical tools that allow us to describe the density field of the Universe statistically.

Correlation function

The homogeneous and isotropic density field has the two-point correlation function

$$\xi(|\boldsymbol{x}-\boldsymbol{x}'|) = \langle \delta(\boldsymbol{x})\delta^*(\boldsymbol{x}')\rangle \ , \tag{2.35}$$

where the asterisk * denotes the complex conjugate and $\langle ...\rangle$ an ensemble average. The ensemble average is, due to the lack of an ensemble of universes, replaced by a volume average of well-separated patches on the sky. The restriction to one of all possible observable universes leads to the so-called cosmic variance, an uncertainty in measurements that is especially present on large scales. Finally, homogeneity is ensured by only including the separation $\boldsymbol{x}-\boldsymbol{x}'$ of two points and isotropy follows from not depending on the direction of the separation vector $|\boldsymbol{x}-\boldsymbol{x}'|$.

Power spectrum

In many occasions, it is useful to quantify the density fluctuations in Fourier representation

$$\delta(\boldsymbol{x},t) = \int_{\mathbb{R}^3} \frac{\mathrm{d}k}{(2\pi)^3}\, \tilde{\delta}(\boldsymbol{k},t)\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \ . \tag{2.36}$$

A Fourier transform of (2.31) shows that the different Fourier modes $\tilde{\delta}$ evolve mutually independently from each other. Of course, this conclusion is limited to the linear regime and mode coupling is an important effect at small scales. The ensemble average can be written as

$$\langle \tilde{\delta}(\boldsymbol{k})\tilde{\delta}^*(\boldsymbol{k}')\rangle = (2\pi)^3\delta_\mathrm{D}(\boldsymbol{k}-\boldsymbol{k}')P(|\boldsymbol{k}|) \ , \tag{2.37}$$

with δ_D being the three-dimensional Dirac delta functional and P the Fourier transform of the correlation function

$$P(|\boldsymbol{k}|) = \int_{\mathbb{R}^3} \mathrm{d}^3y\, \xi(|\boldsymbol{y}|)\mathrm{e}^{-\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{k}} \ . \tag{2.38}$$

Gaussian random fields

Theories, that involve inflation as a phase in the very early Universe, suggest that the primordial density field is a realization of a Gaussian random field. A random

field is characterised by a probability distribution for the occurrence of a specific random realization and can be described using ensemble or volume averages. Let $\tilde{\delta}(\boldsymbol{k}, t)$ be the Fourier transform of the random field $\delta(\boldsymbol{x}, t)$. If we discretise Eq. (2.36) for a finite volume L^3 and a fixed time t in the following way

$$\delta(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}} \frac{\tilde{\delta}(\boldsymbol{k})}{L^3} \mathrm{e}^{-i\boldsymbol{x}\cdot\boldsymbol{k}} := \sum_{\boldsymbol{k}} \delta_{\boldsymbol{k}} \mathrm{e}^{-i\boldsymbol{x}\cdot\boldsymbol{k}} \,, \tag{2.39}$$

then, a Gaussian random field has the following properties: (i) The Fourier components $\delta_{\boldsymbol{k}}$ are mutually statistically independent. (ii) The probability density of a particular $\delta_{\boldsymbol{k}}$ is a Gaussian, which follows from the first property through the central limit theorem. (iii) It can be shown from the first two properties that any Gaussian random field is fully characterised by its two-point correlation function or its power spectrum. So far, there is no observational evidence that higher-order spectra are necessary to describe the data. A so-called primordial non-Gaussianity would indicate that our still incomplete understanding of inflation is wrong (e.g., Bartolo et al., 2004).

Properties the dark matter power spectrum

As discussed before, we only need to specify the initial conditions of the density field and follow the evolution of the primordial density perturbations. Then, we have a complete description of the statistical properties of the density field of the Universe. As long as linear theory of structure evolution is valid, the Gaussianity of the density fluctuations is preserved. Thus, specifying an initial power spectrum fully characterises linear density perturbations at all times.

In the very early Universe, all length scales of interest are much larger than the horizon scale, so that there exists no preferred length scale. The only mathematical function that does not depend on a characteristic scale is a power law with spectral index n_s

$$P_0(k) \propto k^{n_\mathrm{s}} \,. \tag{2.40}$$

If we now require density fluctuations to be scale-invariant, i.e. they all enter the horizon with the same amplitude, we can constrain the spectral index to $n_\mathrm{s} = 1$ (Harrison, 1970; Zeldovich, 1972). Inflationary models predict a spectral index which is close to, but slightly smaller than one ($n_\mathrm{s} \lesssim 1$; Vázquez et al., 2018). Recent measurements from the *Planck* mission constrain the spectral index to $n_\mathrm{s} = 0.9665 \pm 0.0038$, which provides strong evidence in favour of inflation.

Combining the preceding considerations with linear structure growth, we find for the initial power spectrum

$$P_0(k) = A\, k^{n_\mathrm{s}}\, T_k^2 \,, \tag{2.41}$$

and for the evolution of the power spectrum

$$P(k,a) = A\,k^{n_\mathrm{s}}\,T_k^2\,D_+^2(a)\;, \tag{2.42}$$

where A is the amplitude related to the primordial fluctuations, which must be constrained through observations. A measurement of the power spectrum is shown in Fig. 2.3.

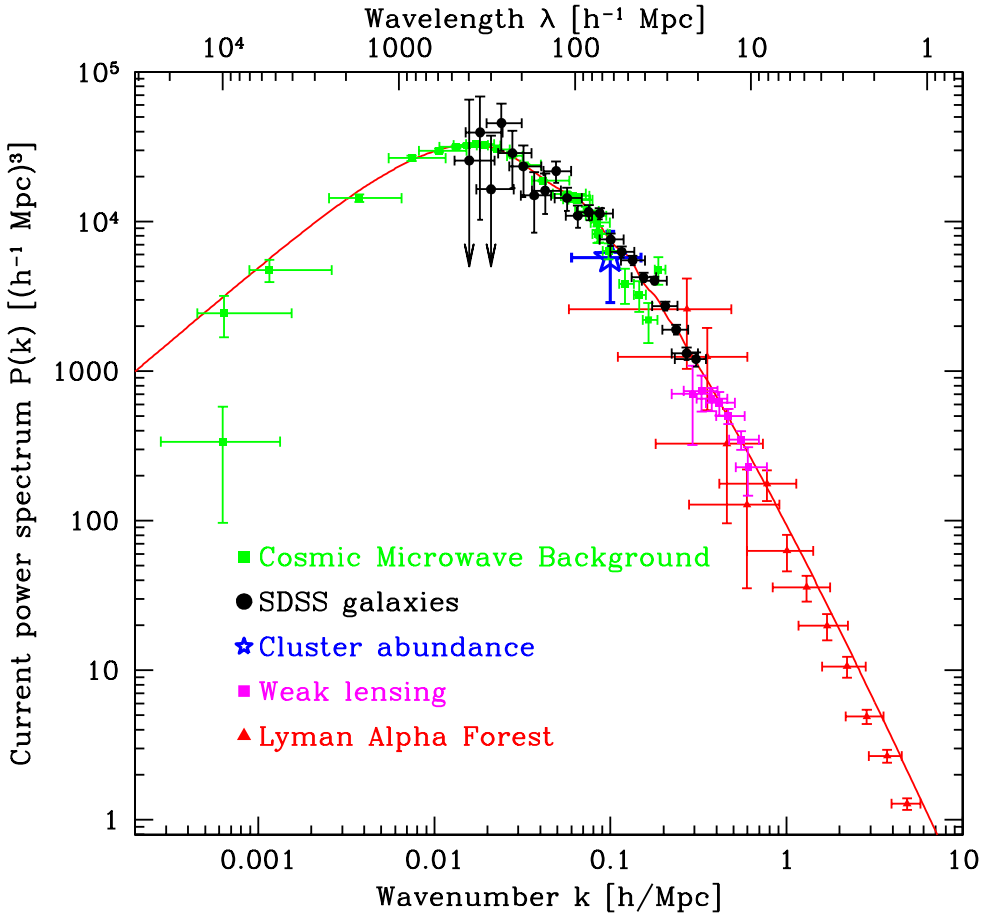


Figure 2.3. – Shown is the matter power spectrum $P(k)$ from various probes of the LSS. A prediction is shown as the red solid line for a ΛCDM cosmology with $h = 0.72$, $\Omega_\mathrm{m} = 0.28$, $\Omega_\mathrm{b} = 0.448$, and $n_\mathrm{s} = 1$. The figure is adopted from Tegmark et al. (2004).

There are many ways to normalise the power spectrum, and the usual representation is in terms of $\sigma_8 := \sigma(R = 8\,h^{-1}\mathrm{Mpc})$. The choice is motivated by the fact that the variance of number counts of galaxies $\Delta N/N$ in an $8\,h^{-1}$Mpc sphere is of order unity. The standard deviation of matter fluctuation at the scale R is defined as

$$\sigma(R,t) = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k\; P(k,t)\,|\hat{W}(kR)|^2\,k^2\;, \tag{2.43}$$

where $\hat{W}(kR)$ is the Fourier transform of the real-space top-hat window function of radius R

$$\hat{W}(kR) = 3\frac{\sin(kR) - kR\,\cos(kR)}{(kR)^3}\,. \tag{2.44}$$

The power spectrum $P(k,t)$ is multiplied by $\hat{W}(kR)$ which allows for the selection of specific scales of interest. Finally, combining (2.42) and (2.43) leads to

$$\sigma_8^2 = \frac{A}{2\pi^2}\int_0^\infty \mathrm{d}k\; k^{2+n_\mathrm{s}}\,T^2(k)\,|\hat{W}(kR)|^2\,. \tag{2.45}$$

Thus, the measurement of σ_8 at the current epoch allows us to empirically constrain the matter power spectrum amplitude A. σ_8 has become one of the key parameters to obtain from cosmological surveys, the *Planck* mission (2018b) reports their best constraint as $\sigma_8 = 0.8102 \pm 0.0060$.

2.4.2. Clustering in the Universe

The Universe is rich in structures that are constantly changing. However, the two components of matter, dark and luminous matter, show a different clustering behaviour with scale and time. Linear perturbation theory fails badly when the density contrast starts to approach unity and baryon physics on small scales cannot be predicted by linear theory in general. Nevertheless, non-linear density structures are of special interest since they are the seeds for galaxies and clusters. For small scales, an accurate prediction must involve higher-order perturbation theory, semi-analytic modelling of selected features, or time-consuming numerical simulations (which are discussed in detail in Sect. 4.1). Each option has advantages and disadvantages and is limited to certain scales only.

Semi-analytic spherical collapse model

In general, semi-analytical models are interesting since they give more insight into the physical processes involved. The spherical collapse model has been a particularly successful, but simple model in describing the density and size of dark matter halos. This model is based on analytically solving the evolution of a spherical overdensity in an averagely-dense Universe. The expansion rate of the small sphere is decreased compared to the rest of the Universe, which increases the density contrast further over time. As soon as the density inside the sphere becomes reaches a threshold value, the expansion comes to a halt and the sphere starts to contract. Since the particles inside the sphere are not expected to follow an exact symmetry, the sphere does not collapse to a single point, but to an overdense, virialised sphere with a radius r_vir that is half its maximum radius during expansion. We call such a collapsed mass concentration inside a spherical region a dark matter halo. According to the

spherical collapse model, its average density is ~ 200 times the mean density of the Universe.

The model is particularly powerful since the non-linear density evolution can be compared to the linear one. A virialised halo in non-linear evolution translates to a spherical overdensity of $\delta_{\rm lin} = 1.69$ in the linear model, where the simplified cosmological model of an Einstein-de-Sitter (EdS) universe with $\Omega_{\rm m} = \Omega_0 = 1$ was used. In this frame, any density perturbation, that had a linearly extrapolated density contrast of $1.69(1+z)$ for any given redshift z, is a collapsed and virialised structure of density $\sim 200\bar{\rho}(z)$. Therefore, dark matter halos at a given mass are more compact in the past than those formed today.

Dark matter halo mass function and profile

Using the spherical collapse model, it is possible to estimate the number density of collapsed halos as a function of time and halo mass. The halo mass function can then be used to compare the predicted abundance of massive halos to the observed one, where usually galaxy cluster masses are used in practice. Bill Press and Paul Schechter (1974) were the first to derive a halo mass function from an initial Gaussian density field, and their model became well known as the Press-Schechter theory. If $\delta_{\rm c}(t, t_{\rm i})$ is the density contrast at $t_{\rm i}$ that is necessary for a collapse at time t, the Press-Schechter model can be written as

$$\frac{{\rm d}n}{{\rm d}M}(M,t) = \frac{\rho_{\rm m,0}}{M}\left(-\frac{{\rm d}\ln\sigma}{{\rm d}M}\right) f\left(\frac{\delta_{\rm c}}{\sigma}\right) \,, \tag{2.46}$$

where f is

$$f\left(\frac{\delta_{\rm c}}{\sigma}\right) = \sqrt{\frac{2}{\pi}}\,\frac{\delta_{\rm c}}{\sigma}\,\exp\left(\frac{\delta_{\rm c}^2}{2\sigma^2}\right) \,. \tag{2.47}$$

An application of this model to data yields constraints on the normalization of the power spectrum σ_8 in the form $\sigma_8 \propto 0.6\,\Omega_{\rm m}^{0.56}$.

The Press-Schechter model traces the abundance of dark matter halos reasonably well. However, it slightly underestimates the abundance of very massive halos while overestimating the abundance of low-mass halos. Using numerical simulations, Jenkins et al. (2001) found an approximate fitting formula that is more accurate than the Press-Schechter theory

$$f(\sigma) = 0.315\exp\left(-\left|\ln\sigma^{-1} + 0.61\right|^{3.8}\right) \,. \tag{2.48}$$

Moreover, Julio Navarro, Carlos Frenk and Simon White (1996, 1997) showed that any ensemble of dark matter halos can be described by a characteristic radius $r_{\rm s}$, a characteristic density $\rho_{\rm s} = 4\rho(r_{\rm s})$ and a slope $\alpha \approx 1$

$$\rho(r) = \frac{\rho_{\rm s}}{\left(\frac{r}{r_{\rm s}}\right)^{\alpha}\left(1+\frac{r}{r_{\rm s}}\right)^{3-\alpha}} \,, \tag{2.49}$$

which is generally known as the NFW model.

Galaxy clustering

So far, we were mainly interested in the evolution of dark matter structures and we identified some observables that we can use to determine the underlying cosmology. However, we cannot observe dark matter directly and we must rely on the observable luminous matter in galaxies. Unfortunately, baryonic matter shows a variety of interactions besides gravity which massively complicate the analytical treatment of galaxy clustering and numerical simulations struggle to give results in good agreement with observations.

As a first approximation one can describe the galaxy correlation function as a power law

$$\xi_{\mathrm{gal}}(r) = \left(\frac{r}{r_0}\right)^{-\gamma} , \tag{2.50}$$

where the slope is $\gamma \simeq 1.7$ and the clustering length r_0 is typically of the order of a few megaparsec. As a general rule, the larger the halo the galaxy resides in, the larger is the clustering length r_0.

A more sophisticated model describing non-linear gravitational clustering, which is in excellent agreement with the observational data, is the halo model (see, e.g., Cooray and Sheth, 2002). This semi-analytic model assumes that all matter in the Universe is located in halos of different size and mass. Therefore, matter particles that live in the same halo are correlated. However, since the distribution of halos is correlated, also matter particles that live in different halos show a correlation that dominates on scales larger than the typical halo size. Thus, in the halo model, the correlation function (and also the power spectrum) is written as the sum of two terms, which we call the 1-halo and the 2-halo term, i.e. $\xi(r) = \xi_{1\mathrm{h}}(r) + \xi_{2\mathrm{h}}(r)$. We can use, e.g., the Press-Schechter theory to make assumptions on the distribution of halo masses and describe the spatial distribution of halos using the matter power spectrum. Alternatively, the halo model can be applied to results from dark matter-only N-body simulations (which are introduced in Sect. 4.1.1) to populate dark matter halos. Depending on the halo mass, a halo can be populated by more than one central galaxy; these galaxies are called satellite galaxies. The halo model is mainly limited by the assumption that the relation of a galaxy and a dark matter halo depends only on the halo mass. Deviations from this behaviour are called assembly bias and hold information about the formation of the halo.

Galaxy bias

The relation between the distribution of dark matter and galaxies is by no means simple and we parametrise our ignorance with the galaxy bias b. In general, the bias parameter depends on many variables such as galaxy type, scale and time; for

a single galaxy, the bias parameter depends on its evolution and surroundings. It is, thus, a fairly complicated function, but it can be linearised on large scales. For linear perturbation theory, the relative number fluctuation of galaxies is related to the underlying density contrast by a linear bias factor b

$$\frac{\Delta n}{n} = b\frac{\Delta \rho}{\rho} \,, \tag{2.51}$$

$$P_{\mathrm{gal}}(\boldsymbol{k}) = b^2 P(\boldsymbol{k}) \,. \tag{2.52}$$

Hence, the bias factor b is closely related to the normalization of the power spectrum, i.e. $\sigma_8 \approx 1/b$. Typical values of the bias factor are $1 \lesssim b \lesssim 2$ depending on scale and cosmic time. Furthermore, a dependence on galaxy type or luminosity is to be expected since early-type galaxies are usually found in galaxy clusters which form inside the heaviest underlying dark matter halos, while late-type galaxies are typically field-galaxies, which can be associated with lower-mass halos.

CHAPTER 3

WEAK GRAVITATIONAL LENSING

In the past 40 years, gravitational lensing has become an invaluable tool to study the Universe. Strong and weak lensing applications lead to precise mass measurements of galaxies and galaxy clusters, robust estimates of various cosmological parameters, and the detection of exoplanets in our galaxy. This thesis focuses on weak gravitational lensing which was first detected in 1990 by Tyson, Valdes and Wenk and has progressed tremendously since then. In the following, I provide a concise introduction of the theory of weak lensing, especially GGL and some applications. For a detailed review, the reader is referred to Bartelmann and Schneider (2001) and Dodelson (2017).

3.1. Basics of gravitational lensing

Light deflection by gravitational fields is one of the main predictions of GR. For a gravitational field that is created by a point mass M, and a light ray from a distant galaxy passing the point mass at a distance ξ, we expect a deflection of the light ray by an angle

$$\hat{\alpha} = \frac{4GM}{c^2\xi} \,, \tag{3.1}$$

which differs from the Newtonian prediction by a factor of two. During the Solar eclipse of 1919, it was possible to measure the apparent positions of stars near the Sun. A comparison between apparent and expected position confirmed Einstein's prediction making it the valid theory of gravity since 100 years.

3.1.1. The lens equation

Gravitational lens systems can be described with geometrical relations as visualised in Fig. 3.1. As a light ray travels from the source galaxy at position $\boldsymbol{\eta}$ to the observer,

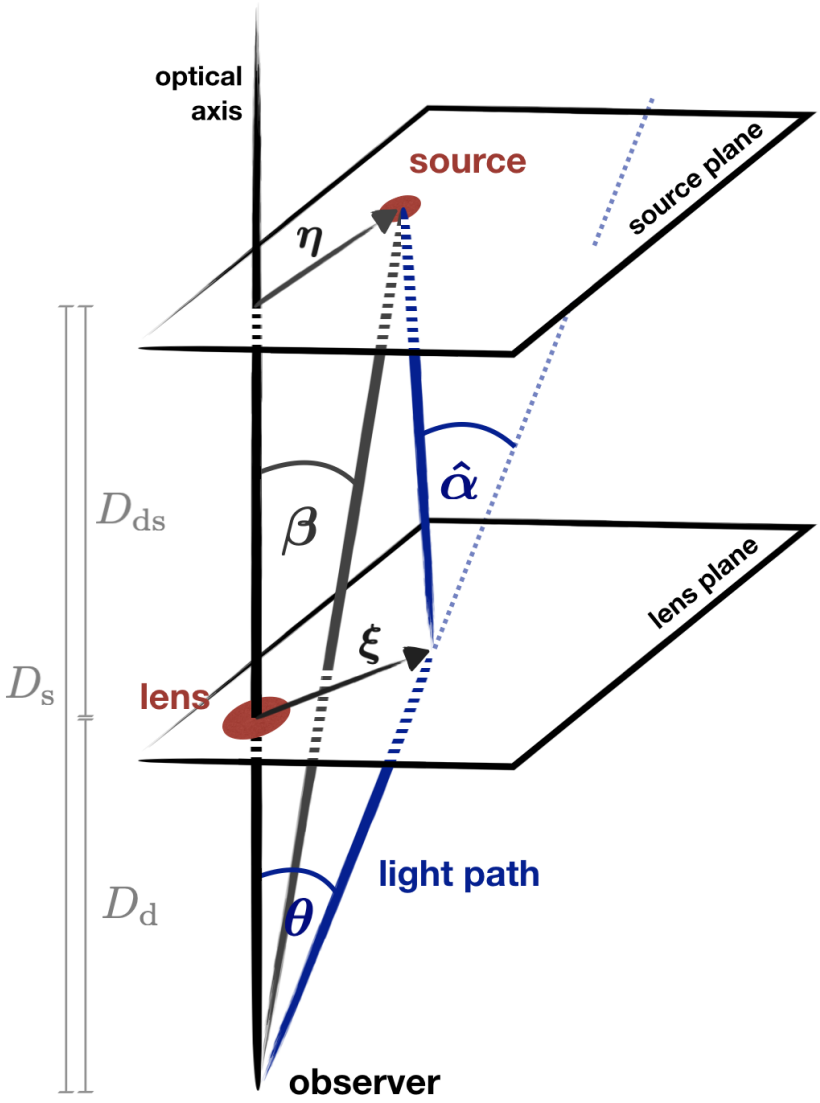


Figure 3.1. – Geometry of a typical lens system where all vectors and angles are two-dimensional quantities and all distances are angular diameter distances.

it passes the gravitational field of the lens galaxy at distance $\boldsymbol{\xi}$ and gets deflected. The relevant angular diameter distances are from us to the source $D(0, z_\mathrm{s}) := D_\mathrm{s}$, from us to the lens (also called the deflector) $D(0, z_\mathrm{d}) := D_\mathrm{d}$, and from the lens to the source $D(z_\mathrm{d}, z_\mathrm{s}) := D_\mathrm{ds}$. In the case that the extent of the lens is much smaller than the distances involved, we can approximate the light curve by two straight lines with

a kink in the lens plane. The kink's magnitude and orientation is described by the deflection angle $\hat{\boldsymbol{\alpha}}$. We can now apply the small angle approximation and express the two-dimensional source position with $\boldsymbol{\eta} = D_\mathrm{s}\boldsymbol{\beta}$ and the impact vector with $\boldsymbol{\xi} = D_\mathrm{d}\boldsymbol{\theta}$, where $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are the angular positions under which the source would appear and the actual image does appear, respectively. Then, we can geometrically relate the true source position $\boldsymbol{\beta}$ to the observed angular position $\boldsymbol{\theta}$ using the deflection angle $\hat{\boldsymbol{\alpha}}$

$$\boldsymbol{\beta}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\boldsymbol{\alpha}}(D_\mathrm{d}\boldsymbol{\theta}) := \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \,, \tag{3.2}$$

where we define the scaled deflection angle $\boldsymbol{\alpha} = D_\mathrm{ds}/D_\mathrm{s}\,\hat{\boldsymbol{\alpha}}$ in the second equation.

The lens equation (3.2) relates the apparent position of a source galaxy to its true position. For a given source position $\boldsymbol{\beta}$ several observed positions $\boldsymbol{\theta}_i$ can satisfy Eq. (3.2). In this case, we observe the same object at multiple positions. In the special case of the deflection around a point mass, as given in (3.1), as well as a source being on the optical axis, we can calculate the solution for $\boldsymbol{\theta}_i$ explicitly. Due to spherical symmetry, we reduce the lens equation to one dimension and obtain

$$\beta = \theta - \frac{D_\mathrm{ds}}{D_\mathrm{d}D_\mathrm{s}}\frac{4GM}{c^2\theta} \,, \tag{3.3}$$

which has the solution

$$\theta_\mathrm{E} = \sqrt{\frac{D_\mathrm{ds}}{D_\mathrm{d}D_\mathrm{s}}\frac{4GM}{c^2}} \,. \tag{3.4}$$

The solution is known as the Einstein angle θ_E and the source galaxy appears as an (Einstein) ring around the lens galaxy. If the distances are known, we can directly infer the unbiased mass of the lens enclosed in the Einstein ring.

3.1.2. Deflection by extended density distributions

If the lens galaxy is extended in three dimensions, the deflection angle can be expressed as the vectorial sum of the deflections from infinitesimal mass elements. A light ray that passes the galaxy has a spatial trajectory described by $\big(\xi_1(\lambda), \xi_2(\lambda), r_3(\lambda)\big)$, where λ is the affine parameter and we chose the coordinates such that the incoming light ray propagates along r_3. The lens deflects the light ray, but if the deflection angle is small we can approximate the light ray as a straight line in the neighbourhood of the deflecting mass, and we can approximate $\boldsymbol{\xi}(\lambda) \approx \boldsymbol{\xi}$ near the lens. This is referred to as the Born approximation and a mass distribution that satisfies this condition is called geometrically-thin.

Since the perception of depth for objects at large distances is not available, it is convenient to define the surface mass density perpendicular to the incoming light

ray

$$\Sigma(\boldsymbol{\xi}) = \int \mathrm{d}r_3 \; \rho(\xi_1, \xi_2, r_3) \; . \tag{3.5}$$

Using the surface mass density, we can express the scaled deflection angle for an arbitrary density distribution as

$$\boldsymbol{\alpha}(\boldsymbol{\xi}) = \frac{4G}{c^2} \frac{D_{\mathrm{ds}}}{D_{\mathrm{s}}} \int \mathrm{d}^2\xi' \; \Sigma(\boldsymbol{\xi}') \, \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} \; . \tag{3.6}$$

Alternatively, we can introduce the dimensionless surface mass density or convergence

$$\kappa(\boldsymbol{\theta}) = \frac{4\pi G}{c^2} \frac{D_{\mathrm{d}} D_{\mathrm{ds}}}{D_{\mathrm{s}}} \Sigma(D_{\mathrm{d}}\boldsymbol{\theta}) \tag{3.7}$$

$$:= \frac{\Sigma(D_{\mathrm{d}}\boldsymbol{\theta})}{\Sigma_{\mathrm{crit}}} \; , \tag{3.8}$$

and the deflection angle becomes

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int \mathrm{d}^2\theta' \; \kappa(\boldsymbol{\theta}') \, \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \; . \tag{3.9}$$

In Eq. (3.8) we defined a characteristic value of lensing systems – the critical surface mass density Σ_{crit}. A sufficient but not necessary condition for a mass distribution to have multiple images is that the surface mass density at at least one point is larger than the critical surface mass density, i.e. $\kappa > 1$. Thus, the ratio of Σ to Σ_{crit} is a quantity that lets us divide lensing systems into 'strong' and 'weak' lenses.

The deflection angle has the property to be curl-free, i.e. it has no rotational modes. It can, therefore, be expressed as a gradient of a deflection potential ψ

$$\boldsymbol{\alpha} = \nabla\psi(\boldsymbol{\theta}) \; , \tag{3.10}$$

where the deflection potential can also be expressed in terms of the dimensionless surface mass density

$$\nabla\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int \mathrm{d}^2\theta' \; \kappa(\boldsymbol{\theta}) \, \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \; . \tag{3.11}$$

These two quantities are therefore related by

$$\nabla^2\psi = 2\kappa \; . \tag{3.12}$$

3.1.3. Image distortions

Gravitational lensing changes the path of light as it travels through the Universe. Since source galaxies are extended objects on the sky, we do not expect the image shape to be the same as the original shape. Still, if the objects are small compared to the scale on which the lens mapping changes considerably, we can describe the image distortion using only a few parameters.

Shear

We investigate the image distortion by looking at the Jacobi matrix of the lens equation (3.2). The components of the Jacobian are defined as $\mathcal{A}_{ij} = \partial\beta_i/\partial\theta_j$ and can be expressed in the following form[1]

$$\mathcal{A}(\boldsymbol{\theta}) = \left(\delta_{ij} - \frac{\partial^2\psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) := \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} , \tag{3.13}$$

where δ_{ij} is the Kronecker delta and we introduced the two components of the shear $\gamma = \gamma_1 + \mathrm{i}\gamma_2 = |\gamma|\mathrm{e}^{2\mathrm{i}\phi}$. Shear quantifies the shape distortion of an image, e.g., an ellipticity is induced in an intrinsically round object due to shear as is illustrated in Fig. 3.2. The convergence, on the other hand, only causes an isotropic scaling of the image.

By slightly rewriting the Jacobian, the isotropic and anisotropic stretch in the image plane can be separated into two terms

$$\mathcal{A}(\boldsymbol{\theta}) = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} , \tag{3.14}$$

where we define the reduced shear $g = \gamma/(1-\kappa)$, which is a complex number as the shear γ. The reduced shear describes the anisotropic distortion of the image shape between the observed image and source galaxy; it is the central measurable quantity in weak gravitational lensing, where only small distortions are considered.

Magnification

According to Liouville's theorem lensing conserves the surface brightness. Due to the shape distortion, however, the observed apparent solid angle ω of the image differs from the one in the absence of lensing, ω_0. Hence, also the flux s that we receive, is enhanced or reduced compared to the unlensed flux s_0. The magnification is defined as

$$\mu(\boldsymbol{\theta}) = \frac{\omega(\boldsymbol{\theta})}{\omega_0} = \frac{s(\boldsymbol{\theta})}{s_0} = \frac{1}{\det\mathcal{A}(\boldsymbol{\theta})} = \frac{1}{[1-\kappa(\boldsymbol{\theta})]^2 - |\gamma(\boldsymbol{\theta})|^2} . \tag{3.15}$$

In a flux-limited survey, the local number density of galaxies is changed due to a rescaling of solid angles by $1/\mu$ and an enhancement or reduction of observed flux by μ. The observed cumulative number density $n(>s)$ is therefore

$$n(>s) = \frac{1}{\mu}\, n_0\left(>\frac{s}{\mu}\right) , \tag{3.16}$$

[1] This definition is correct under the Born approximation and neglecting lens-lens coupling. These two approximations account for the fact that higher redshift mass inhomogeneities slightly distort lower redshift matter distributions as experienced by a light ray on its way to us.

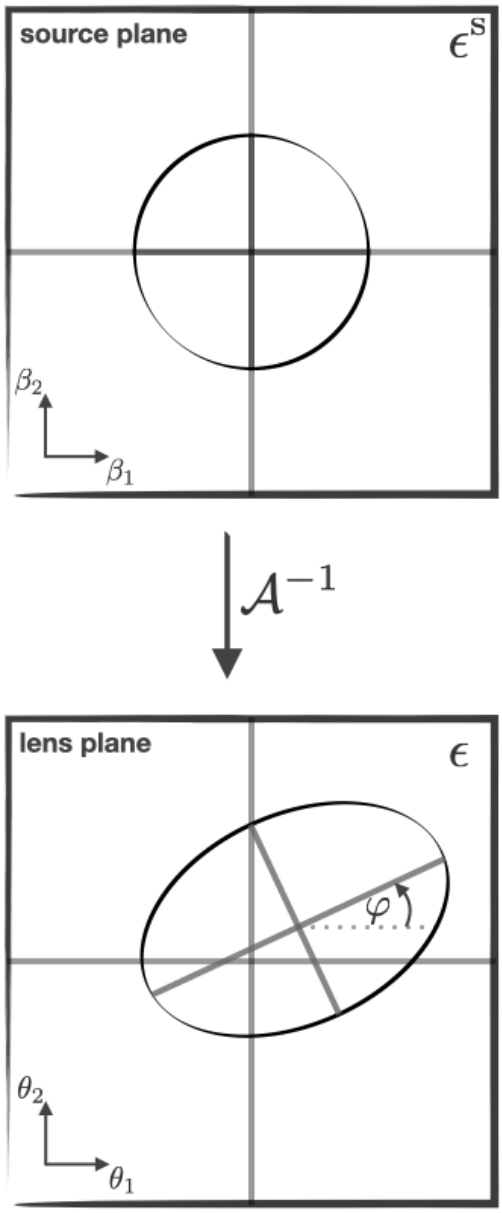


Figure 3.2. – Starting from a round object in the source plane, a mapping $\mathcal{A}^{-1}$ is applied and the object becomes displaced, rotated and elliptical in the weak lensing regime. The area of the object is also changed, which changes the incoming flux described by the magnification μ. The final shape of the object is determined by convergence κ and shear γ.

where n_0 denotes the cumulative number density in the absence of lensing. Whether galaxy counts on the sky are enhanced or reduced depends on the form of n_0 close to the flux limit.

3.2. Gravitational lens models

In general, the lens equation (3.2) must be inverted numerically. In case of an unknown number of multiple images, this is not a trivial task. Thus, simple mass models that allow for an analytical treatment of (3.2) are of special interest. In the following, three mass models are presented that are of importance in this theses.

SIS

A singular isothermal sphere (SIS) is a frequently used gravitational lens model. This model is axi-symmetric and the corresponding spherical mass distribution yields flat rotation curves, as they are observed in late-type galaxies. The density distribution $\rho(r) = \sigma_v^2/(2\pi G r^2)$ depends on the velocity dispersion σ_v and diverges for small (proper) radii. Moreover, the mass $M = 4\pi \int \mathrm{d}r\, r^2 \rho(r)$ is not finite for $r \to \infty$.

Although an SIS has several limitations and is certainly not a good model of gravitational lenses for all scales, it is often used to obtain first insights due to its simplicity. The Einstein radius becomes

$$\theta_\mathrm{E} = 0\farcs576 \left(\frac{\sigma_v}{200\,\mathrm{km/s}}\right)^2 \left(\frac{2D_\mathrm{ds}}{D_\mathrm{s}}\right) , \tag{3.17}$$

where $\sigma_v = 200\,\mathrm{km/s}$ is the typical velocity dispersion of a galaxy. More useful lensing quantities, including the mean convergence $\bar{\kappa}$ inside a circle with radius θ, are

$$\begin{aligned} \kappa(\theta) &= \frac{\theta_\mathrm{E}}{2|\theta|} ; \qquad & \bar{\kappa}(<\theta) &= \frac{\theta_\mathrm{E}}{|\theta|} ; \\ |\gamma|(\theta) &= \frac{\theta_\mathrm{E}}{2|\theta|} ; \qquad & \alpha(\theta) &= \theta_\mathrm{E}\frac{\theta}{|\theta|} . \end{aligned} \tag{3.18}$$

NIS

A non-singular isothermal sphere (NIS) introduces a finite core with size θ_c to remove the divergence of density in the lens centre which makes it a viable alternative to the SIS for numerical considerations. We then find

$$\begin{aligned} \kappa(\theta) &= \frac{\sqrt{\theta_\mathrm{E}^2 + \theta_\mathrm{c}^2}}{2} \frac{2\theta^2 + \theta_\mathrm{c}^2}{(\theta^2 + \theta_\mathrm{c}^2)^{\frac{3}{2}}} , \\ \bar{\kappa}(<\theta) &= \sqrt{\frac{\theta_\mathrm{E}^2 + \theta_\mathrm{c}^2}{\theta^2 + \theta_\mathrm{c}^2}} , \end{aligned} \tag{3.19}$$

where the absolute value of the shear is given by

$$|\gamma|(\theta) = \bar{\kappa}(\theta) - \kappa(\theta) \ . \tag{3.20}$$

NFW

The NFW density profile holds well over nine orders of magnitude in radius and was first introduced in Sect. 2.4.2. However, the analytical solution for the surface mass density and shear is rather complicated. As before r_s is the characteristic scale and ρ_s the characteristic density. The surface mass density reads (Bartelmann, 1996)

$$\Sigma_\mathrm{NFW}(r) = 2r_\mathrm{s}\rho_\mathrm{s}\, g\left(\frac{r}{r_\mathrm{s}}\right) \ , \tag{3.21}$$

where g is a function depending on the dimensionless radius $x = r/r_\mathrm{s}$

$$g(x) = \begin{cases} \frac{1}{x^2-1}\left(1 - \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh}\sqrt{\frac{1-x}{1+x}}\right) & \text{for } x < 1 \\ \frac{1}{3} & \text{for } x = 1 \ . \\ \frac{1}{x^2-1}\left(1 - \frac{2}{\sqrt{x^2-1}} \arctan\sqrt{\frac{x-1}{1+x}}\right) & \text{for } x > 1 \end{cases} \tag{3.22}$$

The shear is then denoted by (Wright and Brainerd, 1999)

$$\gamma_\mathrm{NFW}(r) = \frac{r_\mathrm{s}\rho_\mathrm{s}}{\Sigma_\mathrm{crit}}\, h\left(\frac{r}{r_\mathrm{s}}\right) \ , \tag{3.23}$$

where h is

$$h(x) = \begin{cases} \frac{8\operatorname{arctanh}\sqrt{\frac{1-x}{1+x}}}{x^2\sqrt{1-x^2}} + \frac{4}{x^2}\ln\frac{x}{2} - \frac{2}{x^2-1} + \frac{4\operatorname{arctanh}\sqrt{\frac{1-x}{1+x}}}{(x^2-1)\sqrt{1-x^2}} & \text{for } x < 1 \\ \frac{10}{3} - 4\ln 2 & \text{for } x = 1 \ . \\ \frac{8\arctan\sqrt{\frac{x-1}{1+x}}}{x^2\sqrt{x^2-1}} + \frac{4}{x^2}\ln\frac{x}{2} - \frac{2}{x^2-1} + \frac{4\arctan\sqrt{\frac{x-1}{1+x}}}{(x^2-1)^{\frac{3}{2}}} & \text{for } x > 1 \end{cases} \tag{3.24}$$

3.3. Galaxy-galaxy lensing

In the following, we only consider weak gravitational lensing where small effects on background galaxies are induced, i.e. $\kappa \ll 1$ and $|\gamma| \ll 1$. We cannot detect weak lensing effects for a single object. One of the reasons for this is that we cannot observe an unlensed sky. Moreover, weak distortions are typically of the order of 1% compared to the intrinsic, randomly oriented galaxy ellipticities. However, we can make use of the cosmological principle that the Universe is homogeneous and isotropic. It ensures that, if we average over a sufficiently large number of galaxies, the average ellipticity is zero, i.e. $\langle\epsilon_\mathrm{int}\rangle = 0$; in other words, no direction in the

Universe is singled out. The addition of shear to the intrinsic ellipticity is a non-linear relation in general, but can be linearised in the presence of weak reduced shear. Then, the observed ellipticity averaged over many galaxies,

$$\langle \epsilon_{\rm obs} \rangle \approx \langle \epsilon_{\rm int} \rangle + \langle g \rangle = \langle g \rangle \approx \langle \gamma \rangle \ , \tag{3.25}$$

is actually an unbiased estimator of the average reduced shear and, therefore, at least approximately for the averaged shear. However, it is a noisy estimator since the intrinsic ellipticity dominates the measured shape. This so-called shape noise is the main contributor to noise in most weak lensing measurements (e.g., Niemi et al., 2015). The relative contribution of shape noise decreases with an increasing number of observed galaxies, which is why wide and deep surveys are favoured by lensing enthusiasts. Taking deep data refers to measuring ever fainter objects that are located at larger distances from us, and taking wide data means obtaining preferably full-sky galaxy and shear maps.

GGL probes the average surface mass overdensity around galaxies by cross-correlating the positions of foreground galaxies (lenses) to the shear of background galaxies (sources)

$$\langle \gamma_{\rm t} \rangle(\theta) = \langle \kappa_{\rm g}(\boldsymbol{\theta'}) \gamma_{\rm t}(\boldsymbol{\theta'} + \boldsymbol{\theta}; \boldsymbol{\theta}) \rangle \ , \tag{3.26}$$

where the signal only depends on the absolute value of the projected separation $|\boldsymbol{\theta}| = \theta$, and $\kappa_{\rm g}$ is the projected fractional number-density contrast of lens galaxies. The tangential shear is denoted by $\gamma_{\rm t}$ and measured with respect to the connecting line between lens and source, orthogonal to that line is the tangential shear $\gamma_{\rm t}$, and the cross shear $\gamma_{\rm x}$ is measured with a 45°-rotation (cmp. Fig. 3.3). For a fixed lens position $\boldsymbol{\theta}_{\rm d}$ this corresponds to a rotation of the shear components

$$\gamma_{\rm t}(\boldsymbol{\theta}) + {\rm i}\gamma_{\rm x}(\boldsymbol{\theta}) = -\gamma(\boldsymbol{\theta}) \, \frac{(\boldsymbol{\theta}_{\rm d} - \boldsymbol{\theta})^*}{\boldsymbol{\theta}_{\rm d} - \boldsymbol{\theta}} \ , \tag{3.27}$$

where we also conveniently write the position on the sky in complex notation, i.e. $\boldsymbol{\theta} = \theta_1 + {\rm i}\theta_2$, and an asterisk denotes complex conjugation.

Astronomical objects only induce a tangential shear on background galaxies. The average cross shear $\langle \gamma_{\rm x} \rangle$ must vanish due to parity invariance (Schneider, 2003) and detection of a non-negligible cross shear shows the presence of systematic errors in the data. The average tangential shear $\langle \gamma_{\rm t} \rangle(\theta)$ is related to the azimuthally averaged convergence $\langle \kappa \rangle(\theta)$ and the mean convergence inside a circle with radius θ by

$$\langle \gamma_{\rm t} \rangle(\theta) = \bar{\kappa}(< \theta) - \langle \kappa \rangle(\theta) \ . \tag{3.28}$$

In the case that the redshifts of sources and lenses are known, the most common observable is the excess surface mass density

$$\Delta\Sigma(R) = \bar{\Sigma}(< R) - \Sigma(R) = \gamma_{\rm t}(R) \Sigma_{\rm crit} \ , \tag{3.29}$$

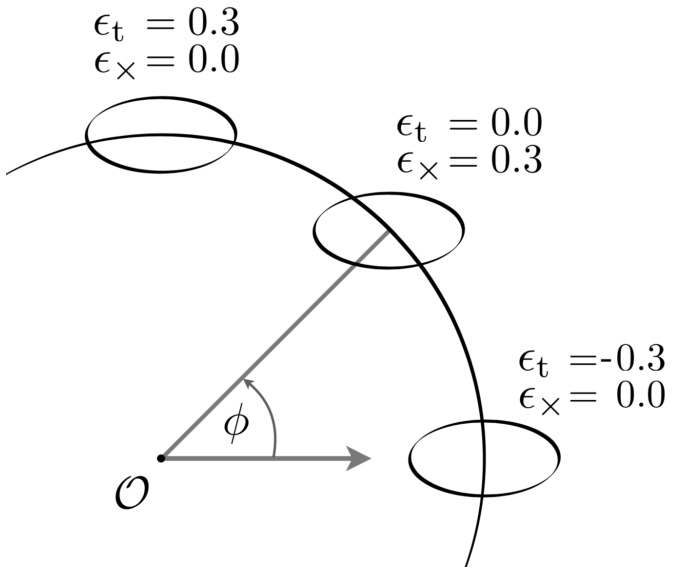


Figure 3.3. – The difference between tangential ϵ_t and cross ellipticity $\epsilon_\times$ is shown. A tangentially and a radially aligned galaxy do not show any cross ellipticity but have different signs of tangential ellipticity. A galaxy that is rotated by $\phi = 45°$ with respect to the circle's tangent does not show any tangential ellipticity.

where R is the physical or comoving distance from the lens centre. In practice, the average tangential shear $\langle\gamma_t\rangle$ is further averaged over many lenses to obtain a high signal-to-noise ratio. The mean tangential shear or excess surface mass density is then analysed as a function of distance from the lens centre and compared to known gravitational lens models as given in the previous section.

3.4. Weak lensing observations

In the last 30 years, the methodology and observation techniques of weak lensing have advanced continuously. Currently, two weak lensing surveys are ongoing: DES and HSC SSP; and KiDS has only recently finished its observations. In the near future, LSST will also provide measurements from the ground, while the space-based missions *Euclid* and WFIRST will obtain measurements untainted by the Earth's atmosphere.

Weak lensing measurements are dominated by noise and the key to successful surveys is a large sample of galaxies, which can be obtained by taking deep and wide data. However, an improved statistical error on measurements has to be accompanied by sufficiently controlled astrophysical and detector-related systematic errors. The lensing community is currently putting huge effort into understanding systematics that affect and will affect weak lensing measurements as well as their implications. The interested reader is kindly referred to Mandelbaum (2018) for an

extensive review regarding sources of systematic errors in the pipeline from galaxy images to scientific results. In the following, I provide an outline of the current challenges in shear measurements and analyses.

telescope systematics In an imaging system, a point-like object is blurred to an extended shape which is described by the point spread function (PSF). The observed brightness profile $I_{\rm obs}$ is related to the true surface brightness I of an image by a convolution with the PSF P

$$I_{\rm obs}(\boldsymbol{\theta}) = \int \mathrm{d}\vartheta^2 I(\boldsymbol{\theta} - \boldsymbol{\vartheta})\, P(\boldsymbol{\vartheta}) \ . \tag{3.30}$$

The atmosphere is a major contributor to the wavelength-dependent blurring effect, which is one of the reasons astronomers are not fond of it. The PSF is modelled using bright stars that are assumed to undergo the same telescope-induced changes as the galaxies and is then carefully interpolated to the galaxy positions. More sources of systematics induced by the detector are the pixel response to incoming light (e.g., non-linear response and the brighter-fatter effect; Guyonnet et al., 2015; Plazas et al., 2016) and charge diffusion between pixels (e.g., charge transfer inefficiency; Kannawadi et al., 2016).

galaxy detection In a simplified picture, galaxies are detected when the integrated flux is brighter than the survey magnitude limit. However, the fainter the magnitudes we observe, the more densely the sky is populated. This leads to overlapping galaxy profiles which is referred to as blending. If multiple objects are detected, the individual galaxy properties, e.g., their redshift, must be analysed very carefully (Dawson et al., 2016). Unresolved galaxies, however, that are just below the detection limit, pose an even bigger challenge as they affect the measured galaxy properties and lead to correlated pixel-noise (Gurvich and Mandelbaum, 2016).

shear estimation The galaxy ellipticities $\langle\epsilon\rangle$ are measured as the second moments of galaxy images

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta\ I(\boldsymbol{\theta}) W(\boldsymbol{\theta})\, \theta_i \theta_j}{\int \mathrm{d}^2\theta\ I(\boldsymbol{\theta}) W(\boldsymbol{\theta})} \ , \tag{3.31}$$

where $\theta_{1,2}$ correspond to Cartesian coordinates. Their origin is the centre of light within a limiting isophote of the image, and $W(\boldsymbol{\theta})$ is a weighting function. The trace of Q contains the information about the size of the image, while the traceless part describes the ellipticity information. The true and the observed second-order brightness moments are related by $Q_{{\rm obs},ij} = Q_{ij} + P_{ij}$, where P_{ij} are the second-order moments of the PSF and the weighting function is chosen to be constant.

Commonly, the ellipticity is defined in terms of the moments as

$$e = e_1 + \mathrm{i}e_2 = \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22}} \,, \tag{3.32}$$

or alternatively as

$$\varepsilon = \varepsilon_1 + \mathrm{i}\varepsilon_2 = \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22} + 2\,|\det Q\,|^{1/2}} \,. \tag{3.33}$$

These definitions have a well-defined response to the lensing shear which allows for averaging across ensembles of galaxies. In the presence of a PSF the observed ellipticity is related to the true ellipticity by

$$e_\mathrm{obs} = \frac{e + T\, e_\mathrm{PSF}}{1 + T} \,, \tag{3.34}$$

with

$$T = \frac{P_{11} + P_{22}}{Q_{11} + Q_{22}} \quad \text{and} \quad e_\mathrm{PSF} = \frac{P_{11} - P_{22} + 2\mathrm{i}P_{12}}{P_{11} + P_{22}} \,, \tag{3.35}$$

where similar equations can be found for ε. These rather simple relations are no longer valid for a non-constant weighting function W, which also complicates the relation (3.25) between the observed ellipticity and true shear.

Uncorrected (residual) telescope systematics may also lead to biased results. This bias is typically assumed to be small but relevant and can be split into a term that scales with the ensemble shear statistics, the multiplicative bias, and a term that is independent of it, the additive bias. Computationally expensive image simulations, null-tests and self-calibration methods exist and are developed to sufficiently control the bias.

redshift estimation To infer cosmological parameters or the galaxy-dark matter halo connection, we must further obtain accurate information about the redshift of the galaxy sample. Obtaining an exact redshift estimate with spectroscopic methods (spec-z) is only feasible for foreground galaxy samples. Background galaxies are too numerous and too faint for spectroscopic analyses. For these galaxies, photometric redshift (photo-z) estimates are obtained. The photo-z estimates are typically done using template-fitting of galaxy fluxes or machine-learning algorithms (Hildebrandt et al., 2010; Wright et al., 2020). The goal (and challenge) is to obtain the true posterior probability function of the whole galaxy sample.

theoretical predictions Weak lensing observations are compared to theoretical predictions to find the best-fitting models. Thus, we are challenged with the need for an accurate description of the non-linear matter power spectrum and

the non-linear galaxy bias. The validity of approximations (as the flat-sky, the Born, and the Limber approximation as well as the impact of higher-order lensing terms) is currently being investigated. Furthermore, analytic theoretical predictions for cosmological studies often disregard the complex baryon physics. Therefore, semi-analytic models or simulations for baryonic effects (Schneider et al., 2020; Chua et al., 2019) and the intrinsic alignment of galaxies (Troxel and Ishak, 2015) are needed. Lastly, inferring astrophysical parameters is based on accurate covariance estimations that still require a computationally expensive numerical estimation (Sellentin and Heavens, 2017).

CHAPTER 4

SHEAR ESTIMATES IN A SIMULATED UNIVERSE

Simulations of the Universe are the astronomer's playground. While having complete control over the input parameters, we can study the outcome. The results can be compared or applied to real data while carefully keeping the limitations of the simulations in mind. For studying cosmology and weak gravitational lensing, powerful simulations have become invaluable tools that we explore in this chapter.

4.1. Simulations for cosmology and lensing

To model the smallest scales in the Universe, the use of numerical simulations is unavoidable. In the past three decades, tremendous progress has been achieved. Analyses of the resulting matter structures extend the linear power spectrum to non-linear scales using fitting formulae. In general, numerical results have several restrictions, e.g., they have a minimum and maximum length scale on which predictions are still expected to be accurate. The results are also limited by mass resolution. Finally, they may or may not involve baryons and their complicated interactions (e.g., star formation, gas cooling, feedback from supernovae and active galactic nuclei (AGN), and the chemical evolution of the intergalactic medium). Dark matter-only simulations need considerably less computation time than a similar-sized cube filled with dark matter and baryons. On the other hand, baryons gravitationally influence the distribution of dark matter on the galactic scale, although they make only a small fraction of the matter content. Negligence of baryons results in a suppression of the power spectrum by up to 20% (Semboloni et al., 2011; van Daalen et al., 2014) compared to a pure dark matter universe, where the suppression is a function of scale.

In general, dark matter-only simulations overpredict the structure growth on scales $\lesssim 1\,\mathrm{Mpc}$ to a largely unknown degree. However, baryon physics is not understood in detail, and any implementation of baryons comes with a huge computational cost compared to scenarios where the only acting force is gravity. For these reasons, dark matter-only simulations are still the best tool to cover the largest simulated volumes and capture the Universe's structure growth over time. Since the mass ratio of baryons to dark matter is roughly 1:5, the Universe is indeed well represented by results from dark matter-only, large-scale simulations. Caution has to be applied on scales comparable to the size of a galaxy, where baryon physics becomes important.

Thus, each simulation must be carefully fine-tuned to the research question to gain an optimal output for given computation time. The maximum size and mass resolution of a simulation are, therefore, linked to the ever-growing computational capabilities. In the following, we focus on the tools used in this thesis – dark matter-only N-body simulations, semi-analytic galaxy formation models and ray-tracing simulations.

4.1.1. Cosmological N-body simulations

N-body simulations use a finite set of particles to mimic the distribution of matter. Dark matter is represented by a set of collisionless particles that only interact via gravitational forces. The first attempts to simulate matter structures in the Universe that way were done in the 1970's by Peebles who simulated the structure of the Coma cluster using a set of 300 particles.

To set up a dark matter-only N-body simulation, the comoving length of the simulated cube has to be chosen as well as the number of particles N and the mass per particle M. The size of the simulated volume must be carefully tailored to the science goals since density fluctuations on scales comparable to or larger than the side length of the simulated cube are completely neglected. These density fluctuations are called super-modes and their impact on the matter power spectrum is not yet fully understood. To minimise the impact of super-sample modes, the size of the simulation cube should be at least as large as the scale on which we observe the largest gravitationally bound structures, i.e. $\approx 200\,h^{-1}\mathrm{Mpc}$.

The initial conditions are set such that the distribution of particles at very early times resembles a Gaussian random field. Further, the power spectrum of the particle distribution is matched to the linear power spectrum of a chosen cosmology. Such a linear power spectrum can be calculated analytically. Lastly, the finite volume of the simulation cube needs to be accounted for. Since the Universe is homogeneous on large scales, the outside of the simulation cube cannot be simply treated as empty, which is equivalent to a sharp drop in density. The problem is solved by periodically extending the simulation cube to all sides with the disadvantage of only having reliable results on scales less than half the cube's side length.

When the initial particle positions are set, the movement of the particles per time step must be calculated. The force exerted on the i-th particle for any given time is

$$\boldsymbol{F}_i = GM^2 \sum_{j \neq i}^{N} \frac{\boldsymbol{r}_j - \boldsymbol{r}_i}{|\boldsymbol{r}_j - \boldsymbol{r}_i|^3} \,, \tag{4.1}$$

where $\boldsymbol{r}_j - \boldsymbol{r}_i$ is the distance of two particles i and j. Crudely calculating (4.1) requires N^2 operations per time step. However, this method can be optimised considerably by separating the gravitational potential of a particle into a short- and long-range component. Short-range interactions are then indeed calculated by summing (4.1). Long-range interactions, however, can be calculated by placing the particles on a regular mesh and using a fast Fourier Transform (FFT). The number of cells is chosen to roughly match the number of particles and, in this case, only $N \log N$ operations are needed to calculate the gravitational force of such a field (Hockney and Eastwood, 1988). Such an algorithm is referred to as a particle-particle particle-mesh (P^3M) algorithm. Further, choosing the length of a time step is also critical for the performance of the simulation. Any particle that has close neighbours experiences faster changes in their position and velocity than particles in a sparse neighbourhood. Thus, the length of the time steps is typically varied from particle to particle.

As the change of position due to the gravitational force is calculated for each time step, outputs are generated. Similar to the limited computational power, also the storage capacities are finite and the format of the output has to be chosen well. Large storages are needed if the resulting positions and velocities of all particles for several time steps are saved. Alternatively, positions and masses of dark matter halos can be identified from the particle positions on-the-fly and then saved with a significant reduction in storage space needed. The virial mass $M_{\rm vir}$, the virial radius $R_{\rm vir}$, and the circular virial velocity $V_{\rm vir}(R_{\rm vir})$ are simply related by

$$M_{\rm vir} = \frac{100}{G} H^2(z) R_{\rm vir}^3 = \frac{V_{\rm vir}^3}{10 G\, H(z)} \,. \tag{4.2}$$

As an example, the Friends-of-Friends (FoF) algorithm from Davis et al. (1985) is commonly used to quickly find dark matter halos by linking two particles if their separation is less than 20% of the mean particle separation. As a disadvantage, the FoF algorithm is unable to identify substructures in a dark matter halo. Substructures are needed to resolve the merger history of each dark matter halo which is a crucial requirement for the application of semi-analytical galaxy formation models (SAMs) that are introduced in the next section. A possible tool for subhalo identification is the SUBFIND algorithm (Springel et al., 2001)[1]; it calculates basic (sub)halo properties as, e.g., cumulative density profiles, maximum circular velocities, velocity

[1] A thorough discussion of the performance of halo and subhalo finders can be found in Knebe et al. (2011) and Onions et al. (2012), respectively.

dispersions, as well as halo shapes and orientations. In the following the N-body simulations used in this work are briefly introduced.

Millennium Simulation

The Millennium Simulation (MS) (Springel et al., 2005) is a dark matter-only N-body simulation that traces 2160^3 particles with mass $8.6 \times 10^8\, h^{-1} M_\odot$ from redshift $z = 127$ to today. Consistent with a combined analysis of the first-year Wilkinson Microwave Anisotropy Probe data (WMAP, `wmap.gsfc.nasa.gov`, Spergel et al., 2003) and the results from the 2-degree Field Galaxy Redshift Survey (2dFGRS, `www.2dfgrs.net`, Colless et al., 2001), the underlying ΛCDM cosmology is $\Omega_\mathrm{m} = \Omega_\mathrm{dm} + \Omega_\mathrm{b} = 0.25$, $\Omega_\mathrm{b} = 0.045$, $\Omega_\Lambda = 0.75$, $h = 0.73$, $n_\mathrm{s} = 1$, and $\sigma_8 = 0.9$. The periodic cube has a comoving side length of $500\, h^{-1}$Mpc, and the spatial resolution limit is $\approx 5\, h^{-1}$kpc comoving, which is equivalent to a dynamic range of 10^5 for the three dimensions in the whole simulation. An illustration of the result can be seen in Fig. 4.1. Using 512 processors of an International Business Machines (IBM) p690-parallel-computer for 350 000 processor hours of central processing unit (CPU) time ($\approx$ 40 years), or 28 days wall-clock time, Springel et al. (2005) calculated up to 11 000 time-steps per individual particle. The raw data volume of almost 20 terabytes holds the full particle data for 64 time steps; in the data $> 10^7$ resolved halos are identified.

Millennium-XXL Simulation

The Millennium-XXL Simulation (MXXL) (Angulo et al., 2012) uses the same underlying cosmology as the MS but differs in size and resolution. 6720^3 particles are distributed on a cube with comoving side length $3\, h^{-1}$Gpc. This makes the simulated volume as large as the observable Universe up to a redshift of $z = 0.72$. The spatial resolution limit is $10\, h^{-1}$kpc resulting in a dynamic range similar to the MS, and the mass resolution is $8.456 \times 10^9\, h^{-1} M_\odot$ with 7×10^8 resolved halos. The MXXL run took 9.3 days of wall-clock time and 2 860 000 CPU hours ($\approx$ 326 years) on 1536 nodes each equipped with two quad-core Intel-X5570-processors. From a starting redshift of $z = 63$ to today, halo and subhalo catalogues are stored for the same redshifts as for the MS. In contrast, the full particle data are only available for a couple of time steps resulting in a total output volume of 100 terabytes.

Scinet light cone simulations

The Scinet light cone simulations (SLICS) (Harnois-Déraps and van Waerbeke, 2015; Harnois-Déraps et al., 2018) are a suite of currently 1025 dark matter-only N-body simulations. The underlying ΛCDM cosmology is based on the 9-year Wilkinson Microwave Anisotropy Probe (WMAP) plus BAO plus supernova cosmology results: $\Omega_\mathrm{m} = \Omega_\mathrm{dm} + \Omega_\mathrm{b} = 0.2905$, $\Omega_\mathrm{b} = 0.0473$, $\Omega_\Lambda = 0.7095$, $h = 0.6898$, $n_\mathrm{s} = 0.969$, and

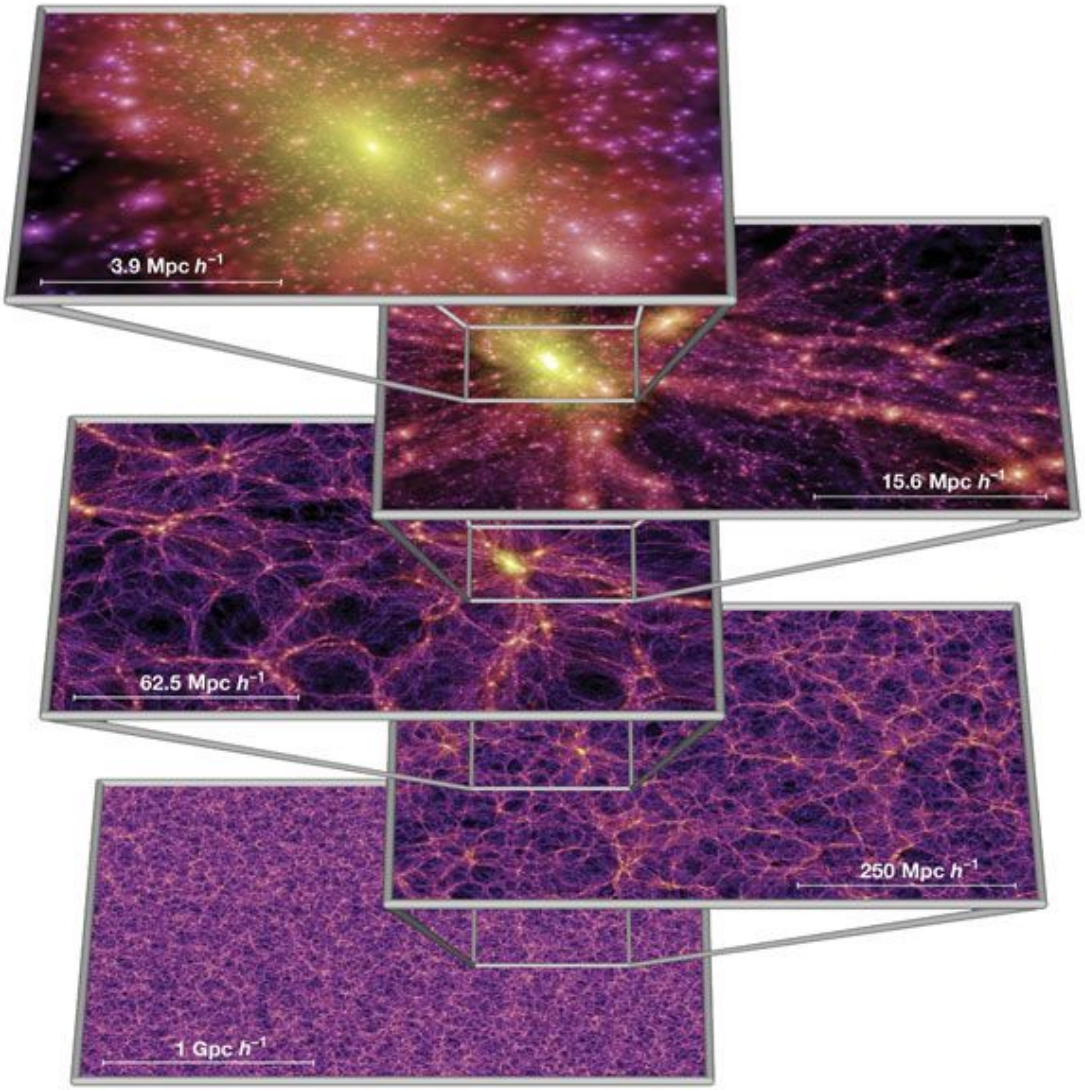


Figure 4.1. – Five zoom-ins of the projected matter density field for redshift $z = 0$ from the MS are shown. Each panel shows the projection of matter from a slice with thickness $15\,h^{-1}$Mpc. The colour traces density and local dark matter velocity dispersion. The figure is adopted from Springel et al. (2005).

$\sigma_8 = 0.826$. The number of particles is 1536^3, and the comoving side length of the periodic cube is $505\,h^{-1}$Mpc. The non-linear evolution of matter structures is traced from $z = 120$ to today with a spatial resolution of $\approx 600\,h^{-1}$kpc. For each simulation, a wall-clock time of 30 hours on 64 nodes of IBM-iDataPlex-DX360M2 machines with two quad-core Intel-E5540-processors was used. The output contains the positions of particles extrapolated on a grid for several redshifts as well as a catalogue with resolved dark matter halos and their positions.

4.1.2. Semi-analytic galaxy formation models

Dark matter-only N-body simulations are powerful tools to visualise the distribution of matter in the Universe. However, the direct detection of dark matter in observations is still elusive. To perform a comparison, we use the light-emitting galaxies that are, at best, biased tracers of the true matter distribution. However, the temporal evolution of galaxies and dark matter halos differs substantially (Springel et al., 2006). Thus, we need a bridge between observational results and dark matter-only simulations. The SAMs use simplified analytic descriptions of baryon interactions which govern the galaxy formation history and assign galaxies with respect to the merger history from their dark matter halos. The model results are then matched to key observations which in return provides insights on the dominating physical processes involved. The so-called merger trees that are needed for the SAMs are obtained from simulations, and therefore all SAMs ultimately are only valid for the chosen cosmology of the N-body simulation, although it is possible to use scaling relations for different cosmologies (Angulo et al., 2012). SAMs depend on a variety of parameters that require fine-tuning. In the following, these parameters and the concept behind them are introduced.

After identifying halos from N-body simulations, baryons are assigned to dark matter halos according to the baryon fraction $f_\mathrm{b} = \Omega_\mathrm{b}/\Omega_\mathrm{m}$. As the density of the baryon gas increases, the gas also heats up due to compression and friction. An efficient gas cooling mechanism must be present for the first, very massive Population III stars to form. The first galaxies are in the form of disk galaxies due to conservation of angular momentum (for a review of galaxy formation see, e.g., Silk, 2013). At this point, the angular momentum of the galaxy is still the same as the one of its hosting halo. After the first galaxies emerge, the reionisation of hydrogen and later helium starts which reaches its mid-point at redshift $z_\mathrm{re} = 7.82 \pm 0.71$ (*Planck* Collaboration et al., 2018a). The Universe is fully ionised by $z \approx 6$ (Bouwens et al., 2015).

Population III stars start to change the composition of the primordial gas. They are very short-lived compared to their successors, the Population II and I stars. As the first stars explode in a supernova, they inject large amounts of kinetic energy into the interstellar medium. This so-called supernova feedback suppresses gas cooling and therefore star formation. A similar form of feedback that affects star formation

is AGN feedback, which is driven by the galaxy's central supermassive black hole that accretes the surrounding, in-spiralling gas and injects kinetic energy into the interstellar medium in the form of jets.

As time progresses, halos and their galaxies undergo minor and major mergers according to the halo's mass ratios. In minor mergers, the more massive halos are barely affected and the low-mass galaxies become satellite galaxies. The satellites are then gradually disrupted and eventually destroyed by the central galaxy. Major mergers, on the other hand, affect the morphology of the colliding galaxies. The merger initiates a starburst with a subsequent quenching of star formation. Depending on the density of the galaxy's local environment, either disk galaxies with a bulge component are formed or, in more dense environments, elliptical galaxies emerge.

Lastly, the SAMs also model the observed fluxes in different filters from artificial galaxies. Using the initial mass function and the star formation history as proxies, assumptions about the stellar composition in the galaxy are made. Then, the integrated spectral energy distribution is inferred from the Hertzsprung-Russell diagram (HRD). Additionally, the presence of dust in the galaxies alters the observed spectral energy distribution. Dust attenuation is known to cause a reddening of the emitted spectra, i.e. suppression of UV and optical photons. This attenuation depends mostly on the optical depth, the column density, and the metallicity of the dust.

Thus, the SAMs incorporate models for stellar formation, supernova and AGN feedback, the initial mass function, merger outcome, as well as chemical evolution of the gas. Some model parameters need to be tuned to observations, where the exact number of parameters depends on the specific SAM. The SAM used later in this work was created by Henriques et al. (2015). They fine-tune nine parameters using mainly results from analyses of early SDSS data (Stoughton et al., 2002), e.g., the stellar mass function as a function of redshift, star formation rate and age of low-redshift galaxies, the gas-phase oxygen abundance as a function of stellar mass from star-forming galaxies, as well as the u-i colour distribution of galaxies. After fine-tuning, the SAMs are validated with independent observational results. Henriques et al. (2015) show that the model reproduces the observed galaxy correlation function, and Saghiha et al. (2017) show that the GGL signal as well as the galaxy-galaxy-galaxy lensing signal from Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS, `www.cfhtlens.org`) is matched.

4.1.3. Ray-tracing simulations

Weak lensing simulations connect lensing observables like shear and convergence to results from N-body simulations. Possible strategies are to perform a ray-tracing simulation through the full three-dimensional gravitational potential or to construct a backward light cone along one or multiple lens planes. Instead of precisely cal-

culating the photon trajectories, a fast solution can be obtained along the straight paths of the Born approximation. An overview and comparison of these and further methods have been recently provided by Hilbert et al. (2020).

The lens equation (3.2) is very generally given by

$$\boldsymbol{\beta}(\boldsymbol{\theta}, z_\mathrm{s}) = \boldsymbol{\theta} - \frac{2}{c^2}\int_0^{\chi_\mathrm{s}} \mathrm{d}\chi_\mathrm{d}\, \frac{f_\mathrm{ds}}{f_\mathrm{d} f_\mathrm{s}} \nabla_{\boldsymbol{\beta}} \Phi[\boldsymbol{\beta}(\boldsymbol{\theta}, z_\mathrm{d}),\, \chi_\mathrm{d}, \chi_\mathrm{s}] \,, \tag{4.3}$$

where $f_\mathrm{ds} = f_\mathrm{K}(\chi_\mathrm{s} - \chi_\mathrm{d})$, $f_\mathrm{d} = f_\mathrm{K}(\chi_\mathrm{d})$, and $f_\mathrm{s} = f_\mathrm{K}(\chi_\mathrm{s})$ are comoving angular diameter distances. The redshift corresponding to $\chi_\mathrm{d,s}$ is $z_\mathrm{d,s}$, and $\nabla_{\boldsymbol{\beta}}$ denotes a gradient with respect to the angular source position $\boldsymbol{\beta}$. Lastly, $\Phi(\boldsymbol{\beta}, \chi_\mathrm{d}, \chi_\mathrm{s})$ denotes the three-dimensional Newtonian potential at position $[\boldsymbol{\beta}(\boldsymbol{\theta}, z_\mathrm{d}), \chi_\mathrm{d}]$. The Jacobian for a flat sky is then

$$\frac{\partial \beta_i}{\partial \theta_j} = \delta_{ij} - \frac{2}{c^2}\int_0^{\chi_\mathrm{s}} \mathrm{d}\chi_\mathrm{d}\, \frac{f_\mathrm{ds}}{f_\mathrm{d} f_\mathrm{s}}\, \frac{\partial^2 \Phi[\boldsymbol{\beta}(\boldsymbol{\theta}, z_\mathrm{d}),\, \chi_\mathrm{d}, \chi_\mathrm{s}]}{\partial \beta_i\, \partial \beta_j}\, \frac{\partial \beta_k(\boldsymbol{\theta}, z_\mathrm{d})}{\partial \theta_j} \,, \tag{4.4}$$

which is connected to the convergence and shear by Eq. (3.13).

Multiple-lens-plane approximation

Using the information of the three-dimensional gravitational potential of the continuous mass distribution is computationally very expensive. Instead, the density contrast can be collapsed onto a series of two-dimensional mass distributions, i.e. $\delta = \sum \delta_\mathrm{2D}$. In the flat sky approximation, the planes lie perpendicular to the central line-of-sight of the observer's backward light cone; the lens planes act as a series of deflectors along the line-of-sight (Blandford and Narayan, 1986), reducing (4.3) and (4.4) to a discrete sum.

Since N-body simulations have a finite size, constructing a backward light cone results in the repetition of the same matter structures at different redshifts. To avoid the introduction of spurious correlations in the lensing data, random translations, parity transformations, and rotations by 90° around the coordinate axis are applied to the distribution of particles. This is possible due to the periodic boundary conditions of the simulation (Martel et al., 2002).

The deflection angle (3.10) for the k^th lens plane as a function of $\boldsymbol{\beta}$ is

$$\boldsymbol{\alpha}^{(k)}(\boldsymbol{\beta}^{(k)}) = \nabla_{\boldsymbol{\beta}^{(k)}} \psi^{(k)}(\boldsymbol{\beta}^{(k)}) \,. \tag{4.5}$$

Starting from the angular position $\boldsymbol{\theta} = \boldsymbol{\beta}^{(1)}$, a light ray is traced back to a plane $k > 1$ by

$$\boldsymbol{\beta}^{(k)}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \sum_{l=1}^{k-1} \frac{f^{(l,k)}}{f^{(k)}} \boldsymbol{\alpha}^{(l)}(\boldsymbol{\beta}^{(l)}) \,, \tag{4.6}$$

where $f^{(l,k)} = f_{\rm K}(\chi^{(k)} - \chi^{(l)})$. Equation (4.6) reduces to the lens equation (3.2) in the case $k = 2$. For practical reasons, (4.6) is rewritten to (Hartlap, 2009)

$$\boldsymbol{\beta}^{(k)} = \left(1 - \frac{f^{(k-1)}}{f^{(k)}}\frac{f^{(k-2,k)}}{f^{(k-2,k-1)}}\right)\boldsymbol{\beta}^{(k-2)} + \frac{f^{(k-1)}}{f^{(k)}}\frac{f^{(k-2,k)}}{f^{(k-2,k-1)}}\boldsymbol{\beta}^{(k-1)} - \frac{f^{(k-1,k)}}{f^{(k)}}\boldsymbol{\alpha}^{(k-1)}(\boldsymbol{\beta}^{(k-1)}) \,, \tag{4.7}$$

which is visualised in Fig. 4.2. Starting again from $\boldsymbol{\theta} = \boldsymbol{\beta}^{(1)}$, the angular positions on other lens planes are efficiently computed by iterating (4.7).

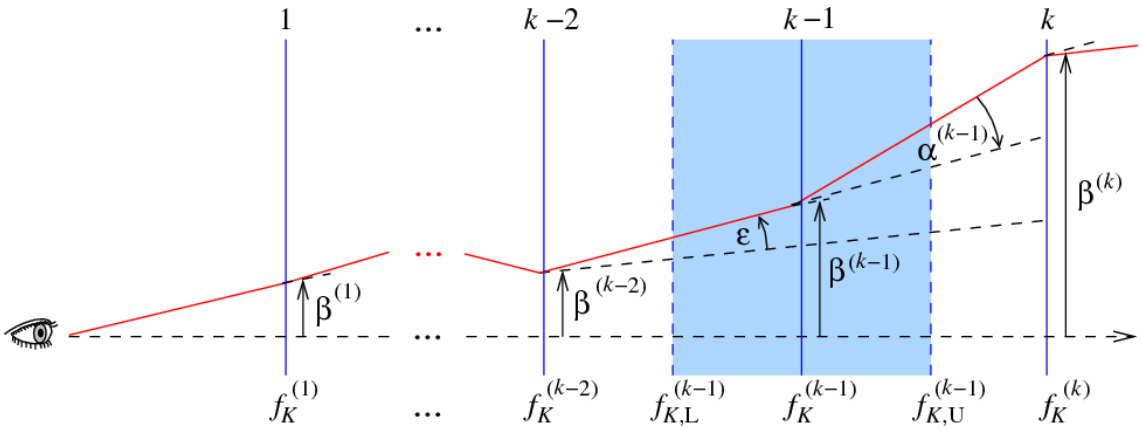


Figure 4.2. – Schematic view of the multiple lens plane and its backward light cone. The continuous matter distributions are projected on discrete lens planes. A light ray originating from the observer's position travels backward through the k lens planes, where it experiences a series of deflections $\boldsymbol{\alpha}^{(k)}$. Following Eq. (4.7), each angular position $\boldsymbol{\beta}$ on plane k is calculated by its two preceding positions $\boldsymbol{\beta}^{(k-1)}$ and $\boldsymbol{\beta}^{(k-2)}$. The figure is adopted from Hilbert et al. (2009).

Following Bartelmann and Schneider (2001), the convergence of a three-dimensional mass distribution in a flat universe in a slice is then given as

$$\kappa^{(k)}(\boldsymbol{\beta}^{(k)}) = \frac{3H_0^2\Omega_{\rm m}}{2c^2}\frac{f^{(k)}}{a^{(k)}}\int_{\chi_{\rm low}^{(k)}}^{\chi_{\rm high}^{(k)}} {\rm d}\chi\; \delta(\boldsymbol{\beta}^{(k)}, \chi) \,, \tag{4.8}$$

where the slice boundaries $\chi_{\rm low}^{(k)}$ and $\chi_{\rm high}^{(k)}$ have to satisfy $\chi_{\rm low}^{(k)} < \chi^{(k)} < \chi_{\rm high}^{(k)}$ and $\chi_{\rm high}^{(k)} = \chi_{\rm low}^{(k+1)}$.

Born approximation

In the Born approximation, we replace $\boldsymbol{\beta}$ by the angular position $\boldsymbol{\theta}$ and then Taylor expand (4.4) to first-order in Φ which yields

$$\boldsymbol{\beta}(\boldsymbol{\theta}, z_{\rm s}) = \boldsymbol{\theta} - \frac{2}{c^2}\int_0^{\chi_{\rm s}} {\rm d}\chi_{\rm d}\; \frac{f_{\rm ds}}{f_{\rm d}f_{\rm s}}\nabla_{\boldsymbol{\theta}}\Phi(\boldsymbol{\theta},\, \chi_{\rm d}, \chi_{\rm s}) \,, \tag{4.9}$$

and

$$\frac{\partial \beta_i}{\partial \theta_j} = \delta_{ij} - \frac{2}{c^2} \int_0^{\chi_\mathrm{s}} \mathrm{d}\chi_\mathrm{d} \, \frac{f_\mathrm{ds}}{f_\mathrm{d} f_\mathrm{s}} \, \frac{\partial^2 \Phi[\boldsymbol{\theta}, \, \chi_\mathrm{d}, \chi_\mathrm{s}]}{\partial \theta_i \, \partial \theta_j} \, . \tag{4.10}$$

This first-order approximation also encompasses neglecting lens-lens coupling, which is sometimes treated as a part of the Born approximation. Differences occur mainly on small angular scales (Hirata and Seljak, 2003).

The convergence (4.8) becomes

$$\kappa(\boldsymbol{\theta}) = \frac{3 H_0^2 \Omega_\mathrm{m}}{2c^2} \int_0^{\chi_\mathrm{s}} \mathrm{d}\chi_\mathrm{d} \, (1 + z_\mathrm{d}) \, \frac{f_\mathrm{d} f_\mathrm{ds}}{f_\mathrm{s}} \delta(\boldsymbol{\theta}, \chi_\mathrm{d}) \, , \tag{4.11}$$

where again the multiple lens plane approximation can be applied. The solution is then obtained by summing (4.11) over the series of two-dimensional projected lens planes δ_2D from the observer to a maximum comoving angular diameter distance. The maximum distance is typically chosen to be larger than the largest redshifts observed in current and future weak lensing surveys.

4.2. Fast extraction of shear estimates

The synthetic lensing data in combination with SAMs provide the perfect tools to explore GGL. However, the vast amount of available galaxy and shear data must first be reduced and linked to observables. It is then possible to numerically study observational challenges as well as test new methods to mitigate biases. The open-source C++-Code `ShearCo` available on `github.com/SandraUnruh/shearco` has been developed throughout the course of this thesis and all results presented in Chapter 5, 6 and 7 are based on variants of this code. In this section, `ShearCo` is described and its performance verified. This method is capable of efficiently dealing with very large data sets due to the employment of FFT.

4.2.1. Obtaining a shear profile

First, a detailed description is given on how azimuthally-averaged tangential shear estimates are obtained as function of distance from the lens centre. All following examples and visualizations refer to results from the MS by Springel et al. (2005, see also Sect. 4.1.1), SAMs by Henriques et al. (2015, see also Sect. 4.1.2), and ray-tracing simulations by Hilbert et al. (2009, see also Sect. 4.1.3).

Shear input

From the ray-tracing simulations, the components of the Jacobi matrix are given on $N_\mathrm{pix} = 4096^2$ pixels, from which the shear γ and convergence κ (Eq. 3.13) as well as

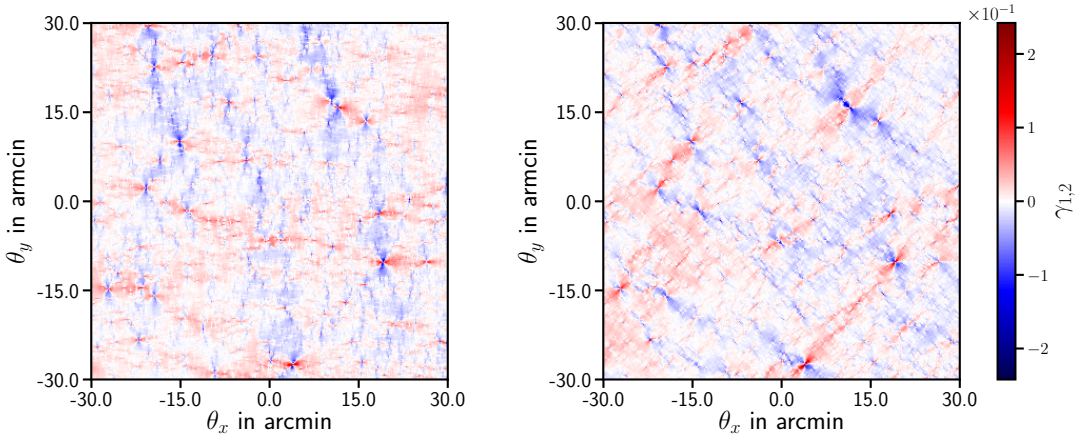


Figure 4.3. – Shown is a visual representation of the Cartesian coordinates of shear $\gamma_{1,2}$. The redshift of the shear field is $z = 0.99$ and the shear catalogue name is 'los_8_2_4_[...]_plane_41'.

the magnification μ (Eq. 3.15) are calculated. Fig. 4.3 shows a visualisation of the Cartesian shear at a redshift $z = 0.99$, where the choice of the specific shear field is unimportant. For consistency, all following shear fields and estimates are shown for the same field. The originally $4° \times 4°$ per shear field are cropped to $1° \times 1°$ for better visibility. A direct comparison of γ_1 and γ_2 shows that apparent structures are rotated by 45° to each other, as expected from the general properties of lensing shear.

Lens input

Lens galaxy positions are extracted from the synthetic galaxy catalogue as obtained from the SAM. Only such galaxies are extracted that fulfil some pre-set conditions, e.g., applying a cut in magnitude, mass, or galaxy colours. It must be kept in mind, that the galaxy positions are directly computed from the positions of the dark matter halos in the N-body simulation. Therefore, they are not a priori confined to the same regular grid as the ray-tracing results.

It is advantageous to project the lens positions on the same regular grid that is used for the shear field. The easiest option is to assign galaxies to their nearest grid point (NGP). Thus, we define a square field of $N_{\rm pix}$ pixels that contains the number of galaxies on each pixel, which is typically zero, one, or two. For this method, analyses close to the centre of the galaxy suffer from discretisation effects on scales comparable to the pixel width 3.5 arcsec.

A more advanced and commonly used method is a so-called cloud-in-cell algorithm (CiC). The CiC interpolates the lens position to the four NGPs. If $x_{\rm g}$ and $y_{\rm g}$ are the true galaxy positions in pixel coordinates, the fractions of a galaxy $f^{\rm CiC}$ on the

nearest grid positions (i, j) to $(i+1, j+1)$ are

$$\begin{aligned} f^{\mathrm{CiC}}_{i,j} &= (x_{i+1} - x_{\mathrm{g}})\,(y_{i+1} - y_{\mathrm{g}})\ , \\ f^{\mathrm{CiC}}_{i+1,j} &= (x_{i} - x_{\mathrm{g}})\,(y_{i+1} - y_{\mathrm{g}})\ , \\ f^{\mathrm{CiC}}_{i,j+1} &= (x_{i+1} - x_{\mathrm{g}})\,(y_{i} - y_{\mathrm{g}})\ , \\ f^{\mathrm{CiC}}_{i+1,j+1} &= (x_{i} - x_{\mathrm{g}})\,(y_{i} - y_{\mathrm{g}})\ , \end{aligned} \tag{4.12}$$

where the sum of all f^{CiC} is normalised to one. For galaxies with positions $x_{\mathrm{g}} > 4095$ or $y_{\mathrm{g}} > 4095$ the respective coordinate is set to 4095.

The density of lens galaxies n_{d} for the grid positions $\boldsymbol{\theta}_i$ for the NGP method is given as

$$n_{\mathrm{d}}(\boldsymbol{\theta}) = \frac{1}{N_{\mathrm{d}}} \sum_{i=1}^{N_{\mathrm{d}}} \delta_{\mathrm{D}}(\boldsymbol{\theta} - \boldsymbol{\theta}_i)\ , \tag{4.13}$$

where for CiC grid positions one also has to sum over all four positions with their appropriate weights f^{CiC}. While the CiC method can be enabled in the code, galaxies are typically assigned using the NGP method since the misplacement is negligible for all cases considered in this thesis.

Fast-Fourier transformation

Averaging the shear field around many lenses is a convolution of the lens density and the shear field. Thus, we can apply the convolution theorem, which states that under certain conditions, the point-wise product of the Fourier transform of two functions is equal to the Fourier transform of their convolution. In this case, the average shear signal is written as

$$\begin{aligned} \langle\gamma\rangle(\boldsymbol{\theta}) &= \int \mathrm{d}^2\theta'\, \gamma(\boldsymbol{\theta} + \boldsymbol{\theta}')\, n_{\mathrm{d}}(\boldsymbol{\theta}') \\ &= \mathcal{F}^{-1}\Big\{\mathcal{F}\{\gamma\} \cdot \mathcal{F}^{*}\{n^{*}_{\mathrm{d}}\}\Big\}\,(\boldsymbol{\theta})\ , \end{aligned} \tag{4.14}$$

where $\mathcal{F}\{f\}$ represents a Fourier transform of a function f and $\mathcal{F}^{-1}\{f\}$ the inverse Fourier transform. The asterisk indicates a complex conjugate which is not needed for the real number density but is mentioned for completeness. Please note that all two dimensional integrals without specific boundaries are assumed to be integrated over the whole simulated field.

To process the shear information efficiently, an FFT is employed. An FFT is a numerical method for discrete values in arbitrary dimensions that is known for its fast calculation speed (Press et al., 1992b). The `C` subroutine library of Frigo and Johnson (2012) is used for the numerical computation of the Fourier transforms. Since an FFT implicitly assumes periodic boundary conditions, shear estimates from lenses near the edges of the field are bound to be inaccurate. To mitigate this effect,

we restrict the area from which we select lenses to the inner $(4^\circ - 2\theta_{\mathrm{out}})^2$ of the field. For numerical calculations in this chapter $\theta_{\mathrm{out}} = 17\farcm5$ is assumed, if not noted otherwise.[2]

The result of convolving the shear field with a lens density field at redshift $z_{\mathrm{d}} = 0.41$ is shown in Fig. 4.4. The GGL signal is averaged over $N_{\mathrm{d}} = 13\,346$ lenses brighter than 22 mag in the r-band. In the first two panels $\gamma_{1,2}$ are displayed as obtained by numerically solving Eq. (4.14). The two shear components still show a rotation of 45° with respect to each other. However, the tangential shear as well as the cross shear are on average different from zero due to the presence of long modes in the density field. A subtraction of the lensing signal around random points mitigates this effect (Singh et al., 2017), which is seen in the last two panels.

Tangential shear

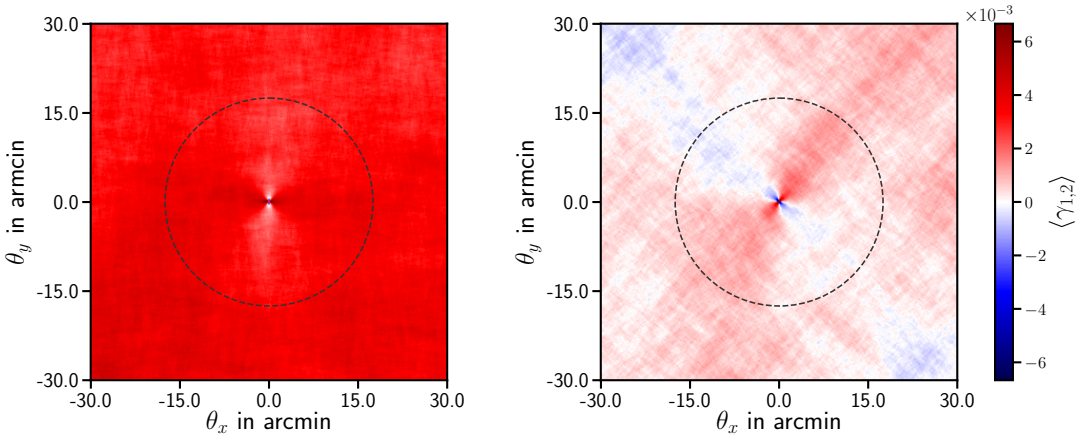
Rotating the Cartesian shear components as introduced in Eq. (3.27) results in the tangential and cross shear. A visual representation of these two shear components is given in Fig. 4.5. The stacked halos around the lens galaxies show an approximately symmetric profile and, therefore, also the measured tangential shear profile is expected to be symmetric around the lens centre. The cross shear, on the other hand, is expected to vanish. In the first two panels of Fig. 4.5 the expectation is certainly not fulfilled, which stresses once more how important the subtraction of the signal around random points is. If the signal around random points is subtracted, the cross shear signal shows only small fluctuations, as can be seen in Fig. 4.5. The fluctuations are due to statistical uncertainties and are reduced if the data of all 64 available lines-of-sight are averaged. In the following, all shear estimates are implicitly given with the signal around random points subtracted, i.e. $\gamma \to \gamma - \gamma_{\mathrm{rand}}$.

The final tangential shear profile is obtained by azimuthally averaging the shear around the lens centre. By using polar coordinates with the lens in the centre, we obtain

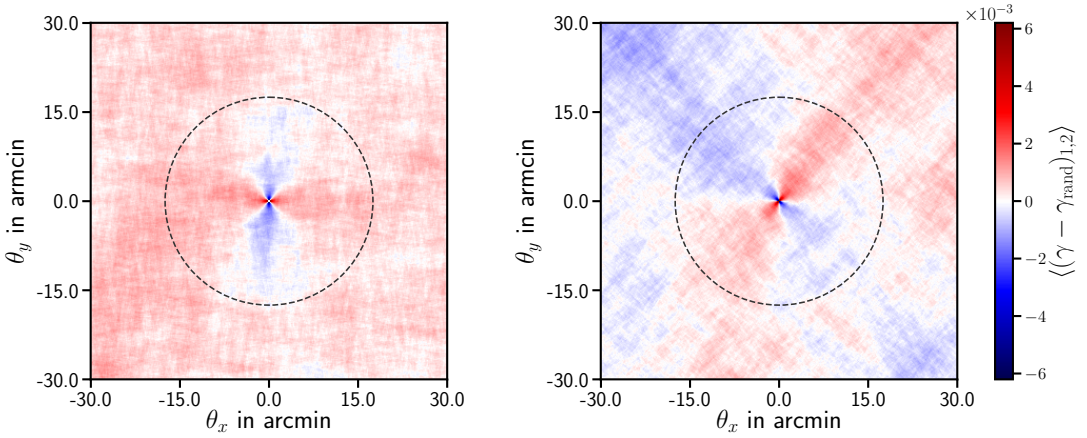
$$\langle \gamma_{\mathrm{t,x}}^{\mathrm{R}} \rangle(\theta) = \int \frac{\mathrm{d}\varphi}{2\pi} \, \langle \gamma_{\mathrm{t,x}} \rangle(\theta, \varphi) \; . \tag{4.15}$$

The averaging is typically done in logarithmically spaced annuli. For 16 annuli in the range $0\farcm6 \leq \theta < 17\farcm5$, the result is shown in Fig. 4.6. The shear profile is the average profile of all 64 lines-of-sight that the MS offers, while errors are obtained by a Jackknife method. The cross shear is compatible with zero and at least three orders of magnitude smaller than the tangential shear. A close inspection of the average shear profile shows a kink at $\theta \sim 4'$ corresponding to a comoving distance of $R \sim 1.8\,\mathrm{Mpc}$. The kink is predicted by the previously discussed halo model and shows the transition from the 1-halo to the 2-halo term.

[2] Alternatively, we also provide a zero-padding method and extend the shear and lens field to $(4^\circ + \theta_{\mathrm{out}})^2$. For this method to work, we need to modify Eq. (4.14) to account for the empty areas, which is discussed in Sect. 4.3 with the inclusion of shape noise and masks.

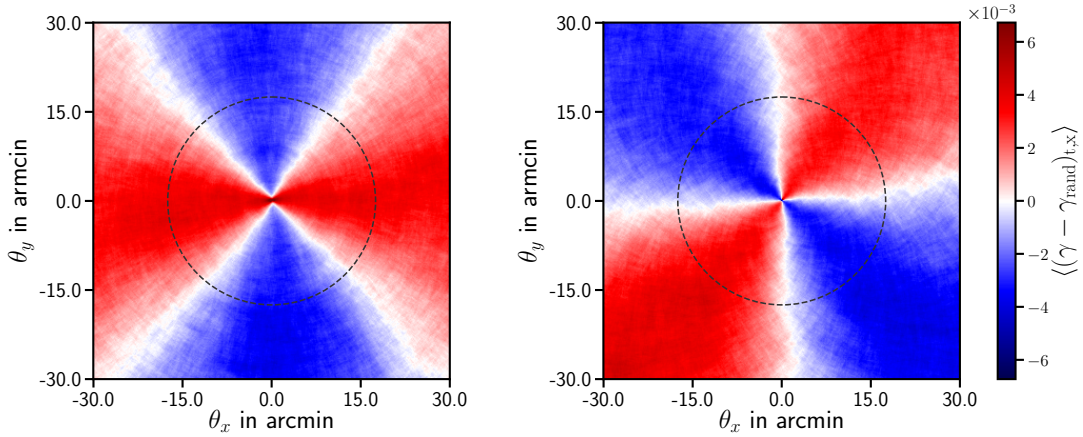


(a) Displayed is the lensing signal $\gamma_{1,2}$ averaged over 13 346 lenses. In this shear field, the average shear signal of the two shear components is different from zero.

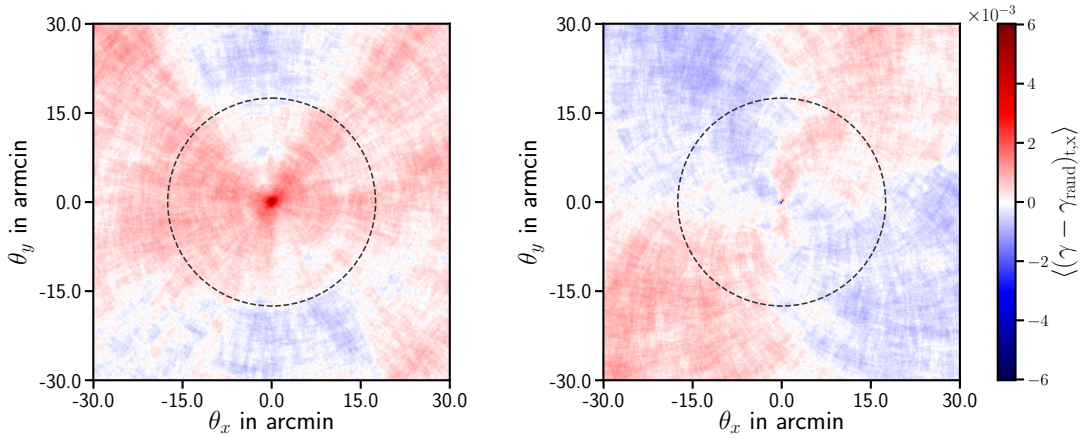


(b) Subtracting the shear signal around random points removes the large-scale signal in the area of interest. The resulting shear field is symmetric but noisy, as naively expected.

Figure 4.4. – Shown is a visual representation of the Cartesian coordinates of shear γ_1 (left panels) and γ_2 (right panels) averaged over lens galaxies in a foreground field at redshift $z_\mathrm{d} = 0.41$ and the shear field shown in Fig. 4.3 at $z_\mathrm{s} = 0.99$. The lens galaxies are brighter than 22 mag in the r-band and the circle with $\theta_\mathrm{out} = 17\farcm5$ indicates the outer boundary of reliable shear estimates. The lens catalogue name is 'los_8_2_4_[...]_Henriques2014_[...]_plane_50'.



(a) These two panels are analogous to Fig. 4.4a for a visualization in tangential and cross shear components. Both signals show strong asymmetries.



(b) In analogy to Fig. 4.4b, the GGL signal around random points is subtracted from results in the upper two panels. Now, the expected symmetry around the lens centre for the tangential shear and the almost vanishing cross shear is seen.

Figure 4.5. – Shown is a visual representation of tangential shear γ_{t} (left panels) and cross shear γ_{x} (right panels) averaged over lens galaxies in a foreground field for the same lenses and sources as in the previous figures.

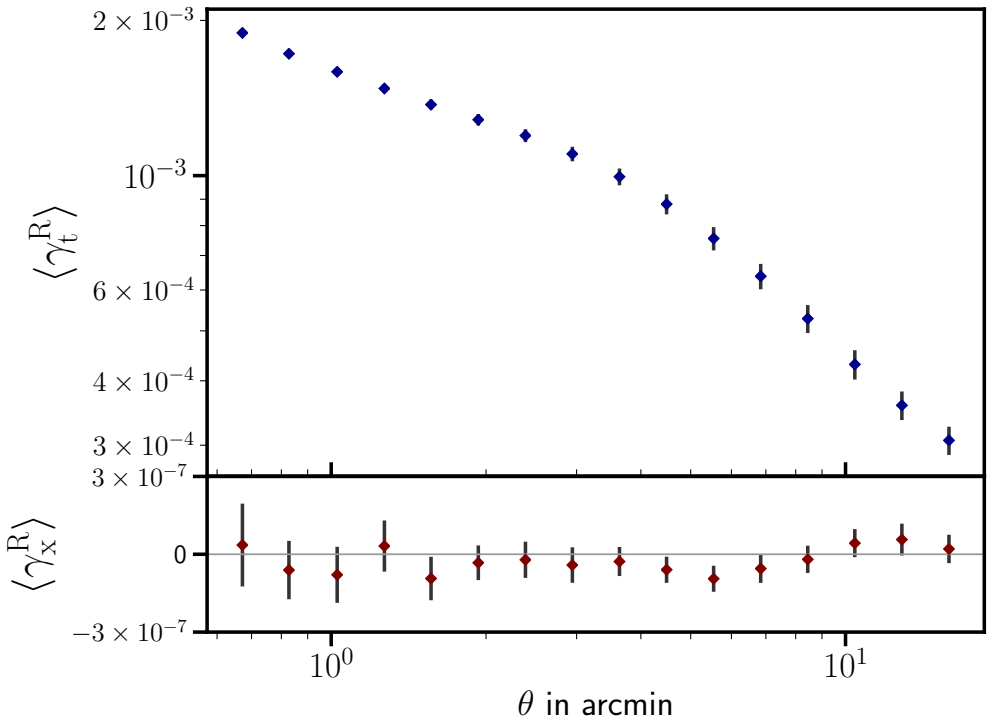


Figure 4.6. – Shown is the tangential and cross shear profile which is obtained by azimuthally averaging the GGL signals from Fig. 4.5b. For this figure, all 64 lines-of-sight are used which contain $N_d = 822\,213$ lenses brighter than 22 mag in the r-band. Errors are obtained by a Jackknife method.

4.2.2. Obtaining an integrated shear profile

To further compress the information of the GGL signal, it is possible to integrate the shear profile shown in Fig. 4.6 with respect to the angle θ. An obvious solution to this problem is to integrate the shear profile (4.15) in the radial direction

$$\langle\bar{\gamma}_\mathrm{t}\rangle = \int_{\theta_\mathrm{in}}^{\theta_\mathrm{out}} \theta\,\mathrm{d}\theta\;\langle\bar{\gamma}_\mathrm{t}^\mathrm{R}(\theta)\,w(\theta)\;, \tag{4.16}$$

where $w(\theta)$ is an optional weighting function, e.g., to optimise the signal-to-noise (S/N). Numerically, one has to choose an appropriate integration scheme and as a fairly simple approach, one could use the trapezoidal rule. By making (correct) assumptions about the true form of the underlying function, it is always possible to increase the level of sophistication. However, the accuracy of the integration method certainly is a function of the number of annuli in the integration range, where the smallest meaningful size of an annulus is given by the size of one pixel. Unfortunately, this approach leads to a considerable increase in computation time.

Thus, it is advantageous to avoid employing an integration scheme at all, since we do not have to carefully balance computation time and numerical accuracy. Therefore, we reconsider the steps made in Eq. (4.14) to (4.16). Instead of averaging the shear field over many lenses first, it is convenient to perform the azimuthal and radial integration first

$$\bar{\gamma}_\mathrm{t} = -\Re\left[\int \mathrm{d}^2\theta'\,\gamma(\boldsymbol{\theta}+\boldsymbol{\theta}')\,\frac{\boldsymbol{\theta}'^*}{\boldsymbol{\theta}'}w(|\boldsymbol{\theta}'|)\right]\;, \tag{4.17}$$

where $\Re(x)$ is the real value of a quantity x and we again add an optional weighting function $w(\theta)$. As an illustrative example, a viable weighting function is

$$w(\theta) = \frac{\theta_\mathrm{E}^2}{\theta^2}\,\frac{\mathcal{H}(\theta-\theta_\mathrm{in})\,\mathcal{H}(\theta_\mathrm{out}-\theta)}{\pi(\theta_\mathrm{out}^2-\theta_\mathrm{in}^2)}\;, \tag{4.18}$$

which is only supported in the region $\theta_\mathrm{in} \leq \theta \leq \theta_\mathrm{out}$ being enforced by the Heaviside step function $\mathcal{H}$. This form of the function also optimises the S/N ratio for a shear profile behaving like $1/\theta$, i.e. an SIS, as is shown in Sect. 4.3.3.

Equation (4.17) is solved with the convolution theorem, where we have to calculate the Fourier transforms $\mathcal{F}\{\gamma\}$ and $\mathcal{F}^*\{[\,\boldsymbol{\theta}^*/\boldsymbol{\theta}\,w(|\boldsymbol{\theta}|)\,]^*\}$. Finally, we simply obtain the integrated shear averaged over all lenses by

$$\langle\bar{\gamma}_\mathrm{t}\rangle = \frac{1}{N_\mathrm{d}}\sum_{i=1}^{N_\mathrm{d}}\bar{\gamma}_\mathrm{t}(\boldsymbol{\theta}_i)\;, \tag{4.19}$$

which must be slightly modified if lens positions are assigned with the CiC-method.

4.3. Masks, shape noise, magnification effects, and signal-to-noise

To mimic real data within its limitations more closely, the methods above must be modified and extended. In this section, numerical methods are described to include the dominant source of noise and sky masking, which is unavoidable in wide-field imaging surveys. We further study effects from lensing magnification in detail. Lastly, the S/N ratio is compared to a simplified analytical model.

4.3.1. Inclusion of shape noise and masks

Shape noise is the main source of noise in weak lensing measurements. Starting from results of the ray-tracing simulation, we artificially add intrinsic ellipticities for each galaxy, and also limit the number density of source galaxies to realistic values.[3]

To mimic a realistic set of galaxies with intrinsic ellipticities, a simplified model of additive, uncorrelated Gaussian noise with zero mean and width $\sigma_\epsilon = 0.3$ is employed. The Gaussian noise is added to each Cartesian shear component as given in the ray-tracing catalogue. Besides, we reduce the density of sources brighter than 24.5 mag in the r-band to $\approx 35\,\mathrm{gal/arcmin}^2$ summed over all redshifts. Accordingly, a set of random source galaxy positions is assigned to every source plane. Analogous to (4.13), we can then define a function n_s for all N_s.

To include the changes in the source density, we weigh Eq. (4.14) by the number of lens-source pairs that can be found with separation $\boldsymbol{\theta}$

$$\langle\gamma\rangle(\boldsymbol{\theta}) = \frac{\int \mathrm{d}^2\theta'\,\gamma(\boldsymbol{\theta}+\boldsymbol{\theta}')\,n_\mathrm{s}(\boldsymbol{\theta}+\boldsymbol{\theta}')\,n_\mathrm{d}(\boldsymbol{\theta}')}{\int \mathrm{d}^2\theta'\,n_\mathrm{s}(\boldsymbol{\theta}+\boldsymbol{\theta}')\,n_\mathrm{d}(\boldsymbol{\theta}')}\,. \tag{4.20}$$

Equivalently, Eq. (4.17) is weighted by the number of sources that can be found in an annulus from $\theta_\mathrm{in} \leq \theta < \theta_\mathrm{out}$

$$\bar{\gamma}_\mathrm{t}(\boldsymbol{\theta}) + \mathrm{i}\bar{\gamma}_\mathrm{x}(\boldsymbol{\theta}) = -\frac{\int \mathrm{d}^2\theta'\,\gamma(\boldsymbol{\theta}+\boldsymbol{\theta}')\,n_\mathrm{s}(\boldsymbol{\theta}+\boldsymbol{\theta}')\,\frac{\boldsymbol{\theta}'^*}{\boldsymbol{\theta}'}w(|\boldsymbol{\theta}'|)}{\int \mathrm{d}^2\theta'\,n_\mathrm{s}(\boldsymbol{\theta}+\boldsymbol{\theta}')\,w(|\boldsymbol{\theta}'|)}\,, \tag{4.21}$$

where in both cases nominators and denominators are calculated with an FFT. These modification encompass empty regions from masks by design. Now, a zero-padding method can be included to avoid biases from the periodic boundary conditions. It is further possible, to enforce a minimum number of lens-source pairs that must be present for a shear estimate to be valid.

The mean separation $\langle\theta\rangle$ of lens-source pairs within an annulus in the presence of masks, survey boundaries, and realistic source number densities can also be obtained

[3] Using the full information from the ray-tracing data leads to an artificial source galaxy density of $17\,000\,\mathrm{gal/arcmin}^2$ per redshift slice and independent of flux-limit.

by an FFT. If galaxy positions are truly random and Poisson noise is low, i.e. the number of lens-source pairs is very high, $\langle\theta\rangle$ per bin is simply calculated as

$$\langle\theta\rangle_i = \sqrt{\frac{\theta_{i+1}^2 + \theta_i^2}{2}} \,, \tag{4.22}$$

where θ_i are the 17 boundaries of the 16 annuli. For the more realistic case, let us denote the denominator of Eq. (4.20) with $N_{\mathrm{ds-pairs}}$, then $\langle\theta\rangle$ per bin is calculated as

$$\langle\theta\rangle_i = \frac{\int \mathrm{d}^2\theta\; \mathcal{H}(\theta - \theta_i)\, \mathcal{H}(\theta_{i+1} - \theta)\; N_{\mathrm{ds-pairs}}(\boldsymbol{\theta})\; |\boldsymbol{\theta}|}{\int \mathrm{d}^2\theta\; \mathcal{H}(\theta - \theta_i)\, \mathcal{H}(\theta_{i+1} - \theta)\; N_{\mathrm{ds-pairs}}(\boldsymbol{\theta})} \,. \tag{4.23}$$

The mean separation of pairs in an annulus is important for connecting the shear profile to its underlying dark matter halo, and comparison to theoretical models in general. Magnification effects that are discussed below are known to change the number density of galaxies on the sky; therefore, also the mean angular distance in a bin is slightly affected. A comparison of the true mean separation (4.22) to the idealised (4.23) may even give the first insights into how much magnification affects the GGL estimate. In any case, caution must be applied since in general is $\langle\gamma(\theta)\rangle \neq \gamma(\langle\theta\rangle)$.

4.3.2. A switch for magnification effects

Magnification changes the observed number density of galaxies on the sky, which includes lens as well as source galaxies. Theoretical predictions tend to ignore magnification effects which is not a viable strategy in the age of wide-field, deep lensing surveys. Neglecting how magnification affects GGL estimates certainly leads to biases in the result. The advantage of ray-tracing simulations is that magnification is simulated accurately. To study effects that arise from ignoring or misestimating magnification, we now describe a way to turn magnification off and on in simulated data.

Lens galaxies

Lens galaxies as extracted from the galaxy catalogues are influenced by magnification by design. Galaxies brighter than a flux-limit $s_{\mathrm{lim,d}}$ just need to be assigned to their NGP. Magnification influences the observed number density on the sky by magnifying the galaxy's flux and by shifting its observed position on the sky. The former effect can be removed easily, but the original galaxy positions cannot be recovered in the available ray-tracing simulations. Therefore, to remove the magnification effect on lens galaxies, a different method must be employed.

We Taylor expand Eq. (3.16) for small magnifications $|\mu - 1| \ll 1$ in the weak lensing limit and obtain to first order

$$\frac{n(> s)}{n_0(> s)} = \mu^{\alpha - 1} \,, \tag{4.24}$$

where α is the local slope at the flux-limit $s_{\rm lim}$

$$\alpha = -\frac{\mathrm{d}\log_{10} n_0(> s)}{\mathrm{d}\log_{10} s}\bigg|_{s_{\rm lim}} \,. \tag{4.25}$$

From (4.24) follows that galaxy counts are enhanced (reduced) for local slopes $\alpha > 1$ ($\alpha < 1$) and are unaffected for $\alpha = 1$. We therefore transform the demagnified lens number density to a distribution with slope one. We let the number of galaxies at the flux-limit stay constant, i.e. $n_0(> s_{\rm lim,d}) = n_0'(> s_{\rm lim,d}')$, and obtain the transformed flux

$$s_0'(s_0) = s_{\rm lim,d}\,\frac{n_0(s_{\rm lim,d})}{n_0(s_0)} \,. \tag{4.26}$$

In the transformed flux system α' is unity, independent of the flux-limit. We calculate a transformed number density $n_{\rm d}'(> \mu s_0')$ and choose only galaxies brighter than the flux-limit $s' = \mu s_0' > s_{\rm lim}$. A catalogue of flux-limited lens galaxies chosen like this is, therefore, not affected by magnification, although this is not true for the individual lens galaxies.

The number of lens galaxies is expected to stay constant within the numerical accuracy when the magnification is switched off. For a magnitude cut of 22 mag and a lens redshift of $z_{\rm d} = 0.41$, 595 348 lenses are found with a cut in observed magnitude and 595 355 lenses with transformed magnitude s'. A detailed analysis showed that for the intrinsic magnitude 917 galaxies became brighter than 22 mag while 924 galaxies became dimmer, which leaves the number of galaxies almost unchanged.

Source galaxies

In principle, the magnification effects for the source galaxies can be removed as was done for the lens galaxies. However, in contrast to the lens galaxies, it is not important to keep track of the true galaxy positions. Therefore, the easiest method is to assign random positions that have the same number density $n_{\rm s}(s_{\rm lim,s})$ as the catalogue source galaxies, where $s_{\rm lim,s}$ is the source galaxy's flux-limit. This allows for the realisation of an arbitrary number of source mock catalogues.

A source mock catalogue, that is subject to magnification effects, is created with the following strategy. We first calculate the cumulative demagnified source number density $n_{\rm s,0}(> s_0, \boldsymbol{\theta})$ from the simulated data. The expected source number density

as a function of the local magnification $\mu(\boldsymbol{\theta})$ is obtained by inserting a flux-limit $s_{\mathrm{lim,s}}$ into Eq. (3.16)

$$n_{\mathrm{s}}(>s_{\mathrm{lim}},\boldsymbol{\theta}) = \frac{1}{\mu(\boldsymbol{\theta})}\, n_0\left(>\frac{s_{\mathrm{lim}}}{\mu(\boldsymbol{\theta})},\boldsymbol{\theta}\right) , \tag{4.27}$$

where the flux-limit is locally modified by the magnification to $s_{\mathrm{lim}}/\mu(\boldsymbol{\theta})$.

We define the threshold of finding a source galaxy at angular position $\boldsymbol{\theta}$

$$T(\boldsymbol{\theta}) = \frac{A\, n_{\mathrm{s}}(>s_{\mathrm{lim}},\boldsymbol{\theta})}{N_{\mathrm{pix}}} , \tag{4.28}$$

with $A = 16\,\mathrm{deg}^2$ being the area of one simulated field and where we enforce $T < 1$. We then draw a random number $P(\boldsymbol{\theta})$ between zero and one and assign a galaxy at position $\boldsymbol{\theta}$ if $T(\boldsymbol{\theta}) > P(\boldsymbol{\theta})$. In general, the number of assigned source galaxies differs from the number in the demagnified case due to magnification effects. Using this method, an arbitrary number of mock catalogues can be generated. By inserting $\mu(\boldsymbol{\theta}) = 1$ into (4.27), we recover an assignment of purely random positions, i.e. the 'magnification switched off' case.

4.3.3. Signal-to-noise in theory and simulation

Using simplifying assumptions about the underlying average mass profile, it is fairly easy to calculate the expected S/N ratio for GGL. To derive this equation, we closely follow the considerations given in Bartelmann and Schneider (2001).

We assume that the gravitational lens has the form of an SIS as discussed in Sect. 3.2. Then, we azimuthally average the tangential shear profile using N_{s} source galaxies. The tangential shear of such a deflector only depends on the radial distance to the centre and is denoted as ϵ_{t}. We can then define an estimator for the lensing signal X, including for now arbitrary factors $w(\theta_i) = w_i$

$$X = \sum_{i=1}^{N_{\mathrm{s}}} w_i\, \epsilon_{\mathrm{t},i} , \tag{4.29}$$

where X vanishes when there is no lensing signal. Using (3.18), the expectation values of X and X^2 of an SIS yield

$$E(X) = \theta_{\mathrm{E}} \sum_{i=1}^{N_{\mathrm{s}}} \frac{w_i}{2\theta_i} , \tag{4.30}$$

$$\begin{aligned} E(X^2) &= \sum_{i,j=1}^{N_{\mathrm{s}}} w_i w_j\, E(\epsilon_{\mathrm{t},i},\epsilon_{\mathrm{t},j}) \\ &= \Big[E(X)\Big]^2 + \sigma_\epsilon^2 \sum_{i=1}^{N_{\mathrm{s}}} w_i^2 . \end{aligned} \tag{4.31}$$

Then the S/N becomes

$$\frac{\mathrm{S}}{\mathrm{N}} = \frac{\theta_\mathrm{E}}{2\sigma_\epsilon} \frac{\sum_{i=1}^{N_\mathrm{s}} w_i/\theta_i}{\sqrt{\sum_{i=1}^{N_\mathrm{s}} w_i^2}} \,. \tag{4.32}$$

The S/N maximises for factors $w_i \propto 1/\theta_i$, as is seen when (4.32) is differentiated with respect to w_j. Inserting this choice of the w_i in (4.31) yields the sum $\left\langle \sum_i \theta_i^{-2} \right\rangle$ which can be replaced by its ensemble average over the annulus

$$\begin{aligned} \left\langle \sum_i \theta_i^{-2} \right\rangle = N_\mathrm{s} \left\langle \theta_i^{-2} \right\rangle &= \frac{2N_\mathrm{s}}{\theta_\mathrm{out}^2 - \theta_\mathrm{in}^2} \int_{\theta_\mathrm{in}}^{\theta_\mathrm{out}} \theta \mathrm{d}\theta \; \theta^{-2} \\ &= 2n_\mathrm{s}\pi \ln(\theta_\mathrm{out}/\theta_\mathrm{in}) \,. \end{aligned} \tag{4.33}$$

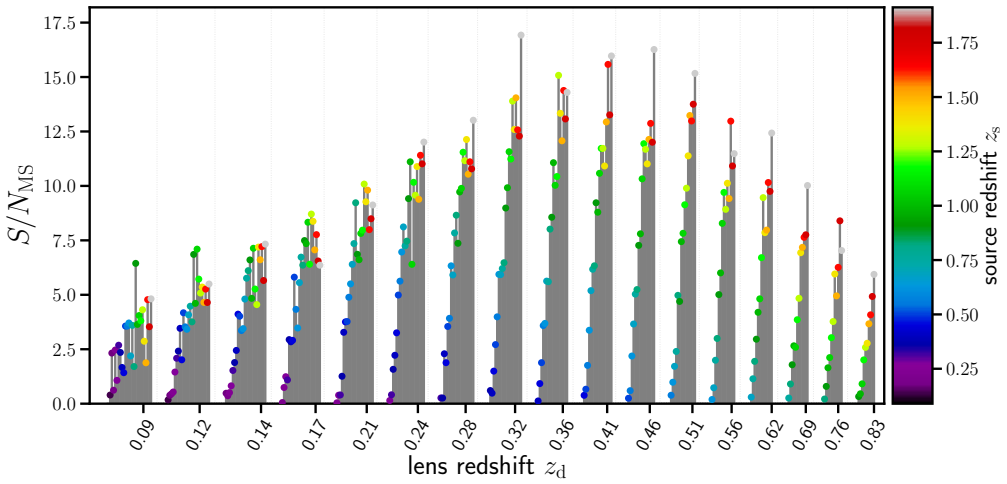
Finally, substituting the Einstein angle (3.4) leads to

$$\begin{aligned} \frac{\mathrm{S}}{\mathrm{N}} &= \frac{\theta_\mathrm{E}}{\sqrt{2}\sigma_\epsilon} \sqrt{\pi n_\mathrm{s}} \sqrt{\ln(\theta_\mathrm{out}/\theta_\mathrm{in})} \\ &= 0.7 \sqrt{\frac{n_\mathrm{s}}{35\,\mathrm{arcmin}^{-2}}} \left(\frac{\sigma_\epsilon}{0.3}\right)^{-1} \left(\frac{\sigma_v}{200\,\mathrm{km/s}}\right)^2 \sqrt{\frac{\ln(\theta_\mathrm{out}/\theta_\mathrm{in})}{\ln 10}} \left\langle \frac{D_\mathrm{ds}}{D_\mathrm{s}} \right\rangle \,, \end{aligned} \tag{4.34}$$

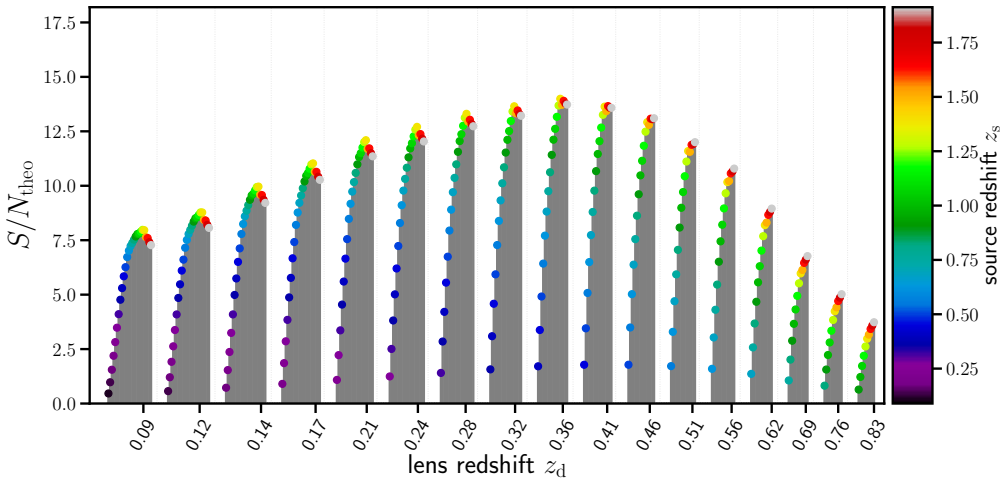
for a single deflector with typical velocity dispersion $\sigma_v \sim 200\,\mathrm{km/s}$. All values have been chosen as used in this thesis. It can be seen that the GGL signal cannot be detected for a single lens galaxy; a stacking of lensing signals increases the S/N by $\sqrt{N_\mathrm{d}}$. Thus, at least a hundred lenses are needed for a significant detection.

In Fig. 4.7 the resulting S/N from the MS for various lens-source combinations is shown compared to results from Eq. (4.34). The signal from the MS is the integrated shear profile in an annulus from $\theta_\mathrm{in} = 0\farcm6$ to $\theta_\mathrm{out} = 17\farcm5$ estimated as in Eq. (4.21) with weighting function (4.18). As before, lens galaxies brighter than 22 mag in the r-band are used. The mean velocity dispersion $\langle\sigma_v\rangle$ is not calculated by the SAM, but it is related to the stellar mass M_* by $\sigma_v \propto M_*^\alpha$ with slope $\alpha \approx 0.3$ (Zahid et al., 2016). We use fitting parameters from Zahid et al. (2016) that are listed in their Table 1 (rows with labelling 'SHELS') and shown in their Fig. 7.[4] The number of source galaxies is equal to the expected number of sources with a magnitude cut of 24.5 mag in the r-band such that the cumulative number density is $n_\mathrm{s} = 35\,\mathrm{arcmin}^{-2}$; the source positions are then randomly distributed. The shape noise is set to $\sigma_\epsilon = 0.3$. Errors are obtained through a Jackknife method of the 64 lines-of-sight. The resulting S/N is seen in the upper panel of Fig. 4.7. The model assumes a simplified density profile and we just considered shape noise as a source of noise. Therefore, a perfect match between theory and simulated data is not expected. In general, the simulated S/N is smaller than the theoretical one due to the presence of subdominant noise terms, e.g., sample variance. Overall, theory and data show good agreement.

[4] Zahid et al. (2016) used data from the Smithsonian Hectospec Lensing Survey (SHELS) that has an r-band magnitude limit of 20.9 mag and contains central galaxies up to redshift 0.7.



(a) The S/N as obtained from the MS for various lens redshifts (see x-axis) and source redshifts (see colour code).



(b) For the same redshift combinations as in the upper panel, the theoretical S/N is calculated.

Figure 4.7. – The theoretical S/N as given in (4.34) is compared to the one obtained by data from the MS for different lens and source redshifts. In general, the theoretical S/N overestimates the true S/N but for the highest lens and source redshifts. The grey bars under the data points are provided for better visibility of the general trend.

CHAPTER 5

MAGNIFICATION BIAS IN THE SHEAR-RATIO TEST

Using the numerical tool `ShearCo` *described in Sect. 4.2, the influence from magnification on the so-called SRT is explored. This chapter is complemented by Appendix A by studying the break-down of the thin-lens approximation (TLA) and its influence on the SRT. The following chapter was first published as Unruh et al. (2019) in the journal Astronomy & Astrophysics, Volume 623, pages A94-A103, reproduced with permission ©ESO.*

Abstract: Using the same lens galaxies, the ratios of tangential shears for different source galaxy redshifts are equal to the ratios of their corresponding angular-diameter distances. This is the so-called shear-ratio test (SRT) and it is valid when effects induced by the intervening large-scale structure (LSS) can be neglected. The dominant LSS effect is magnification bias which, on the one hand, induces an additional shear, and on the other hand, causes a magnification of the lens population. Our objective is to quantify the magnification bias for the SRT and show an easy-to-apply mitigation strategy that does not rely on additional observations. We use ray-tracing data through the Millennium Simulation to measure the influence of magnification on the SRT and test our mitigation strategy. Using the SRT as a null-test, we find deviations from zero up to 10% for a flux-limited sample of lens galaxies, which is a strong function of lens redshift and the lens-source line-of-sight separation. Using our mitigation strategy, we can improve the null-test by a factor of $\sim$100.

5.1. Introduction

Light bundles from background galaxy images get coherently distorted as they travel through the inhomogeneous Universe. We refer to that effect as gravitational lensing. The distortion includes a change of the intrinsic shape of the galaxies (lensing shear) as well as a magnification effect which affects the observed number density of galaxies. In the regime of weak gravitational lensing, the change of galaxy shapes cannot be seen for single galaxies. A statistical approach is needed where information is extracted from thousands to millions of lens galaxies. Amplitude and direction of the distortion depend on the integrated tidal gravitational field along the line-of-sight as well as the curvature of the Universe, which makes weak lensing measurements a powerful cosmological probe (see, e.g., Kilbinger, 2015, for a recent review). Arguably, it is the most powerful method to constrain the equation of state of Dark Energy (Albrecht et al., 2006). Ongoing surveys like the Hyper Suprime-Cam SSP Survey (Aihara et al., 2018), Kilo Degree Survey (KiDS) (Kilo Degree Survey, de Jong et al., 2017) and Dark Energy Survey (DES) (Abbott et al., 2018) already put tight constraints on cosmological parameters (Troxel et al., 2018; Hildebrandt et al., 2017) especially in combination with other probes (van Uitert et al., 2018; Joudaki et al., 2018; Harnois-Déraps et al., 2017). In the future even bigger projects are planned with the *Euclid* mission (Laureijs et al., 2011), the Square Kilometre Array Observatory (SKA, Blake et al., 2004), and the Large Synoptic Survey Telescope (LSST, Ivezić et al., 2019).

The correlation between the positions of a foreground lens population and the shear of more distant background source galaxies has been named galaxy-galaxy lensing (GGL; see, e.g., Hoekstra, 2013). The excess projected mass around lens galaxies within an aperture (Squires and Kaiser, 1996; Schneider, 1996) is reflected by the so-called tangential shear γ_t. In 2003, Jain & Taylor proposed the shear-ratio test (SRT) as a purely geometrical probe for cosmology. If the maximum separation over which galaxies are correlated with the large-scale structure (LSS) is small compared to the angular-diameter distance between lens and source, the ratio of tangential shear values from two different source populations with the same lens population does only depend on a ratio of angular-diameter distances. Hence, there is no contribution from the lens properties anymore, while the distance ratios depend on cosmology through the distance-redshift relation. Therefore, shear-ratios were originally constructed for probing cosmological parameters (Jain and Taylor, 2003; Bernstein and Jain, 2004). However, as it turns out, the dependence on cosmology is rather weak (Taylor et al., 2007, Zhao & Schneider in prep.) and correspondingly, first applications of this SRT returned only weak constraints (Kitching et al., 2007; Taylor et al., 2012). Alternative probes, e.g., CMB (*Planck* Collaboration et al., 2016), Supernovae Type 1a (Scolnic et al., 2018), and cosmic shear (Hildebrandt et al., 2017), provide far more accurate constraints on cosmological parameters, which essentially renders the SRT non-competitive for its original purpose. Yet, we

can turn the argument around and use the SRT as a null test to detect remaining systematics (see, e.g., Schneider, 2016; Prat et al., 2018).

The main challenge in ongoing and future weak lensing surveys is to obtain an unbiased estimate of shear from faint background galaxies and their corresponding redshift distributions. Great efforts to understand what influences/biases the data beside shape distortion by the weak lensing effect (see, e.g., Mandelbaum, 2018; Harnois-Déraps et al., 2017; Amon et al., 2018; Zuntz et al., 2018; Abbott et al., 2018) uncovered, e.g., noise bias (Bartelmann et al., 2012; Melchior and Viola, 2012), underfitting bias (Voigt and Bridle, 2010), and intrinsic alignment effects (Troxel and Ishak, 2015). Moreover, photometric redshift estimates suffer from dust obscuration in galaxies, the lack of a sufficient number of spectroscopic galaxy redshifts for the calibration, and a limited number of spectral bands and galaxy template spectra. Since biases in the data, if uncorrected, can be a magnitude larger than the effects from weak lensing, different strategies including null tests have been proposed to check for remaining systematics. Commonly used null tests are correlations between corrected galaxy shapes and uncorrected stellar ellipticity (e.g., Heymans et al., 2012) or other quantities that should be independent of shear, as well as checking for B-mode statistics with the cross shear or using COSEBIs (Complete Orthogonal Sets of E-/B-mode Integrals, Schneider et al., 2010; Asgari et al., 2017). The SRT emerged among other more recently introduced probes for systematics (see, e.g., Cai et al., 2016; Li et al., 2017, for alternative probes).

Magnification bias in GGL and the SRT has been considered by Ziour and Hui (2008). Hilbert et al. (2009) showed that magnification bias suppresses the GGL signal expected from shear that is induced by matter correlated with the lens galaxies, by up to 20%. For galaxies in the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) the magnification bias has a confirmed impact of $\sim 5\%$ (Simon and Hilbert, 2018). The LSS between us and the lens will shear source galaxies. While this shear is not correlated with the true positions of the lens galaxies, magnification by the LSS also affects the observed number density of lenses and thus leads to a correlation between observed lens positions and source shear (Bartelmann and Schneider, 2001). The induced correlation between foreground and background galaxy populations has already been measured in different surveys (see, e.g., Scranton et al., 2005; Garcia-Fernandez et al., 2018). Despite the fairly large effect of this magnification bias on the GGL signal shown by Hilbert et al. (2009), it appears to have been neglected in (almost) all observational studies of GGL and their quantitative interpretation.

In this paper, we will investigate the effect of magnification bias on the SRT. Since magnification bias is a function of source and lens redshift, the SRT can fail even if shear and redshift data are sound (Ziour and Hui, 2008). We use simulated data to quantify the magnification bias in the SRT, describe its properties and, most importantly, we will present a simple mitigation strategy. The main advantage of this mitigation method is that no additional measurements are needed. We also consider

a second mitigation strategy that employs stellar velocity dispersion measurements to estimate the magnification.

This paper is organised as follows: in Sect. 5.2 we will revisit the basics of the SRT in more detail. In Sect. 5.3 we will briefly describe our synthetic lensing data taken from ray-tracing through the Millennium Simulation as well as the mock lens catalogue. The effects of the magnification bias and its mitigation strategies will be described in Sect. 5.4 and Sect. 5.5. We will conclude in Sect. 5.6.

5.2. Weak lensing and the shear-ratio test

5.2.1. Cosmological distances

The comoving distance $\chi(z_1, z_2)$ of a source at redshift z_2 from a lens at redshift z_1 is given by an integration over the Hubble parameter $H(z)$ as a function of redshift z

$$\chi(z_1, z_2) = \int_{z_1}^{z_2} \frac{c\,\mathrm{d}z'}{H(z')} \,, \tag{5.1}$$

where

$$\left(\frac{H(z)}{H_0}\right)^2 = \Omega_\mathrm{m}(1+z)^3 + (1-\Omega_\mathrm{m}) \,, \tag{5.2}$$

for a flat Universe with matter density Ω_m in units of the present-day critical density $\rho_{\mathrm{crit},0} = 3H_0^2/(8\pi G)$. The comoving distance χ is related to the angular-diameter distance D via:

$$D(z_1, z_2) = \frac{\chi(z_1, z_2)}{1+z_2} \,. \tag{5.3}$$

Note that $\chi(z_1, z_2) = \chi(0, z_2) - \chi(0, z_1)$ but $D(z_1, z_2) \neq D(0, z_2) - D(0, z_1)$ except for $z_2 - z_1 \ll 1$. In the following, we will omit the argument zero in the distances, i.e. $D(0, z) := D(z)$.

5.2.2. Galaxy-galaxy lensing

Foreground matter at a redshift z_d will deflect light rays from background galaxies and induce a shear pattern. In complex notation the shear reads

$$\gamma(\boldsymbol{\theta}) = \gamma_1(\boldsymbol{\theta}) + \mathrm{i}\gamma_2(\boldsymbol{\theta}) \,. \tag{5.4}$$

Here, $\boldsymbol{\theta}$ is the position on the sky and $\gamma_{1,2}$ are the Cartesian shear components at angular position $\boldsymbol{\theta}$. In GGL, the shear is measured with respect to the connecting line between a lens at position $\boldsymbol{\theta}_\mathrm{d}$ and a source galaxy – orthogonal to that line is

the tangential shear γ_t, and the cross shear $\gamma_\times$ is measured with a 45°-rotation. For a fixed lens position $\boldsymbol{\theta}_\mathrm{d}$ this corresponds to a rotation of the shear components

$$\gamma_\mathrm{t}(\boldsymbol{\theta}) + \mathrm{i}\gamma_\times(\boldsymbol{\theta}) = -\gamma(\boldsymbol{\theta})\,\frac{(\boldsymbol{\theta}-\boldsymbol{\theta}_\mathrm{d})^*}{\boldsymbol{\theta}-\boldsymbol{\theta}_\mathrm{d}}\;, \tag{5.5}$$

where we also conveniently write the position on the sky in complex notation, i.e. $\boldsymbol{\theta} = \theta_1 + \mathrm{i}\theta_2$, and an asterix denotes complex conjugation.

Shear is caused by a foreground line-of-sight overdensity. The lensing strength factorises into a part containing all the lens properties, and one containing the angular-diameter distances between us and source, $D(z_\mathrm{s}) = D_\mathrm{s}$, as well as between lens and source, $D(z_\mathrm{d}, z_\mathrm{s}) = D_\mathrm{ds}$ (see, e.g., Schneider et al., 1992). The lens properties are characterised by its matter distribution and its angular-diameter distance, D_d. Shear that is caused by the matter associated with the lens galaxies will be denoted with γ_g, where the 'g' refers to galaxy. The expectation value of this tangential shear measurement can be written as

$$\gamma_\mathrm{g}(\boldsymbol{\theta}; z_\mathrm{d}, z_\mathrm{s}) = \gamma_{\mathrm{g},\infty}(\boldsymbol{\theta}; z_\mathrm{d})\,\frac{D_\mathrm{ds}}{D_\mathrm{s}} := \gamma_{\mathrm{g},\infty}(\boldsymbol{\theta}; z_\mathrm{d})\,\beta(z_\mathrm{d}, z_\mathrm{s})\;, \tag{5.6}$$

where z_d is the redshift of the lens for background sources at redshift z_s. The lensing efficiency β is a ratio of angular-diameter distances which is scaled by $\gamma_{\mathrm{g},\infty}$. If not noted otherwise further expressions of shear in this paper are tangential shear estimates.

5.2.3. The classical shear-ratio test

We can calculate a weighted integral to obtain a mean shear estimate $\bar{\gamma}_\mathrm{g}$

$$\bar{\gamma}_\mathrm{g}(\boldsymbol{\theta}) = \int \mathrm{d}^2\theta'\,\gamma_\mathrm{g}(\boldsymbol{\theta}+\boldsymbol{\theta}')w(|\boldsymbol{\theta}'|)\;, \tag{5.7}$$

where

$$w(\theta) = \frac{1}{2\pi\theta^2}\,\mathcal{H}(\theta-\theta_\mathrm{in})\,\mathcal{H}(\theta_\mathrm{out}-\theta)\;, \tag{5.8}$$

is a weight function different from zero only in the annulus $\theta_\mathrm{in} \leq \theta \leq \theta_\mathrm{out}$ and where $\mathcal{H}$ is the Heaviside step function. This form of the weight function optimises the signal-to-noise ratio for a shear profile behaving like $1/\theta$, as is the case for an isothermal profile (Bartelmann and Schneider, 2001).

If the tangential shear can be factorised as in Eq. (5.6), i.e. into a factor that depends only on lens properties and the lensing efficiency, we can consider shear measurements from two different source populations at z_i and z_j behind the same lens galaxy or a population of lens galaxies at fixed redshift z_d. Then the ratio, R, of those shear measurements is independent of the lens properties and is solely

determined by the geometry of the observer-lens-source system as a ratio of lensing efficiencies,

$$\frac{\bar{\gamma}_{\mathrm{g}}(z_{\mathrm{d}}, z_j)}{\bar{\gamma}_{\mathrm{g}}(z_{\mathrm{d}}, z_i)} = \frac{\beta(z_{\mathrm{d}}, z_j)}{\beta(z_{\mathrm{d}}, z_i)} =: R(z_{\mathrm{d}}; z_i, z_j) \; . \tag{5.9}$$

Equation (5.9) can be written as a null test which can be applied in cosmic shear measurements as a consistency check that does not require any additional data. This method is purely based on geometrical considerations and as such it is independent of structure growth in the Universe. This makes the SRT easy-to-apply since it does not require the use of simulations.

5.3. Mock data

5.3.1. Millennium simulation data

In this work we make use of the Millennium Simulation (MS, Springel et al., 2005). The MS is an N-body simulation tracing the evolution of 2160^3 dark matter particles of mass $8.6 \times 10^8 \, h^{-1}\mathrm{M}_\odot$ enclosed in a $(500 \, h^{-1}\mathrm{Mpc})^3$-cube, where h is the dimensionless Hubble parameter defined as $H_0 = 100h \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$. In the MS, 64 snapshots are available in the redshift range from $z = 127$ to today. The underlying cosmology is a flat ΛCDM cosmology with matter density parameter $\Omega_{\mathrm{m}} = 0.25$, baryon density parameter $\Omega_{\mathrm{b}} = 0.045$, dark energy density parameter $\Omega_\Lambda = 1 - \Omega_{\mathrm{m}} = 0.75$, a dimensionless Hubble parameter $h = 0.73$, a scalar spectral index $n_{\mathrm{s}} = 1$ and a power spectrum normalization of $\sigma_8 = 0.9$. These values agree with a combined analysis of 2dFGRS (Colless et al., 2001) and first-year WMAP data (Spergel et al., 2003).

Various catalogues of galaxies have been added to the simulation using semi-analytic galaxy-formation models. Saghiha et al. (2017) showed that the galaxy catalogue from the model by Henriques et al. (2015) matches best with the observed GGL and galaxy-galaxy-galaxy lensing signal from the CFHTLenS. We use lens galaxies from the redshift slices 59 to 43 corresponding to $z_{59} = 0.0893$ and $z_{43} = 0.8277$.

Furthermore, we use simulated lensing data obtained by a multiple-lens-plane ray-tracing algorithm in 64 realizations with a $4 \times 4 \, \mathrm{deg}^2$-field-of-view (Hilbert et al., 2009), where we consider the source redshift planes 58 to 34 that correspond to redshifts $z_{58} = 0.1159$ and $z_{34} = 1.9126$. To perform the ray-tracing, the matter distribution in each redshift slice is carefully mapped to the midplane. Then, a multiple lens plane algorithm is used to calculate shear, magnification and convergence information on a grid in each midplane. With a box size of $500 \, h^{-1}\mathrm{Mpc}$, the MS is not large enough to contain a full light cone out to high redshifts. However, a simple stacking of simulation cubes would result in a light ray that meets the same matter

structures several times on its way to $z = 0$ due to periodicity. Therefore, Hilbert et al. (2009) decided to perform the ray-tracing using a skewed angle through the box, yet making use of its periodic boundary conditions. Hence, no random rotation or translation of the matter in the box has been done, which preserves the galaxy-matter correlation. Effectively, a light ray can travel a comoving distance of $5\,h^{-1}$Gpc before encountering the same matter structures.

To avoid double counting, the indices in Eq. (5.9) will be restricted to $i < j$ since $R(z_\mathrm{d}; z_i, z_j) = 1/R(z_\mathrm{d}; z_j, z_i)$. We will also only use consecutive redshift bins for taking ratios due to the relation $R(z_\mathrm{d}; z_i, z_k) = R(z_\mathrm{d}; z_i, z_j)\,R(z_\mathrm{d}; z_j, z_k)$.

5.3.2. Obtaining a tangential shear estimate

We can calculate the weighted mean tangential shear

$$\bar{\gamma}_\mathrm{t}(\boldsymbol{\theta}) = -\Re\left[\int \mathrm{d}^2\theta'\,\gamma(\boldsymbol{\theta}+\boldsymbol{\theta}')\,\frac{\boldsymbol{\theta}'^*}{\boldsymbol{\theta}'}w(|\boldsymbol{\theta}'|)\right]\,, \tag{5.10}$$

for every grid point of the $4\times4\,\mathrm{deg}^2$-field. It defines a convolution which reduces to a simple multiplication in Fourier space. We thus use fast Fourier Transform (FFT) from the library of Frigo and Johnson (2012) to compute the shear estimator (5.10). An FFT implicitly assumes periodic boundary conditions. Since we convolve the shear field with a function of finite support, results from the FFT will be wrong in a stripe of the thickness θ_out around the field edge. Thus, we do not consider shear data in this stripe and focus on the data in the inner $(4^\circ - 2\theta_\mathrm{out}) \times (4^\circ - 2\theta_\mathrm{out})$ area of the field. For the inner boundary of the annulus we will set $\theta_\mathrm{min} = 0\farcm5$ in the paper if not explicitly noted otherwise. The exact value of θ_min is not crucial, since the weighting function will give the same shear signal per logarithmic bin if the mass profile is isothermal. However, we cannot go to arbitrarily small θ_in due to the finite resolution of the simulation.

We extract the positions of galaxies from the Henriques et al. (2015) model catalogue that fulfill our selection criteria, e.g., a flux limit or a cut in halo mass. The positions of the N_L galaxies are then assigned to their nearest grid point. Each pixel in the grid has a size of $(3.5\,\mathrm{arcsec})^2$. Thus, as long as we choose the size of our integration area in (5.10) in the square arcminute regime, we will not suffer from discretization effects. We then simply average the shear signal (5.10) for all lens galaxies

$$\langle\bar{\gamma}_\mathrm{t}\rangle = \frac{1}{N_\mathrm{L}}\sum_{i=1}^{N_\mathrm{L}}\bar{\gamma}_\mathrm{t}(\boldsymbol{\theta}_i)\,. \tag{5.11}$$

When we apply a cut at 24 mag in the r-band, we typically find 1500 lens galaxies per field in the lowest redshift slice and around 30 000 lens galaxies in the highest redshift slice. To reduce the statistical error, we use the shear information on all

4096^2 pixel (in contrast to using shear information only on those positions where a source galaxy is located according to the Henriques galaxy catalogue).

We repeat the whole process, but this time we randomise the lens galaxy positions. Singh et al. (2017) showed that subtracting the signal around random points from the actual shear signal usually leads to a more optimal estimator with a decreased error budget. Although this poses only a minor contribution in our case, we change our shear estimator to $\bar{\gamma}_\mathrm{t} \rightarrow \bar{\gamma}_\mathrm{t} - \bar{\gamma}_\mathrm{t,rand}$.

5.4. The magnification bias

5.4.1. The conventional shear-ratio test

We calculate ratios of observed shear estimates and consider the null-test

$$\left\langle \frac{\bar{\gamma}_\mathrm{g}(z_\mathrm{d}, z_j)}{\bar{\gamma}_\mathrm{g}(z_\mathrm{d}, z_i)} \right\rangle - \frac{\beta(z_\mathrm{d}, z_j)}{\beta(z_\mathrm{d}, z_i)} = 0 \,. \tag{5.12}$$

We use the shear-ratios for all statistically independent redshift bins and show the result of the SRT in Fig. 5.1. Since our method is unaffected by shape noise and Poisson noise of source galaxies, we expect only minor deviations from zero in the SRT. However, this is not the case. The deviation from zero gets worse for higher lens redshift and smaller separations between lens and sources. In the right figure, we applied a cut in magnitude which includes the magnification (and is thus the observed magnitude), in contrast to the left figure where we used all available galaxies in the catalogue which is equivalent to a stellar mass-limited sample. In the lowest-redshift bins (i.e. $z_\mathrm{d} < 0.15$) the results are almost identical since $> 95\%$ of the lenses are brighter than 24 mag in the r-band. For these redshift bins, the classical SRT (5.12) performs as expected. In the medium-redshift range $0.15 \leq z_\mathrm{d} \leq 0.4$ a sensible choice of lens and source redshifts will keep the deviations small. The effect for the magnitude-limited sample is quantitatively smaller but qualitatively similar. The bias is strongest for high-redshift lenses with $z_\mathrm{d} > 0.4$ where even widely separated lenses and sources show deviations from zero at the percent level.

5.4.2. Magnification effects

Until now, we have only considered the shear caused by the matter that is associated with the lens galaxies at fixed redshift z_d. However, there exists intervening LSS between us and the source that can induce an additional shear signal. This shear is not correlated with the true positions of lens galaxies. However, the LSS also alters the distribution of lens galaxies on the sky by magnification effects. This leads to a correlation between shear caused by the LSS and the *observed* distribution of galaxies. For most practical purposes magnification bias is the dominant second-order effect (Hui et al., 2007).

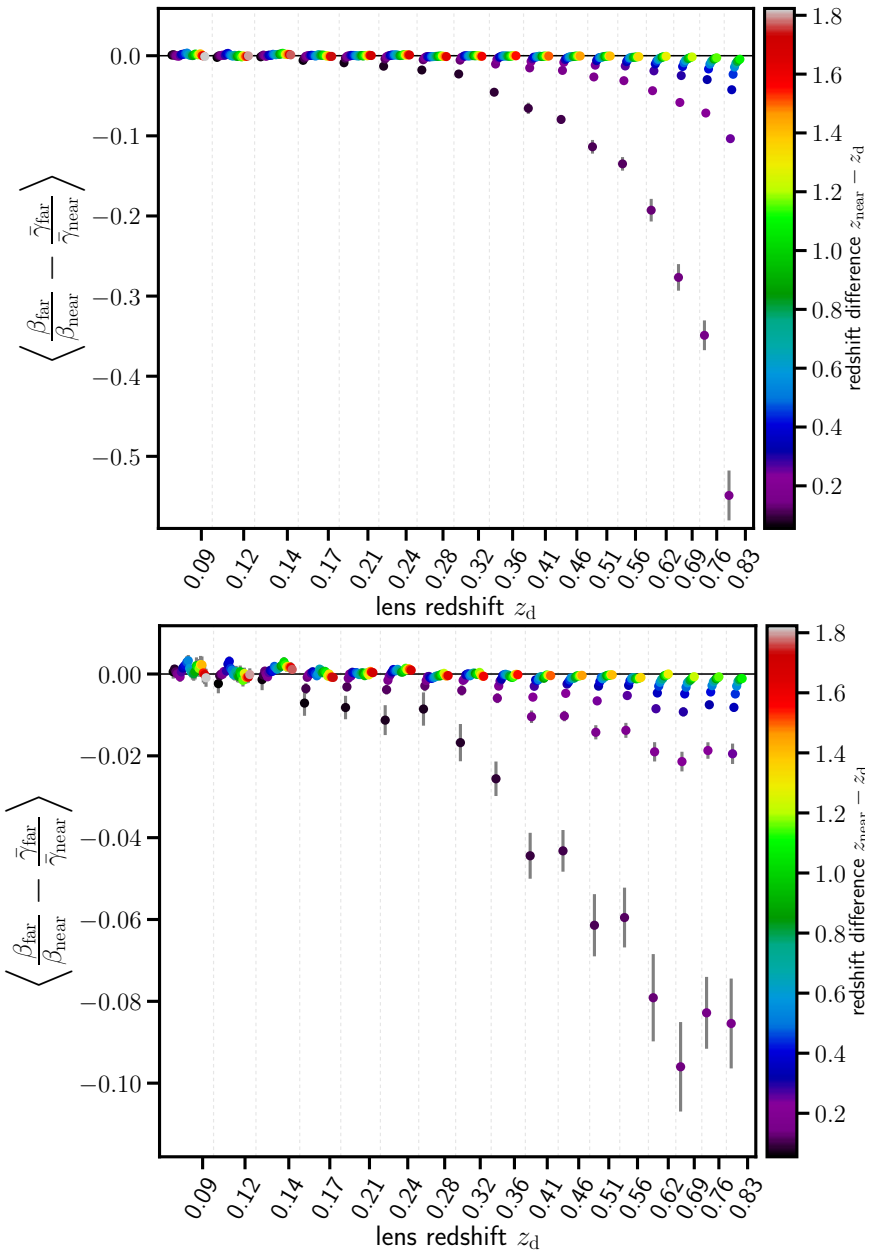


Figure 5.1. – Shear-ratio test of two populations of galaxies that are sheared by the same lens, i.e. $z_{\rm far} > z_{\rm near} > z_{\rm d}$ with source redshifts in consecutive bins. The dotted lines separate shear-ratios from different lens galaxies, while source redshifts increase from left to right, which is highlighted by colour code. Each color represents a combination of two sources at $z_{\rm near}$ and $z_{\rm far}$. Ideally, the outcome of the SRT (5.12) is zero. Data are taken from ray-tracing through the MS. The shear estimator $\langle \bar{\gamma}_{\rm t} \rangle$ is defined in Eq. (5.11) with $\theta_{\rm in} = 0.'5$ and $\theta_{\rm out} = 5'$. Errors are obtained by a Jackknife method from the 64 different realizations per redshift. **Upper**: all simulated galaxies have been used. **Lower**: a magnitude cut of 24 mag in the r-band is applied for the lens galaxies.

We can approximate the influence of magnification by the LSS with

$$\bar{\gamma}_{\mathrm{t}}(z_{\mathrm{d}}, z_{\mathrm{s}}) = \bar{\gamma}_{\mathrm{g}}(z_{\mathrm{d}}, z_{\mathrm{s}}) + \bar{\gamma}_{\mathrm{LSS}}(z_{\mathrm{s}}) \; . \tag{5.13}$$

The shear at the source redshift is a superposition of the shear induced by matter associated with the lens galaxies at redshift z_{d} and the LSS between us and the source galaxies. An average of the shear contribution from the LSS over a sufficient number of source galaxy must be zero. However, we use lens galaxies to obtain an averaged tangential shear estimate and the observed position of these galaxies is altered by the intervening LSS. This leads to a correlation between LSS-induced shear and lens galaxy positions:

$$\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{s}}) = \langle\bar{\gamma}_{\mathrm{g}}\rangle(z_{\mathrm{d}}, z_{\mathrm{s}}) + \langle\bar{\gamma}_{\mathrm{LSS}}\rangle(z_{\mathrm{d}}, z_{\mathrm{s}}) \; . \tag{5.14}$$

The relative contribution of the foreground LSS to the observed lensing signal depends on the size of the annulus. We expect that for a small annulus close to the lens position, the mean lensing signal is dominated by shear associated with matter at the lens redshift z_{d}. Hence, we investigate the impact of θ_{out} for two different θ_{in} for the shear estimator $\bar{\gamma}_{\mathrm{t}}$. We concentrate on the realistic case of a flux-limited sample and for each θ_{in} we vary θ_{out} for an SRT as done in Fig. 5.1. For clarity we only plot the SRT for the lens-source-source combination $z_{\mathrm{d}} = 0.83$, $z_{\mathrm{near}} = 0.91$ and $z_{\mathrm{far}} = 0.99$ which corresponds to one the largest deviations from zero in our SRT. The result shown in Fig. 5.2 follows our expectations. The deviation from zero is less pronounced for small annuli but it is still present. For larger integration ranges, the deviation is larger, but stays approximately constant for $\theta_{\mathrm{out}} \geq 5'$.

We verify that it is indeed magnification that affects the shear estimate around galaxies. Since the influence grows with redshift, with smaller redshift differences of lens and source galaxies (Fig. 5.1), and with the size of the annulus (Ziour and Hui, 2008, see Fig. 5.2), we choose (again) our highest redshift bins with $z_{\mathrm{d}} = 0.83$ and the two consecutive redshift bins as source galaxies with $z_{\mathrm{near}} = 0.91$ and $z_{\mathrm{far}} = 0.99$. We set the integration range to $\theta_{\mathrm{out}} = 17\farcm5$. We bin the lens galaxies in magnification such that each bin contains a roughly equal number of lenses. Then, we measure the shear in the two source planes, $\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{near}})$ and $\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{far}})$, for each bin and plot it against the average magnification per bin. Furthermore, we want to visualise the influence the LSS has on the measured shear result. Since we chose lens and source plane close to each other, almost all the relevant LSS is also in front of the lenses. Then, we can just measure the shear signal $\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{d}})$ around the lens galaxies at the lens redshift to get an approximate measurement of $\langle\bar{\gamma}_{\mathrm{LSS}}\rangle(z_{\mathrm{d}}, z_{\mathrm{s}})$.

The result is displayed in Fig. 5.3, where we also plot the mean shear for all lens galaxies irrespective of magnification as dashed lines. As expected from the form of the lensing kernel, the shear from sources with higher redshift $\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{far}})$ is larger than $\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{near}})$. Shear and magnification show a clear correlation. The red dashed line, however, is naively expected to be consistent with zero, as it would

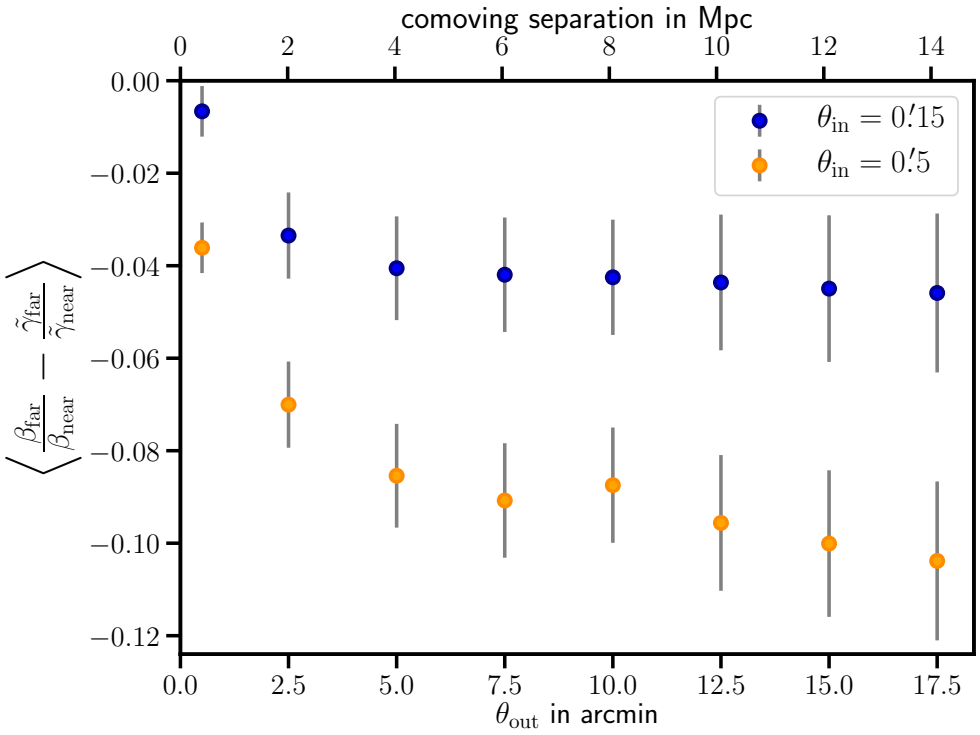


Figure 5.2. – For an identical setup as the right Fig. 5.1 we perform the SRT; we vary the integration range by choosing two different $\theta_{\rm in}$ and by altering $\theta_{\rm out}$ in the shear estimator (5.10) for each $\theta_{\rm in}$. The leftmost points are shear estimates around a thin annulus at $\theta = 0.'15$ (blue) and $\theta = 0.'5$ (orange). For clarity only the outcome of the SRT for the combination $z_{\rm d} = 0.83$, $z_{\rm near} = 0.91$ and $z_{\rm far} = 0.99$ is shown. The SRT differs strongly from zero regardless of $\theta_{\rm out}$, only at a thin annulus very close to the lens is the signal almost compatible with zero.

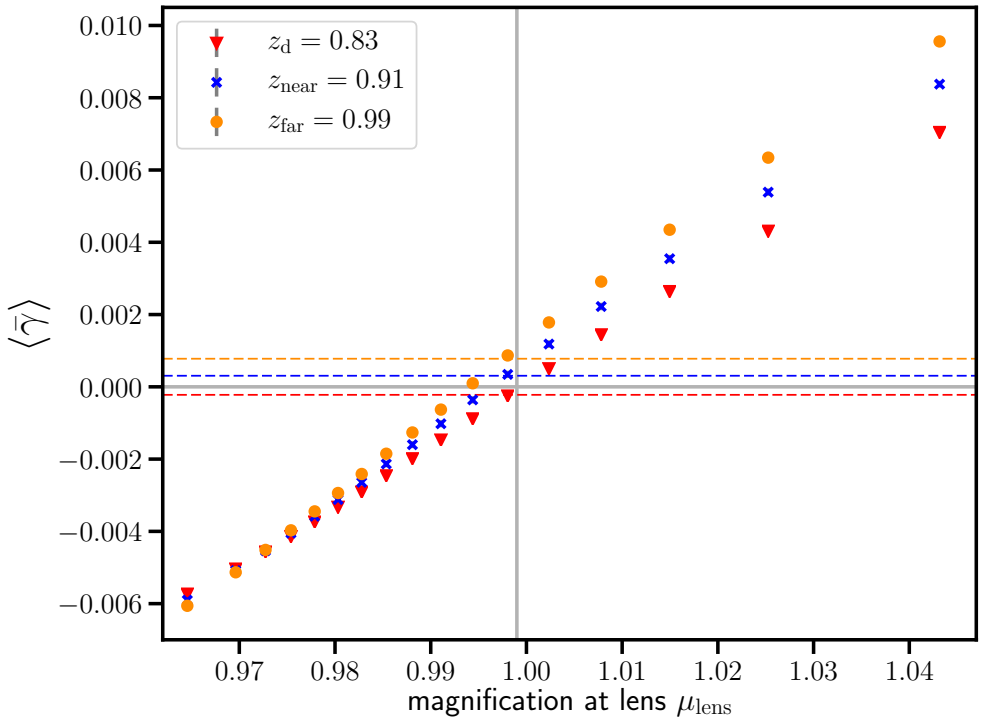


Figure 5.3. – The shear estimate $\langle\bar{\gamma}_\mathrm{t}\rangle(z_\mathrm{d} = 0.83, z_i)$ with $\theta_\mathrm{in} = 0.'5$ and $\theta_\mathrm{out} = 17.'5$ is shown. A grey vertical line indicates the mean magnification of all lens galaxies. The lens galaxies have been binned in magnification with a roughly equal number of lenses in each bin. We omitted the two highest bins in magnification at mean magnification $\mu_\mathrm{lens} = 1.11$ and $\mu_\mathrm{lens} = 1.6$ for clarity. The dashed lines represent the mean shear estimates for all lenses. Due to the magnification effects of the intervening LSS, the shear around lens galaxies at lens redshift (red triangles) differs from zero as does the red dashed line which is an approximate measure for $\bar{\gamma}_\mathrm{LSS}$ in Eq. (5.14). A clear correlation between shear and magnification can be seen.

be the case if all the contributions to the lensing signal came from matter associated with the lens galaxies. However, there exists an additional contribution from shear caused by the LSS in front of the lens galaxies. Due to the selection of lens galaxies in the foreground (in contrast to random positions) the LSS-induced shear does not vanish. It can be seen that the red dashed line, a measure for $\bar{\gamma}_{\mathrm{LSS}}$, is of the same order of magnitude as the blue dashed line, the shear signal for sources $z_{\mathrm{s}} > z_{\mathrm{d}}$.

5.5. Mitigation strategies

As we have seen in the previous section, the influence of the magnification bias is negligible for low lens redshifts as well as for lenses and sources separated widely in redshift. This was already pointed out in Moessner and Jain (1998). However, as we approach Stage IV experiments to infer the equation of state parameter for Dark Energy, we go to higher and higher lens and source redshifts. A mitigation strategy is thus crucial. Using theoretical considerations for lensing power spectra, Ziour and Hui (2008) derived a mitigation in their Eq. (38) for the SRT. It involves the knowledge of the easily obtainable number count slope as well as the linear galaxy bias factor at of lenses as a function of mass, which is notoriously difficult to obtain. The assumption of a linear bias factor will eventually break down for small angular scales, and even on large scales, a linear bias is not necessarily a sufficiently accurate approximation (Hui et al., 2007).

In the following, we will introduce a new mitigation strategy. Its main advantage is that it does not require additional observations or simulations. As can be seen in Fig. 5.3, the shear signal $\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{d}})$ is not zero if measured at the lens redshift. However, the shear induced by matter associated with the lens galaxy $\langle\bar{\gamma}_{\mathrm{g}}\rangle(z_{\mathrm{d}}, z_{\mathrm{d}})$ is certainly zero. Thus, what we measure is due to the intervening LSS

$$\langle\bar{\gamma}_{\mathrm{t}}\rangle(z_{\mathrm{d}}, z_{\mathrm{d}}) = \langle\bar{\gamma}_{\mathrm{LSS}}\rangle(z_{\mathrm{d}}, z_{\mathrm{d}}). \tag{5.15}$$

In general, the influence of the LSS grows with redshift, and if the separation of lenses and sources is moderate, we can introduce a scaling factor $\lambda \gtrsim 1$ that parametrises this similarity as $\langle\bar{\gamma}_{\mathrm{LSS}}\rangle(z_{\mathrm{d}}, z_{\mathrm{s}}) \approx \lambda\langle\bar{\gamma}_{\mathrm{LSS}}\rangle(z_{\mathrm{d}}, z_{\mathrm{d}})$. From the form of the lensing kernel, we can deduce that the main contribution of the LSS to the shear signal is located at about half the distance between us and the source. Therefore, if we increase the source redshift only slightly, we will also increase λ slightly. Using these assumptions, it follows naturally that the SRT performs better for high source redshifts at fixed lens redshift. Relative to the shear contribution from the LSS, the shear from matter associated with the lens galaxies shows a strong dependence on source distance. Thus, the LSS-induced shear has less impact for larger line-of-sight separations of lens and sources.

5.5.1. Improved SRT – quantifying the foreground contribution

The scaling factor can be obtained from the data by correlating the observed lensing signal $\langle\bar{\gamma}_\mathrm{t}\rangle(z_\mathrm{d}, z_\mathrm{s})$ at the source redshift with the foreground part $\langle\bar{\gamma}_\mathrm{LSS}\rangle(z_\mathrm{d}, z_\mathrm{d})$. Using Eq. (5.14), we find

$$\begin{aligned}\langle\bar{\gamma}_\mathrm{t}(z_\mathrm{d}, z_\mathrm{s})\bar{\gamma}_\mathrm{LSS}(z_\mathrm{d}, z_\mathrm{d})\rangle &= \langle\bar{\gamma}_\mathrm{g}(z_\mathrm{d}, z_\mathrm{s})\bar{\gamma}_\mathrm{LSS}(z_\mathrm{d}, z_\mathrm{d})\rangle \\ &\quad + \lambda(z_\mathrm{d}, z_\mathrm{s})\,\langle\bar{\gamma}_\mathrm{LSS}(z_\mathrm{d}, z_\mathrm{d})\bar{\gamma}_\mathrm{LSS}(z_\mathrm{d}, z_\mathrm{d})\rangle\ , \end{aligned} \tag{5.16}$$

where averages are taken as in Eq. (5.11) for each lens. The first term on the right-hand side, $\langle\bar{\gamma}_\mathrm{g}\bar{\gamma}_\mathrm{LSS}\rangle$, vanishes since the shear from matter associated with the lens galaxies is uncorrelated with the shear caused by the foreground matter. Exploiting Eq. (5.15), we can calculate the scaling factor by

$$\lambda(z_\mathrm{d}, z_\mathrm{s}) = \frac{\langle\bar{\gamma}_\mathrm{t}(z_\mathrm{d}, z_\mathrm{s})\bar{\gamma}_\mathrm{LSS}(z_\mathrm{d}, z_\mathrm{d})\rangle}{\langle\bar{\gamma}_\mathrm{LSS}(z_\mathrm{d}, z_\mathrm{d})\bar{\gamma}_\mathrm{LSS}(z_\mathrm{d}, z_\mathrm{d})\rangle} = \frac{\langle\bar{\gamma}_\mathrm{t}(z_\mathrm{d}, z_\mathrm{s})\bar{\gamma}_\mathrm{t}(z_\mathrm{d}, z_\mathrm{d})\rangle}{\langle\bar{\gamma}_\mathrm{t}(z_\mathrm{d}, z_\mathrm{d})\bar{\gamma}_\mathrm{t}(z_\mathrm{d}, z_\mathrm{d})\rangle}\ . \tag{5.17}$$

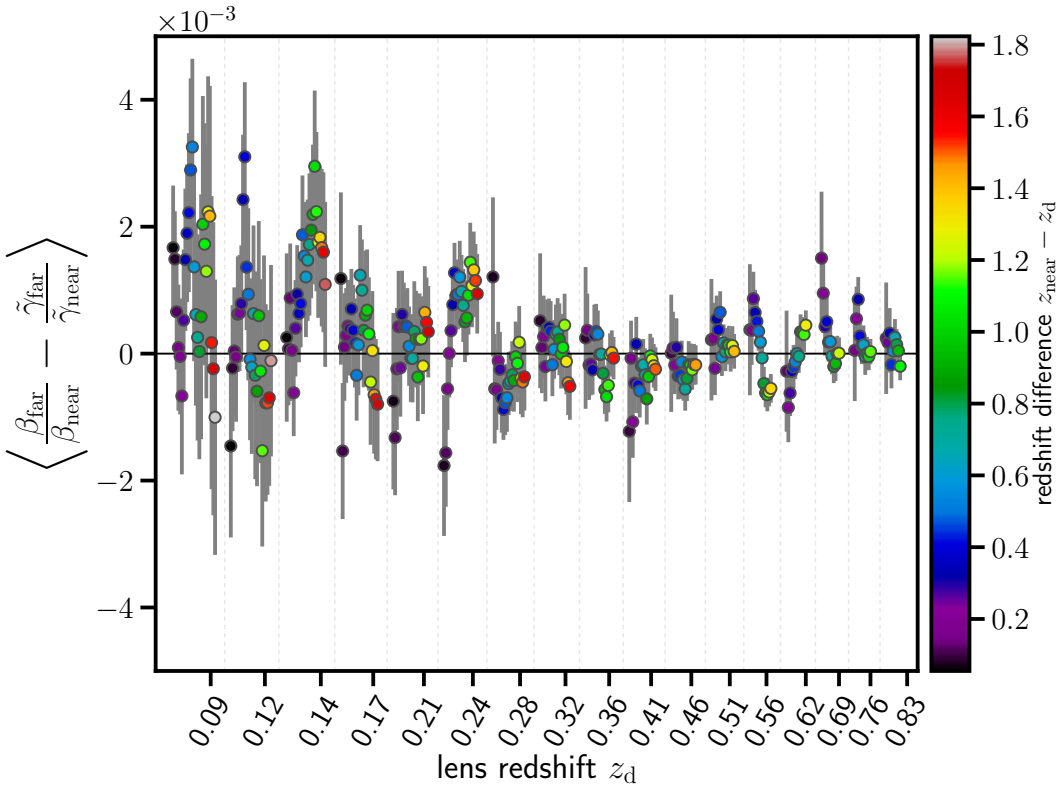


Figure 5.4. – Shown is a shear-ratio test with the same properties as the right panel of Fig. 5.1 but with a modified shear estimator (5.18) that corrects for magnification bias. While the lowest-redshift bins are almost unchanged, the high-redshift bins show a significant improvement. Note, the change of y-axis compared to the right panel of Fig. 5.1.

For brevity, we will introduce a new tangential shear estimator

$$\langle\hat{\bar{\gamma}}_\mathrm{t}\rangle(z_\mathrm{d}, z_\mathrm{s}) = \langle\bar{\gamma}_\mathrm{t}\rangle(z_\mathrm{d}, z_\mathrm{s}) - \lambda(z_\mathrm{d}, z_\mathrm{s})\langle\bar{\gamma}_\mathrm{t}\rangle(z_\mathrm{d}, z_\mathrm{d}) \,, \tag{5.18}$$

In Fig. 5.4, we show the results for the modified estimator in the realistic case of a flux-limited sample[1]. Correcting for magnification bias indeed improves the SRT by two orders of magnitude. The scaling factor λ ranges from 1.1 for redshifts adjecant to the lens redshifts to $\lambda \approx 1.7$ for $\Delta z \approx 1$ to $\lambda \approx 2.5$ for $\Delta z \approx 2$. This behaviour is fairly independent of lens redshift, while λ-values tend to be slightly higher for low lens redshifts than for high lens redshifts.

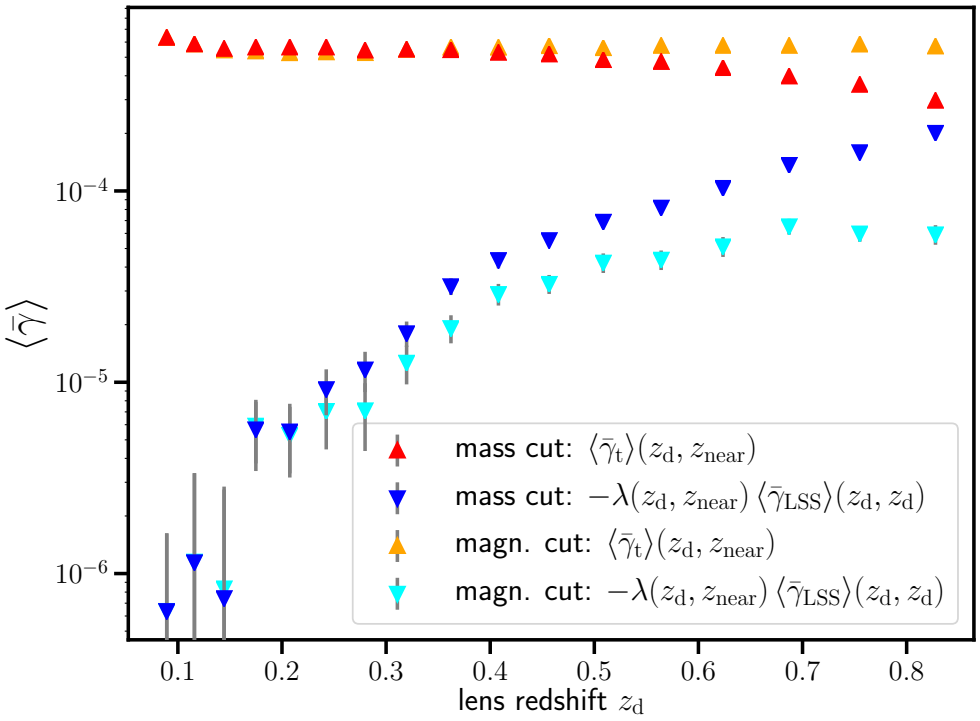


Figure 5.5. – A comparison of $\bar{\gamma}_\mathrm{t}$ and $\lambda\bar{\gamma}_\mathrm{LSS}$, where we used the shear estimator (5.11) within an annulus $\theta_\mathrm{in} = 0.'5$ and $\theta_\mathrm{out} = 5'$. Red and blue symbols correspond to a mass-limited sample while the orange and cyan symbols correspond to a magnitude cut $r = 24$. z_near is the adjacent redshift slice of the lens redshift z_d. For low redshifts the scaling factor λ is around 1.2 and it decreases to 1.1 for the high redshift bins. The contribution from the LSS, $\bar{\gamma}_\mathrm{LSS}$, to the shear is a steep function of lens redshift. Whereas for low redshifts the effect is negligible, the ratio of $\bar{\gamma}_\mathrm{LSS}/\bar{\gamma}_\mathrm{t}$ approaches 1 for the highest lens redshift bin considered. The flux-limited sample is in general less affected but shows qualitatively a similar behavior.

To visualise the contribution from magnification effects, Fig. 5.5 shows $\langle\bar{\gamma}_\mathrm{LSS}\rangle$ in comparison to the shear signal $\langle\bar{\gamma}_\mathrm{t}\rangle$ measured from the source redshift slice that is

[1]The result for the stellar mass-limited sample are almost identical to those in Fig. 5.4.

adjacent to the lens redshift slice. For low lens redshifts, $\langle\bar{\gamma}_{\rm LSS}\rangle$ makes a negligible contribution. The ratio $\langle\bar{\gamma}_{\rm LSS}\rangle/\langle\bar{\gamma}_{\rm t}\rangle$ rises with increasing redshift. The red and blue lines correspond to a mass-limited sample and show a stronger contribution by the LSS to the shear signal than the orange and cyan lines, which represent a flux-limited sample of lens galaxies; this behaviour was already seen in Fig. 5.1. The reason for this is the way magnification changes the observed number density of galaxies in the sky. For a stellar mass-limited sample the ratio of lensed galaxy number counts, n, over unlensed galaxy number counts, n_0, depends solely on the magnification, $n/n_0 = \mu^{-1}$. A flux-limited sample, on the other hand, also depends on the slope of cumulative galaxy counts evaluated at the limiting flux, $\alpha = -\left(\mathrm{d}\ln n_0/\mathrm{d}\ln s\right)|_{s_{\rm lim}}$. The ratio of lensed over unlensed galaxy number counts changes to $n/n_0 = \mu^{\alpha-1}$.

5.5.2. Alternative ways of obtaining the scaling factor λ

In the previous section, we showed that the scaling factor λ can be obtained by correlating the lensing signal at the lens redshift with that at the source redshift. We can also divide the lens population into sub-samples that show a different dependence on magnification to obtain an estimate for the scaling factor. For each of the $N \geq 2$ sub-samples Eq. (5.14) holds

$$\bar{\gamma}_{\rm t}^{i}(z_{\rm d}, z_{\rm s}) = \lambda(z_{\rm d}, z_{\rm s})\bar{\gamma}_{\rm LSS}^{i}(z_{\rm d}, z_{\rm d}) \,+\, \bar{\gamma}_{\rm g}^{i}(z_{\rm d}, z_{\rm s})\;. \tag{5.19}$$

A binning of the lenses in their magnification is, of course, dependent on magnification but unfortunately, this property is not directly observable. As proof of principle, we will show the result of this approach with the magnification readily available in the ray-tracing catalogues. We already presented the dependence of shear on magnification in Fig. 5.3 where we split the lens sample into 19 sub-samples. For a high number of lens galaxies, it is sufficient to split the lens population into two samples, for example, in samples with magnification bigger or smaller than the mean magnification. Then we can calculate the scaling factor λ with

$$\lambda = \frac{\bar{\gamma}_{\rm t}^{1} - \bar{\gamma}_{\rm t}^{2}}{\bar{\gamma}_{\rm LSS}^{1} - \bar{\gamma}_{\rm LSS}^{2}}\;, \tag{5.20}$$

where '1' corresponds to $\mu_{\rm lens} < \langle\mu\rangle$ and '2' to $\mu_{\rm lens} \geq \langle\mu\rangle$. Furthermore, we set $\bar{\gamma}_{\rm g}^{1}(z_{\rm d}, z_{\rm s}) = \bar{\gamma}_{\rm g}^{2}(z_{\rm d}, z_{\rm s})$ since the shear induced by matter correlated with the lens galaxies must be independent of the lens magnification.

We checked this approach and found similar results to our previous method for lens redshifts $z_{\rm d} > 0.15$ with slightly worse performance for high source redshifts. For the low-redshift bins, we have only a comparatively low number of galaxies available, while at the same time the average of the LSS-induced shear $\bar{\gamma}_{\rm LSS}$ is of the order 10^{-6}, which is a factor of 1000 lower than the actual shear signal. This leads to a very noisy estimate of the scaling factor λ.

Magnification estimates are challenging since the intrinsic variation of galaxy properties is very broad. Nonetheless, viable techniques to estimate magnification exist. One of these techniques makes use of the Fundamental Plane (FP, Bertin and Lombardi, 2006) for elliptical galaxies. In the FP, the intrinsic effective radius of a galaxy $R_{\rm eff}^{\rm FP}$ is related to the galaxy's surface brightness and stellar velocity dispersion (neither of which are affected by lensing magnification). The measured effective radius $R_{\rm eff}$, on the other hand, is magnified. A comparison of $R_{\rm eff}$ and $R_{\rm eff}^{\rm FP}$ estimates the magnification of a galaxy $\mu \approx (R_{\rm eff}/R_{\rm eff}^{\rm FP})^2$. A modified version of the FP has been successfully applied by Huff and Graves (2011) to the photometric catalogue from the Sloan Digital Sky Survey (SDSS).

The intrinsic scatter in the FP is $\sim 20\%$ (Bernardi et al., 2003) which corresponds to a $\sigma_\mu \sim 40\%$ scatter in the magnification estimate. We repeat our analysis with a scaling factor as measured in Eq. (5.20) but now add Gaussian noise with mean zero and width σ_μ to the magnification from the ray-tracing catalogue. Fortunately, the introduced scatter does not invalidate the mitigation for lens redshifts $z_{\rm d} > 0.2$. The reason for this is the almost linear relation between magnification and shear, as seen in Fig. 5.5. For lenses with $z_{\rm d} \leq 0.2$ the noise in the magnification estimate enhances the variance in the SRT by a factor of ~ 20. This is again due to the very small foreground signal (compare Fig. 5.5). The difference in the foreground signal $\bar{\gamma}_{\rm LSS}$ between bin '1' and '2' is so small that the introduced scatter in magnification can bring the difference very close to zero and can thus lead to unreasonably high values in the scaling factor. The enhanced scatter in the SRT can be suppressed by enforcing the scaling factor to be smaller than 5, which is physically justified. With the added constraint on λ, the alternative mitigation strategy for the magnification bias performs nearly as well as shown in Fig. 5.4.

5.5.3. Impact of shape noise

In this work, we made use of ray-tracing simulation that have only minor contributions of noise. In observations, the largest source of uncertainty in weak lensing measurements is shape noise (e.g., Niemi et al., 2015). It arises because the measured galaxy ellipticities are dominated by the intrinsic galaxy shapes, with a much weaker contribution from lensing shear. The intrinsic ellipticity is expected to be randomly distributed and thus, the average over a sufficiently large number of galaxies vanishes. We employ a simplified model of additive, uncorrelated Gaussian noise to obtain an estimate of how shape noise affects our mitigation strategy. We add a Gaussian with zero mean and width $\sigma_\epsilon = 0.3$ to each Cartesian shear component from the ray-tracing catalogue before estimating the tangential shear around lens galaxies. Furthermore, we limit the density of source galaxies to $< 35\,{\rm gal/arcmin}^2$ which is certainly fulfilled with a magnitude cut in the r-band at 24 mag. The last adjustment we make for the limiting magnitude for the lens galaxies, it is reduced to 22 mag which leaves $\sim 1/3$ of the lens galaxies with a 24 mag-cut. Since we reduced

the density of background galaxies by a factor of ~ 2500, the SRT shows a way more noisy result. For some combinations of lens and source redshifts, the corresponding variance increased by a factor of 10 000. As one of the reasons, we identify the decreased statistics; however, we must also take into account how we constructed the null hypothesis of the SRT (5.12). It contains a ratio of shear values, nominator as well as denominator are noisy quantities and if the denominator is close to zero, we obtain very high values in the SRT. Hence, we propose to change the null hypothesis to

$$\langle \bar{\gamma}_\mathrm{g}(z_\mathrm{d}, z_j) \rangle \, \beta(z_\mathrm{d}, z_i) - \langle \bar{\gamma}_\mathrm{g}(z_\mathrm{d}, z_i) \rangle \, \beta(z_\mathrm{d}, z_j) = 0 \, . \tag{5.21}$$

With the new estimator the noise increased for all combinations of lens and source redshifts roughly by a factor 100 without extreme outliers as before.

To still verify that our mitigation strategy improves the SRT, we perform a reduced χ^2 test and the result is shown in table 5.1. As expected, the mitigation

Table 5.1. – Results of a reduced χ^2 test.

lens mag limit	source mag limit	mitigation	χ^2_red
none	none	no	80.90
none	none	yes	1.06
24mag	none	no	10.97
24mag	none	yes	1.02
22mag	24.5mag	no	2.28
22mag	24.5mag	yes	1.36

Notes. We perform the test for comparability for all cases with the modified null hypothesis (5.21). The first two rows correspond to a mass-limited sample, the next two to a magnitude limited sample and the last two also consider shape noise. For each, we show the χ^2_red result with and without mitigation. As can be seen by eye, the mitigation drastically improves the result in the first two cases. In the case that includes noise, the χ^2_red test still performs better when the mitigation is used.

improves the χ^2_red value significantly for the cases shown in Fig. 5.1. For the shape noise dominated case, an SRT with corrected magnification bias still performs better than without mitigation, although the difference is less pronounced. It is important to note that the level of shape noise reduces with an increased observed area in the sky. On the other hand, the magnification bias is fairly independent of the observed area. Thus, future surveys with observed areas $\gtrsim 1000\,\mathrm{deg}^2$ will have less impact on shape noise than we considered here (and vice versa).

5.5.4. Impact of intrinsic alignments

So far we ignored that the intrinsic ellipticities of galaxies may be correlated with the positions of other galaxies that are close in real space (e.g. Joachimi et al.,

2015). This flavor of intrinsic alignment (IA) may impact our estimate of the LSS contribution to the shear signal at the lens redshift. In particular, Eq. (5.15) is modified

$$\langle\bar{\gamma}_\mathrm{t}\rangle(z_\mathrm{d}, z_\mathrm{d}) = \langle\bar{\gamma}_\mathrm{LSS}\rangle(z_\mathrm{d}, z_\mathrm{d}) + \langle\bar{\gamma}_\mathrm{IA}\rangle(z_\mathrm{d}, z_\mathrm{d}). \tag{5.22}$$

The IA contribution is

$$\langle\bar{\gamma}_\mathrm{IA}\rangle(z_\mathrm{d}, z_\mathrm{d}) = \int_{\theta_\mathrm{in}}^{\theta_\mathrm{out}} 2\pi\,\theta\,\mathrm{d}\theta\,\langle\gamma_\mathrm{IA}\rangle(\theta)\,\frac{1}{2\pi\theta^2}. \tag{5.23}$$

The intrinsic tangential ellipticity $\langle\gamma_\mathrm{IA}\rangle(\theta)$ as a function of angular separation can be roughly estimated by:

$$\langle\gamma_\mathrm{IA}\rangle(\theta) \approx \frac{w_\mathrm{dI}(\theta\chi_\mathrm{d}, z_\mathrm{d})}{\Delta\chi(\Delta z) + w_\mathrm{dd}(\theta\chi_\mathrm{d}, z_\mathrm{d})}. \tag{5.24}$$

Here, $w_\mathrm{dI}(r, z)$ denotes the projected cross correlation at transverse comoving separation r and redshift z between the lens galaxy positions and the tangential components of the source galaxy intrinsic ellipticities, $w_\mathrm{dd}(r, z)$ denotes the projected cross correlation of the lens and source galaxy positions, and $\Delta\chi$ is the projection depth corresponding to the redshift interval Δz used to select source galaxies around the lens redshift.

We estimate the IA contribution using measurements of the projected IA correlations by Hilbert et al. (2017) in the Illustris simulation (Vogelsberger et al., 2014). The correlation w_dI, and thus $\langle\bar{\gamma}_\mathrm{IA}\rangle(z_\mathrm{d}, z_\mathrm{d})$, depends on the source and lens galaxy sample selection criteria, and also on how much weight the galaxy image shear estimator gives to the outskirts of the galaxy images.

The resulting $\langle\gamma_\mathrm{IA}\rangle(\theta)$ is shown in Fig. 5.6 for lenses at $z_\mathrm{d} = 0.83$ and sources between $z_\mathrm{s} = z_\mathrm{d} - \Delta z/2$ and $z_\mathrm{s} = z_\mathrm{d} + \Delta z/2$, where $\Delta z = 0.05$. This yields $\langle\bar{\gamma}_\mathrm{IA}\rangle(z_\mathrm{d}, z_\mathrm{d}) \approx -3 \times 10^{-4}$ for radially weighted moments (such as Kaiser et al., 1995, KSB), which is noticeable compared to the LSS contribution $\langle\bar{\gamma}_\mathrm{LSS}\rangle(z_\mathrm{d}, z_\mathrm{d})$ (see Fig. 5.5).

Equation (5.24) shows that the IA contribution $\langle\bar{\gamma}_\mathrm{IA}\rangle(z_\mathrm{d}, z_\mathrm{d})$ could be reduced by substantially increasing the source redshift window size Δz. Furthermore, IA correlations appear to be dominated by galaxies in the same halo and do not reach beyond a few tens of Mpc. Thus, the IA contribution could be avoided if one can reliably select source-lens pairs such that the sources are at least a few tens of Mpc in front of the lenses. Moreover, the IA contribution $\langle\bar{\gamma}_\mathrm{IA}\rangle(z_\mathrm{d}, z_\mathrm{d})$ can be substantially reduced by increasing the lower integration bound θ_in. For example, $\langle\bar{\gamma}_\mathrm{IA}\rangle$ practically vanishes for KSB-like estimators and $\theta_\mathrm{in} \geq 1\,\mathrm{arcmin}$.

5.6. Discussion & Conclusion

In this article, we take a closer look at the influence of magnification bias on the shear-ratio test as introduced in Jain and Taylor (2003) as well as a viable mitigation

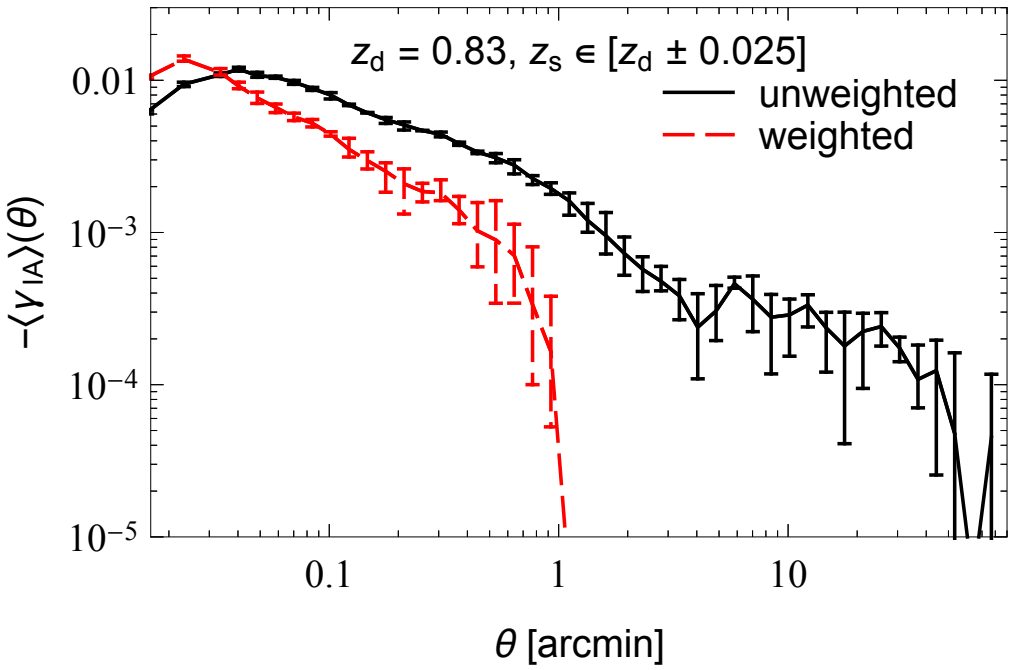


Figure 5.6. – The mean intrinsic tangential ellipticity $\langle \gamma_{\rm IA} \rangle(\theta)$ as a function of angular separation θ (estimated from the Illustris simulation, Vogelsberger et al., 2014; Hilbert et al., 2017) for lenses at redshift $z_{\rm d} = 0.83$ and sources at redshifts $z_{\rm s} \in [0.805, 0.855]$. The IA signal strongly depends on whether the shape estimator uses unweighted or radially weighted second moments of the galaxy image light distribution.

strategy. An advantage of the SRT is that it can be applied to the same data as obtained from cosmic shear surveys. Moreover, it is a purely geometrical method and does not rely on any assumptions of structure growth. As such, this null test has the potential to uncover remaining systematics in shear measurement and redshift estimation. Schneider (2016) even extended the SRT in such a way that it does not depend on the choice of cosmology anymore.

The SRT is based on a ratio of shears that are induced by matter correlated with the lens galaxies, and it does not take LSS effects into account. LSS mainly changes the observed shear signal due to magnification which alters the observed number density of lens galaxies on the sky. Hilbert et al. (2009) showed that, depending on the observed scale, the tangential shear can deviate by up to 20% from the GGL signal expected from shear that is induced by matter correlated with the lens galaxies, though this fact seems to have been largely ignored in subsequent observational studies. By using shear power spectra, Ziour and Hui (2008) derived relations that suggested that magnification bias influences the SRT quite heavily.

We made use of ray-tracing results through the Millennium Simulation (Hilbert et al., 2009) and galaxy catalogues from semi-analytic models (Henriques et al., 2015) to obtain accurate estimates for the tangential shear around galaxies at several redshifts. We used lenses in the redshift range $0.09 \leq z_\mathrm{d} \leq 0.83$ and sources in the range $z_\mathrm{d} < z_\mathrm{s} \leq 1.9$. With that, we were able to quantify the impact of magnification bias on the SRT as can be seen in Fig. 5.1 and 5.2. We find: (1) the higher the lens redshift, the larger is the deviation of the SRT from its expected value – for the lens redshifts considered the deviation from 0 increases from 10^{-3} by a factor of $\sim$100, (2) lenses and sources must be well separated along the line-of-sight – the relative impact of the magnification bias on the SRT is largest when the source and lens galaxies are close, and (3) magnification bias depends on the range over which shear is estimated.

For our mitigation strategy, we assume that the shear signal is a superposition of the shear induced by matter correlated with the lens galaxies at redshift z_d and shear due to matter between us and the source galaxies, where the LSS in the redshift range $z_\mathrm{d} < z < z_\mathrm{s}$ is irrelevant for the magnification bias on the shear signal. To extract the LSS-induced shear signal from the data, we measure the tangential shear around lens galaxies at the lens redshift and use a scaling factor to approximate its value at the source redshift. The scaling factor can be calculated as a correlation between these two shear components (Eq. 5.17). Subtracting the scaled LSS-induced shear from the measured shear signal will yield the shear that is induced by matter correlated with the lens galaxies. The latter is what is needed for the SRT, and the good performance of this mitigation approach can be seen in Fig. 5.4. We further introduced an alternative way of obtaining the scaling factor that relies on dividing the lens population into sub-samples with different magnification. Estimating the magnification with the Fundamental Plane for early-type galaxies leads to results that perform well for all redshifts where magnification bias is important.

Furthermore, we estimated the impact of shape noise on our mitigation strategy. All redshifts show an increased scatter by a factor of ~ 100. A $\chi^2_{\rm red}$ analysis showed that applying the mitigation still improves the SRT. Also, shape noise can be reduced by observing a larger area on the sky in contrast to the magnification bias. We used roughly $1000\,\mathrm{deg}^2$ for our analysis, future experiments like *Euclid* will surpass this by a factor of ≥ 10. Finally, we discussed the possible impact of IA on our mitigation strategy. Since the IA contribution to the shear signal at the lens redshift might be substantial compared to the LSS contribution, modifications to our mitigation strategy that reduce the impact of IA (e.g. by estimating the LSS contribution using sources slightly in front of the lenses) should be explored in more detail in future work.

Magnification bias is present on all relevant scales and needs to be corrected for. It affects not only the performance of the SRT but must be considered in all applications of GGL and its generalization to groups and clusters. A viable mitigation strategy is, therefore, crucial for ongoing and future experiments.

CHAPTER 6

MAGNIFICATION EFFECTS IN GALAXY-GALAXY LENSING

Building on the results of the previous chapter, the influence from magnification effects on shear estimates is generalised. This chapter was accepted by the journal Astronomy & Astrophysics, reproduced with permission ©ESO.

Abstract: Magnification changes the observed local number density of galaxies on the sky. This biases the observed tangential shear profiles around galaxies: the so-called galaxy-galaxy lensing (GGL) signal. Inference of physical quantities, such as the mean mass profile of halos around galaxies, are correspondingly affected by magnification effects. We used simulated shear and galaxy data from the Millennium Simulation to quantify the effect on shear and mass estimates from the magnified lens and source number counts. The former is due to the large-scale matter distribution in the foreground of the lenses; the latter is caused by magnification of the source population by the matter associated with the lenses. The GGL signal is calculated from the simulations by an efficient fast Fourier Transform, which can also be applied to real data. The numerical treatment is complemented by a leading-order analytical description of the magnification effects, which is shown to fit the numerical shear data well. We find the magnification effect is strongest for steep galaxy luminosity functions and high redshifts. For a KiDS+VIKING+GAMA-like survey with lens galaxies at redshift $z_\mathrm{d} = 0.36$ and source galaxies in the last three redshift bins with a mean redshift of $\bar{z}_\mathrm{s} = 0.79$, the magnification correction changes the shear profile up to 2%, and the mass is biased by up to 8%. We further considered an even higher redshift fiducial lens sample at $z_\mathrm{d} = 0.83$, with a limited magnitude of 22 mag in the r-band and a source redshift of $z_\mathrm{s} = 0.99$. Through this, we find that a magnification correction changes the shear profile up to 45% and that the mass is

biased by up to 55%. As expected, the sign of the bias depends on the local slope of the lens luminosity function $\alpha_{\rm d}$, where the mass is biased low for $\alpha_{\rm d} < 1$ and biased high for $\alpha_{\rm d} > 1$. While the magnification effect of sources is rarely more than 1% of the measured GGL signal, the statistical power of future weak lensing surveys warrants correction for this effect.

6.1. Introduction

Gravitational lensing is a powerful tool in unveiling the true distribution of matter in the Universe and probing cosmological parameters (see, e.g. Kilbinger, 2015, for a recent review). The lensing signal is sensitive to all matter, regardless of its nature, and is observed as the distortion of light bundles travelling through the Universe. In the weak lensing regime, this distortion is small and must be studied with large statistical samples. Therefore, large and deep surveys are required, for example, the Kilo Degree Survey[1](KiDS), the Dark Energy Survey[2] (DES), the Hyper Surprime-Cam Subaru Strategic Program[3] (HSC SSP), or the near future surveys with *Euclid*[4], and the Large Synoptic Survey Telescope[5] (LSST). To maximise the scientific output from these surveys, the scientific community is currently putting great efforts into understanding nuances in the theoretical framework.

Galaxy-galaxy lensing (GGL) correlates the position of foreground galaxies to the distortion of background galaxies (see, e.g. Hoekstra, 2013). The distortion is typically measured in terms of mean tangential shear with respect to the lens position. This shear signal, as a function of separation from the lens centre, can be related to the underlying mass properties of the parent halo. A major challenge is obtaining an unbiased mass estimate. Biases arise if the underlying model does not describe all contributions to the matter-shear correlation function sufficiently. Also, the galaxy bias that connects the position of galaxies to its surrounding matter distribution must be carefully taken into account. For current and future surveys, we must further consider second-order effects to the galaxy-matter correlation, such as, for example, magnification effects (Ziour and Hui, 2008; Hilbert et al., 2009) and intrinsic alignment of galaxies (Troxel and Ishak, 2015). In this work, we focus on the former effect.

Magnification is the change of the observed solid angle of an image compared to the intrinsic solid angle, or, since the surface brightness remains constant, the ratio of observed flux to the intrinsic one. Like the shear, it is a local quantity, a direct prediction of the lensing formalism, and is caused by all matter between the observed galaxy population and us. However, direct measurements of magnification

[1] https://www.kids.strw.leidenuniv.nl
[2] https://www.darkenergysurvey.org
[3] https://hsc.mtk.nao.ac.jp/ssp/
[4] https://www.euclid-ec.org
[5] https://www.lsst.org

are challenging because the intrinsic flux is typically unknown. Yet, the change in size and magnitude results in a changed spatial distribution of the galaxy population. This so-called number count magnification has been measured (e.g., Chiu et al., 2016; Garcia-Fernandez et al., 2018). Consequently, magnification by the large-scale structure (LSS) also changes the GGL signal compared to a signal that is just given by matter correlated with the lens galaxies. We stress that the magnification changes the number counts of the source as well as lens galaxies on the sky. The impact of magnification of the lens galaxies on the GGL signal for surveys like Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) is $\sim 5\%$ (Simon and Hilbert, 2018), but can be as large as 20% for other lens samples (Hilbert et al., 2009). Although these results suggest a fairly large impact of magnification on GGL lensing estimates, quantitative analyses widely neglect the influence of magnification. Unruh et al. (2019) studied the impact of the number count magnification of lens galaxies on shear-ratios. They found that the shear-ratio test (Jain and Taylor, 2003) is affected by lens magnification and that its effect must be mitigated, especially for high lens redshifts.

In this paper, we quantify the impact of magnification on observed tangential shear profiles and halo-mass estimates from GGL. We consider both the effect of magnification of the sources by the lenses, as well as the effect of magnification of the lenses by the LSS. For this, we compared the GGL signal with and without magnification using simulated data. We then derived mean halo masses in both cases employing a halo model to quantify the expected mass bias. We complemented the numerical results with analytic estimates of the effects.

This article is organised as follows. The theoretical framework is briefly described in Sect. 6.2. Section 6.3 features an analytical description of how magnification affects shear estimates, as well as a brief discussion of how numerical results were obtained in this study. The numerical procedure is then more thoroughly explained in Appendix 6.8. In Sect. 6.4, we discuss the impact of magnification of source galaxies by the lenses, and in Sect. 6.5 we discuss magnification of the lens galaxies by the LSS. We study the magnification bias on mass estimates in Sect. 6.6 and conclude in Sect. 6.7.

6.2. Theory

In the following, we introduce the theoretical concepts of gravitational lensing for this work. For a more general and extensive overview, the reader is kindly referred to Bartelmann and Schneider (2001).

6.2.1. Cosmological distances

For a flat universe, the Hubble parameter $H(z)$ can be written as

$$\left(\frac{H(z)}{H_0}\right)^2 = \Omega_\mathrm{m}(1+z)^3 + 1 - \Omega_\mathrm{m} \,, \tag{6.1}$$

where z is the redshift, H_0 denotes the Hubble constant, and Ω_m is the matter density in units of today's critical density $\rho_\mathrm{crit} = 3H_0^2/(8\pi G)$; with the vacuum speed of light c and the gravitational constant G. The comoving distance travelled by a photon between redshift z_1 and z_2 reads

$$\chi(z_1, z_2) = \int_{z_1}^{z_2} \frac{c\,\mathrm{d}z'}{H(z')} \,, \tag{6.2}$$

and the angular-diameter distance is

$$D(z_1, z_2) = \frac{\chi(z_1, z_2)}{1+z_2} \,. \tag{6.3}$$

For a redshift $z_1 = 0$, which is the observer's position, we write $D(0, z) =: D(z)$. In addition, the dimensionless Hubble parameter h is used to parametrise our ignorance about the true value of today's Hubble parameter, defined as $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. In the following, all distances are angular-diameter distances.

6.2.2. Gravitational lensing distortions and magnification

Gravitational lensing distorts the appearance of galaxy images. In the weak lensing regime, this distortion can locally be described as a linear mapping from the background (source) plane to the foreground (lens) plane. The Jacobian $\mathcal{A}$ of the local mapping can be written as

$$\mathcal{A} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} = (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \,, \tag{6.4}$$

where κ is the convergence, $\gamma_{1,2}$ are the two Cartesian shear components, and $g_{1,2} = \gamma_{1,2}/(1-\kappa)$ are the two Cartesian reduced shear components, which all depend on the position in the lens plane. The convergence causes an isotropic scaling of the galaxy image, while the shear leads to an anisotropic stretching, and thus causes an initially circular object to appear elliptical.

This scaling of the galaxy image changes the apparent solid angle ω of the image, compared to one in the absence of lensing, which we denote by ω_0. Likewise, the flux is affected by gravitational lensing, the unlensed flux s_0 is enhanced or reduced

to the observed flux s. The ratio of these quantities defines the magnification μ and can also be calculated from the Jacobian by

$$\mu = \frac{\omega}{\omega_0} = \frac{s}{s_0} = \frac{1}{\det \mathcal{A}} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} \ . \tag{6.5}$$

Magnification changes the observed local number density of galaxies on the sky. The cumulative observed number density of galaxies on the sky $n(>s)$, brighter than flux s, is locally

$$n(>s) = \frac{1}{\mu}\, n_0 \left(> \frac{s}{\mu} \right) \ , \tag{6.6}$$

where n_0 denotes the cumulative number density in absence of lensing. The prefactor $1/\mu$ is due to the scaling of the solid angle. The flux in the argument of n_0 must also be scaled by $1/\mu$ to account for the flux enhancement or reduction.

Magnification effects in the weak lensing limit are small, specifically $|\mu - 1| \ll 1$, and we Taylor expand Eq. (6.6) in $(\mu - 1)$ to obtain to first order

$$\frac{n(>s)}{n_0(>s)} = \mu^{\alpha - 1.} \tag{6.7}$$

Where the exponent α is the local slope at the flux limit $s_{\rm lim}$, it is defined as

$$\alpha = -\frac{\mathrm{d}\log_{10} n_0(>s)}{\mathrm{d}\log_{10} s}\bigg|_{s_{\rm lim}} \ . \tag{6.8}$$

For $\alpha > 1$, the galaxy counts are enhanced, and for $\alpha < 1$ they are depleted. In the case of $\alpha = 1$, no magnification bias is present.

6.2.3. Galaxy-galaxy lensing

In GGL, the positions of foreground galaxies (lenses) are correlated with the shear of background galaxies (sources). For a position $\boldsymbol{\theta}$, the complex shear is written as $\gamma(\boldsymbol{\theta}) = \gamma_1(\boldsymbol{\theta}) + \mathrm{i}\gamma_2(\boldsymbol{\theta})$. The tangential shear $\gamma_{\rm t}$ and the cross shear $\gamma_{\rm x}$ at source position $\boldsymbol{\theta}_{\rm s}$ for a given lens at position $\boldsymbol{\theta}_{\rm d}$ are

$$\gamma_{\rm t}(\boldsymbol{\theta}_{\rm s}; \boldsymbol{\theta}_{\rm s} - \boldsymbol{\theta}_{\rm d}) \ + \ \mathrm{i}\gamma_{\rm x}(\boldsymbol{\theta}_{\rm s}; \boldsymbol{\theta}_{\rm s} - \boldsymbol{\theta}_{\rm d}) = -\gamma(\boldsymbol{\theta}_{\rm s})\, \frac{(\boldsymbol{\theta}_{\rm s} - \boldsymbol{\theta}_{\rm d})^*}{\boldsymbol{\theta}_{\rm s} - \boldsymbol{\theta}_{\rm d}} \ , \tag{6.9}$$

where an asterisk denotes complex conjugation. The GGL signal $\langle \gamma_{\rm t} \rangle(\theta)$ is defined as the correlator between the positions of foreground galaxies and the tangential shear,

$$\langle \gamma_{\rm t} \rangle(\theta) = \langle \kappa_{\rm g}(\boldsymbol{\theta}')\ \gamma_{\rm t}(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta}) \rangle \ , \tag{6.10}$$

where $\kappa_\mathrm{g}(\boldsymbol{\theta})$ is the fractional number-density contrast of foreground lens galaxies on the sky. The corresponding correlator for the cross-component of the shear is expected to vanish, due to parity invariance.

A practical estimator for the GGL signal averages the tangential and cross shear over many lens-source pairs in bins of separation θ:

$$\hat{\gamma}_{\mathrm{t}/\times}(\theta) = \frac{\sum_{ij} \Delta\left(\theta, |\boldsymbol{\theta}_\mathrm{s}^{(i)} - \boldsymbol{\theta}_\mathrm{d}^{(j)}|\right) \gamma_{\mathrm{t}/\times}\left(\boldsymbol{\theta}_\mathrm{s}^{(i)}; \boldsymbol{\theta}_\mathrm{s}^{(i)} - \boldsymbol{\theta}_\mathrm{d}^{(j)}\right)}{\sum_{ij} \Delta\left(\theta, |\boldsymbol{\theta}_\mathrm{s}^{(i)} - \boldsymbol{\theta}_\mathrm{d}^{(j)}|\right)} \,. \tag{6.11}$$

Here, $\boldsymbol{\theta}_\mathrm{s}^{(i)}$ denotes the position of the i-th source, $\boldsymbol{\theta}_\mathrm{d}^{(j)}$ denotes the position of the j-th lens, and the binning function $\Delta(\theta, \theta')$ is unity if θ' falls into the corresponding θ bin, and zero if it does not.

6.3. Magnification effects in GGL

In this section, we consider the effect of magnification on the GGL signal. As we show, magnification of sources and lenses leads to a bias of the estimator (6.11), which is a function of limiting magnitudes for the lens and source population, as well as their redshifts, since these determine the local slope (6.8) at the limiting magnitude. Magnification is typically assumed to be a minor effect in GGL measurements, and most theoretical predictions do not account for it. While the impact of the magnification of lenses has already received some attention (e.g., Ziour and Hui, 2008; Hartlap, 2009), the source magnification is less well known.

6.3.1. Magnification of lenses by large-scale structure

Magnification, caused by the LSS between us and the lenses, changes the number density of the lens galaxy sample, while simultaneously inducing a shear on background galaxies. Thus, the observed shear signal differs from what is typically considered as the GGL signal, which is a correlation of lens galaxy positions and the shear on background galaxies (Eq. 6.11). In this correlation, the lens galaxies are connected by the galaxy bias to their surrounding matter, which induces a shear in the background galaxies. Magnification by intervening matter structures alters this rather simple picture. Since for larger lens redshifts more intervening matter is present, the impact of magnification effects grows with increasing redshift. On the other hand, the impact is reduced with increasing line-of-sight separations of lenses and sources. In the following, we consider a lowest-order correction for the magnification of the lenses by the LSS for the GGL signal of a flux- or volume-limited lens sample (see, e.g., Ziour and Hui, 2008; Hartlap, 2009; Thiele et al., 2020). We stress that this correction ignores the magnification of sources, which is treated in the next sub-section.

In the weak lensing regime, we can approximate the magnification by $\mu \approx 1 + 2\kappa$, valid if $\kappa \ll 1$, $|\gamma| \ll 1$. Then, the number count magnification (6.7) of the observed number density of lenses $n_\mathrm{d}(\boldsymbol{\theta})$ at redshift z_d on the sky is, for a flux-limited sample,

$$n_\mathrm{d}(\boldsymbol{\theta}, z_\mathrm{d}) = n_{\mathrm{d},0}(\boldsymbol{\theta}, z_\mathrm{d}) + 2\,[\alpha_\mathrm{d}(z_\mathrm{d}) - 1]\;\kappa^\mathrm{LSS}(\boldsymbol{\theta}, z_\mathrm{d})\,\bar{n}_\mathrm{d}(z_\mathrm{d})\;, \tag{6.12}$$

where $n_{\mathrm{d},0}$ denotes the lens number density without magnification, $\bar{n}_\mathrm{d}$ denotes the mean lens number density, α_d denotes the local slope of the lenses at the limiting magnitude, and κ^LSS denotes the convergence due to matter structures between us and the lenses. Thus, in the presence of magnification, the expected signal is modified to

$$\gamma_\mathrm{t}(\theta|z_\mathrm{d}, z_\mathrm{s}) = \gamma_\mathrm{t}^\mathrm{nomagn}(\theta|z_\mathrm{d}, z_\mathrm{s}) + 2\,[\alpha_\mathrm{d}(z_\mathrm{d}) - 1]\;\gamma_\mathrm{t}^\mathrm{LSS}(\theta|z_\mathrm{d}, z_\mathrm{s})\;, \tag{6.13}$$

where $\gamma_\mathrm{t}^\mathrm{nomagn}$ denotes the tangential shear signal without magnification, and the LSS shear signal is

$$\gamma_\mathrm{t}^\mathrm{LSS}(\theta|z_\mathrm{d}, z_\mathrm{s}) = \frac{9H_0^3\Omega_\mathrm{m}^2}{4c^3}\int_0^\infty \mathrm{d}\ell\;\ell\,\mathrm{J}_2(\ell\theta)\int_0^{z_\mathrm{d}}\mathrm{d}z\;\frac{H_0}{H(z)} \times\;\frac{D(z, z_\mathrm{d})\,D(z, z_\mathrm{s})}{D_\mathrm{d}\,D_\mathrm{s}}\,P_\mathrm{m}\left(\frac{\ell + 1/2}{(1+z)\,D(z)}; z\right)\;, \tag{6.14}$$

which is shown explicitly in Hartlap (2009) and Simon and Hilbert (2018). We set $D(z_\mathrm{s}) = D_\mathrm{s}$ and $D(z_\mathrm{d}) = D_\mathrm{d}$ and by $\mathrm{J}_n(x)$ we denote the nth-order Bessel function of the first kind. In this work, we use the revised `Halofit` model (Takahashi et al., 2012) for the spatial matter power spectrum $P_\mathrm{m}(k; z)$ at wavenumber k and redshift z. The argument in the matter power spectrum arises through the application of the wide-angle corrected Limber projection, which was recently put forward by Kilbinger et al. (2017); the denominator is the comoving angular-diameter distance $f_k(z) = (1+z)D(z)$ at redshift z. The corresponding expression of Eq. (6.13) in the absence of a flux limit, meaning for a volume-limited lens sample, can be obtained by setting $\alpha_\mathrm{d} = 0$.

For $\alpha_\mathrm{d} < 1$, the magnification by the LSS suppresses the GGL signal. For $\alpha_\mathrm{d} = 1$, the magnification effect vanishes. For $\alpha_\mathrm{d} > 1$, the LSS contribution enhances the GGL signal.

6.3.2. Magnification of sources by lenses

Galaxies are correlated with the mass distribution, and thus the location of the galaxies correlates with the magnification induced on the background sources. This implies that the number density of sources is correlated with the positions of the lens galaxies. Assuming for a moment that the number-count slope of sources α is larger than unity, one expects that the number density of sources is more enhanced

close to lens galaxies living in a dense environment. The estimator (6.11) therefore contains a disproportionately high number of lens-source pairs for those lenses living in a dense environment compared to those located in less dense regions.

The expected number density of sources is

$$\begin{aligned} n_{\rm s}(\boldsymbol{\theta}, > s) &= \frac{1}{\mu(\boldsymbol{\theta})}\, n_{\rm s0}\left(> \frac{s}{\mu(\boldsymbol{\theta})}\right) \approx n_{\rm s0}(> s)\, \mu^{\alpha_{\rm s}-1}(\boldsymbol{\theta}) \\ &\approx n_{\rm s0}(> s)\, [1 + 2(\alpha_{\rm s} - 1)\kappa(\boldsymbol{\theta})] \ , \end{aligned} \tag{6.15}$$

where in the second step we used the first-order Taylor expansion leading to Eq. (6.7), and in the last step we again made the weak lensing approximation $\mu \approx 1 + 2\kappa$.

The expectation value of the estimator (6.11) of the GGL signal is therefore affected by the local change of the source number density and becomes

$$\begin{aligned} \langle\hat{\gamma}_{\rm t}\rangle(\theta) &= \left\langle \kappa_{\rm g}(\boldsymbol{\theta}')\, \gamma_{\rm t}(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta})\, \frac{1}{\mu(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta})}\, \frac{n_{\rm s0}[> s/\mu(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta})]}{n_{\rm s0}(> s)} \right\rangle \\ &\approx \langle \kappa_{\rm g}(\boldsymbol{\theta}')\, \gamma_{\rm t}(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta})\, \mu^{\alpha_{\rm s}-1}(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta}) \rangle \\ &\approx \langle\gamma_{\rm t}\rangle(\theta) + 2(\alpha_{\rm s} - 1) \langle \kappa_{\rm g}(\boldsymbol{\theta}')\, \gamma_{\rm t}(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta})\, \kappa(\boldsymbol{\theta}' + \boldsymbol{\theta}; \boldsymbol{\theta}) \rangle \ . \end{aligned} \tag{6.16}$$

Thus, in the case of small magnifications, the bias is given by a third-order cross-correlation between the number density of foreground (lens) galaxies and the shear and convergence experienced by the background galaxies. This correlation is caused by the lensing effect of matter associated with the lens galaxies; hence, the bias is caused by magnification of sources by the matter at $z_{\rm d}$. Equation (6.16) ignores the effect of intervening matter since this is sub-dominant for the source galaxy sample. Given that the bias term differs from the GGL signal by one order in the convergence, and that the characteristic convergence dispersion is of order 10^{-2}, we expect that magnification of sources biases the GGL signal at the level of $\sim 1\%$.

Interestingly, the third-order correlator in the final expression of Eq. (6.16) is related to the galaxy-shear-shear correlator that was introduced by Schneider and Watts (2005) as one of the $G_{\pm}$ quantities of galaxy-galaxy-galaxy lensing, since κ and γ are linearly related. Thus, from measurements of the galaxy-shear-shear correlations in a survey, this bias term can be directly estimated. We note that such measurements have already been successfully conducted (e.g. Simon et al., 2008, 2013). A more quantitative description of this correction, which will be relevant for precision GGL studies in forthcoming surveys like *Euclid* and LSST, is beyond the scope of this paper and will be done at a later stage.

An approximate, more intuitive way of describing the magnification of sources by lenses is provided by assuming that each lens is located at the centre of a halo of mass m. In the case of no magnification, the expected tangential shear signal

$\gamma_\mathrm{t}(\theta) = \langle\hat{\gamma}_\mathrm{t}(\theta)\rangle$ can be expressed as[6]

$$\gamma_\mathrm{t}(\theta) = \int \mathrm{d}z_\mathrm{s}\, p_{z_\mathrm{s}}(z_\mathrm{s}) \int \mathrm{d}z_\mathrm{d}\, p_{z_\mathrm{d}}(z_\mathrm{d}) \int \mathrm{d}m\; p_{m|z_\mathrm{d}}(m|z_\mathrm{d}) \gamma_\mathrm{t}(\theta|m, z_\mathrm{d}, z_\mathrm{s}) \;, \tag{6.17}$$

for a population of sources with redshift distribution $p_{z_\mathrm{s}}(z_\mathrm{s})$, a population of lenses with redshift distribution $p_{z_\mathrm{d}}(z_\mathrm{d})$, and a conditional distribution $p_{m|z_\mathrm{d}}(m|z_\mathrm{d})$ of the masses m of the halos in which the lens galaxies reside. The mean tangential shear profile $\gamma_\mathrm{t}(\theta|m, z_\mathrm{d}, z_\mathrm{s})$ for lenses with halo mass m at redshift z_d and sources at redshift z_s can be factorised,

$$\gamma_\mathrm{t}(\theta|m, z_\mathrm{d}, z_\mathrm{s}) = \gamma_\infty(\theta|m, z_\mathrm{d}) \frac{D_\mathrm{ds}}{D_\mathrm{s}} \;, \tag{6.18}$$

where $D_\mathrm{ds} = D(z_\mathrm{d}, z_\mathrm{s})$, and γ_∞ is the mean shear profile for (hypothetical) sources at infinite distance.

Equation (6.17) assumes that the source number density is statistically independent of the lens positions. However, the observed number density of sources may change behind lenses due to magnification by the lenses. The expected magnification is a function of angular separation, the source and lens redshift, and the lens halo mass. For a flux-limited sample, the expected shear signal (6.17) then changes to:

$$\begin{aligned}
\gamma_\mathrm{t}(\theta) = &\left[\int \mathrm{d}z_\mathrm{s}\, p_{z_\mathrm{s}}(z_\mathrm{s}) \int \mathrm{d}z_\mathrm{d}\, p_{z_\mathrm{d}}(z_\mathrm{d}) \int \mathrm{d}m\; p_{m|z_\mathrm{d}}(m|z_\mathrm{d}) \right. \\
&\left. \times\; \mu(\theta|m, z_\mathrm{d}, z_\mathrm{s})^{\alpha_\mathrm{s}(z_\mathrm{s})-1}\right]^{-1} \\
&\times \int \mathrm{d}z_\mathrm{s}\, p_{z_\mathrm{s}}(z_\mathrm{s}) \int \mathrm{d}z_\mathrm{d}\, p_{z_\mathrm{d}}(z_\mathrm{d}) \int \mathrm{d}m\; p_{m|z_\mathrm{d}}(m|z_\mathrm{d}) \\
&\times\; \mu(\theta|m, z_\mathrm{d}, z_\mathrm{s})^{\alpha_\mathrm{s}(z_\mathrm{s})-1}\, \gamma_\mathrm{t}(\theta|m, z_\mathrm{d}, z_\mathrm{s}) \;,
\end{aligned} \tag{6.19}$$

where $\mu(\theta|m, z_\mathrm{d}, z_\mathrm{s})$ denotes the mean magnification of sources at redshift z_s and separation θ by lenses at redshift z_d with halo mass m, and $\alpha_\mathrm{s}(z_\mathrm{s})$ denotes the slope of the source counts at redshift z_s at the source flux limit as in Eq. (6.8). We obtained the corresponding expression for a volume-limited source sample by replacing α_s by zero in Eq. (6.19).

We may also assume that $\gamma_\mathrm{t}(\theta|m, z_\mathrm{d}, z_\mathrm{s})$ and $\mu(\theta|m, z_\mathrm{d}, z_\mathrm{s})$ are larger when the halo mass m is larger (in the weak-lensing regime). Then, lens galaxies in more massive halos appear under represented in the estimator (6.11), and the expected shear signal (6.19) is lower than the prediction (6.17) ignoring magnification when $\alpha_\mathrm{s} < 1$. For $\alpha_\mathrm{s} = 1$, the effect vanishes. For $\alpha_\mathrm{s} > 1$, the number of source-lens pairs with more massive lenses in the GGL estimator (6.11) is enhanced more by magnification,

[6] A more comprehensive halo model is discussed in Sect. 6.6.

and thus one expects a shear signal (6.19) that is larger than the prediction (6.17). When neglecting magnification, this may cause biases in the estimation of the mean halo mass.

As Eq. (6.19) indicates, magnification may also affect the observed redshift distributions of the lenses and sources. Furthermore, the magnification profile $\mu(\theta|m, z_\mathrm{d}, z_\mathrm{s})$ usually has a strong radial dependence, being large for radii close to the Einstein radius, but rapidly dropping to values close to unity for larger radii. Thus, the magnification effects on the GGL signal are stronger for smaller radii. When neglecting magnification, this may cause additional biases when estimating parameters such as the halo concentration, and also when estimating the width of the halo-mass distribution.

6.3.3. Mock data production

We made use of ray-tracing results through the Millennium Simulation (Springel et al., 2005; Hilbert et al., 2009), from which we obtained tangential shear profiles using a fast Fourier Transform (FFT) method. The method follows and improves the one presented in Unruh et al. (2019). Specific details are given in Appendix 6.8.

For the radial binning of the tangential shear profile, we chose $N_\mathrm{bin} = 16$ logarithmically spaced bins between $\theta_\mathrm{in} = 0.6'$ and $\theta_\mathrm{out} = 17.5'$. We estimated the error on the shear signal with a jackknife method that measures the field-to-field variance of the 64 fields. Then, we repeated the whole process while replacing the lens galaxies with random positions to obtain the shear estimate that is caused by the long modes in the matter density field, as well as boundary effects as recommended by Singh et al. (2017). The random signal $\hat{\gamma}_\mathrm{rand}$ was subtracted from original signal $\hat{\gamma}_\mathrm{t} \to \hat{\gamma}_\mathrm{t} - \hat{\gamma}_\mathrm{rand}$. For convenience, we dropped the hat to distinguish the estimator $\hat{\gamma}_\mathrm{t}$ from theoretical expectations $\gamma_\mathrm{t} = \langle\hat{\gamma}_\mathrm{t}\rangle$ in the following.

For our analyses, we wanted to obtain samples of galaxies with different local slopes α of the source counts, which depend on redshift and limiting magnitude. For a given lens galaxy sample with a magnitude cut corresponding to a flux limit s_lim, we estimated the local slope (6.8) by finite differencing around s_lim. Figure 6.1 illustrates the cumulative number of galaxies at redshift $z_\mathrm{d} = 0.41$ for the whole simulated field of $64 \times 16\,\mathrm{deg}^2$. Several magnitude cuts are indicated, as well as one example of a tangential curve at $s_\mathrm{lim} = 21\,\mathrm{mag}$ with slope $\alpha = 1.06$. For all GGL measurements in this work, we fixed the source redshift to $z_\mathrm{s} = 0.99$. To find the local slopes α_s of the source galaxy counts, we varied the limiting magnitude in the r-band filter (see Table 6.1a). For the local slope α_d of the lens galaxy counts, we further varied the lens redshift z_d, as well as the limiting magnitude, as shown in Table 6.1b and 6.1c.

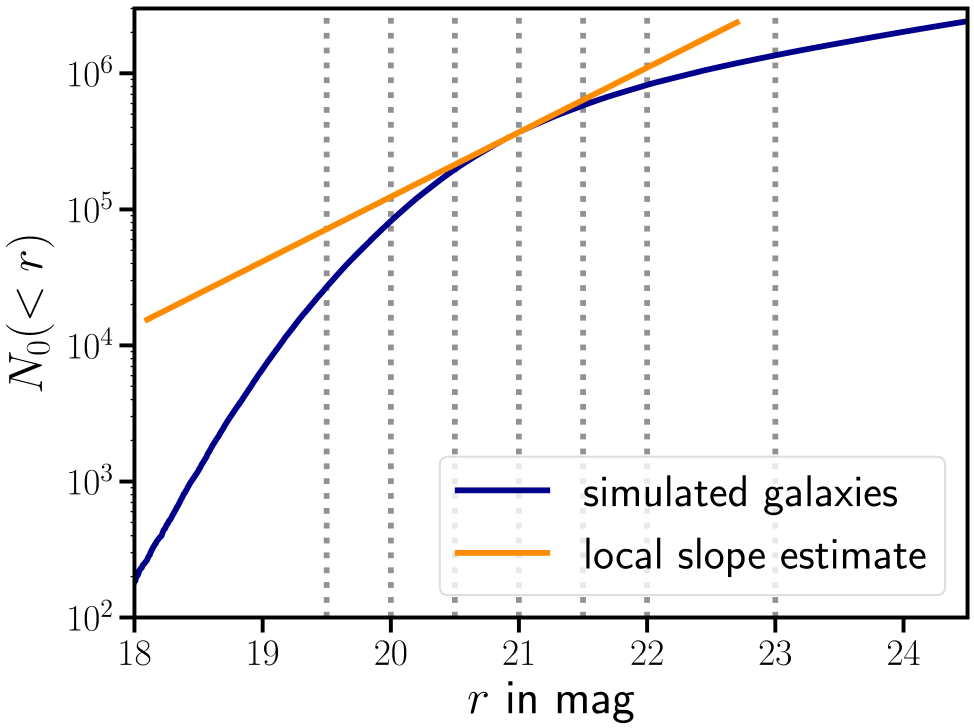


Figure 6.1. – Cumulative number of galaxies N_0 is shown as a function of r-band magnitude for a field of $64 \times 16\,\mathrm{deg}^2$ at redshift $z = 0.41$. The dotted vertical lines indicate the magnitude cuts listed in Table 6.1b. Finally, the tangential curve shows the local slope α at $r = 21$ mag.

Table 6.1. – Local slopes for the source galaxy counts α_s and the lens galaxy counts α_d given as a function of redshift and as a function of the limiting flux.

(a) fix $z_\mathrm{s} = 0.99$		(b) fix $z_\mathrm{d} = 0.41$		(c) fix $s_\mathrm{lim,d} = 22$	
$s_\mathrm{lim,s}$	α_s	$s_\mathrm{lim,d}$	α_d	z_d	α_d
22.0	2.89	19.5	2.71	0.24	0.46
22.5	2.33	20.0	2.03	0.41	0.66
23.0	1.75	20.5	1.56	0.51	0.98
23.5	1.29	21.0	1.06	0.62	1.58
24.0	0.93	21.5	0.85	0.83	2.49
24.5	0.68	22.0	0.66		
25.0	0.58	23.0	0.44		
26.0	0.45				

Notes. Either redshift or limiting flux is always kept fixed, while the other quantity is varied, as shown below. The limiting flux is given in terms of r-band magnitude.

6.4. Magnification effects on background galaxies

Lens and source galaxies are affected by magnification. To understand this impact in more detail, we first discuss how the shear profile changes when only source galaxies are magnified. In the following, we describe how we generated an arbitrary number of mock source catalogues, with and without magnification bias included. Results from the Millennium Simulation are then presented, and the magnification-induced bias in the GGL signal is compared to the prediction of the analytical model presented in Sect. 6.3.2.

6.4.1. Magnification switched off

To switch magnification off, we simply chose random source positions. We set the number of galaxies to $N_{\mathrm{s},0} = 10^7$ per $16\,\mathrm{deg}^2$ field to keep the impact of noise low.

6.4.2. Magnification switched on

Magnification changes the number counts of observed galaxies on the sky. Using the ray-tracing data, we obtained the cumulative number counts of the galaxies as a function of magnification-corrected flux, $n_{\mathrm{s},0}(> s_0)$. We obtained the local expected number counts of galaxies by adjusting the flux limit to $s_{\mathrm{lim,s}}/\mu(\boldsymbol{\theta})$ at each position $\boldsymbol{\theta}$ and using the first equality in Eq. (6.15). We further scaled the number counts so that for $\mu = 1$ the expected number of source galaxies is $N_{\mathrm{s}} = 10^7$ per field of solid angle $A = 4^\circ \times 4^\circ$. The threshold of finding a source at a grid position $\boldsymbol{\theta}$ is then $T(\boldsymbol{\theta}) = n_{\mathrm{s}}(> s_{\mathrm{lim,s}}; \boldsymbol{\theta})\, A/N_{\mathrm{pix}}$, where we restricted $T(\boldsymbol{\theta})$ to be smaller than unity. Finally, we drew a uniform random number $P(\boldsymbol{\theta})$ between zero and one for each position. A source galaxy was placed at a position $\boldsymbol{\theta}$ if $T(\boldsymbol{\theta}) > P(\boldsymbol{\theta})$. The 'magnification off' method can be recovered if we insert $\mu = 1$ for all $\boldsymbol{\theta}$ in Eq. (6.15).

6.4.3. Results

The relative impact of magnification of sources on a tangential shear profile is shown by the orange 'upward' triangles in Fig. 6.2. As expected, the net effect depends on the local slope α_{s}; the effect is typically of the order of 1 to 2% per bin. In the two panels on the left-hand side in Fig. 6.2, the local slope α_{s} is larger than unity, and the shear signal is enhanced; while in the two panels on the right, $\alpha_{\mathrm{s}} < 1$, which reverses the effect. Also, the magnification effect is stronger for smaller separations θ. This is seen more clearly in Fig. 6.3, where the absolute difference of source-magnified to magnification-corrected shear profiles is compared. The shear profiles vary with α_{s} for constant redshifts $z_{\mathrm{d,s}}$ according to Table 6.1a. The difference between the expected and 'measured' shear profiles in Fig. 6.3 is the bias that we estimated in Sect. 6.3.2 and gave in Eq. (6.16). We calculated the first and the third line of (6.16)

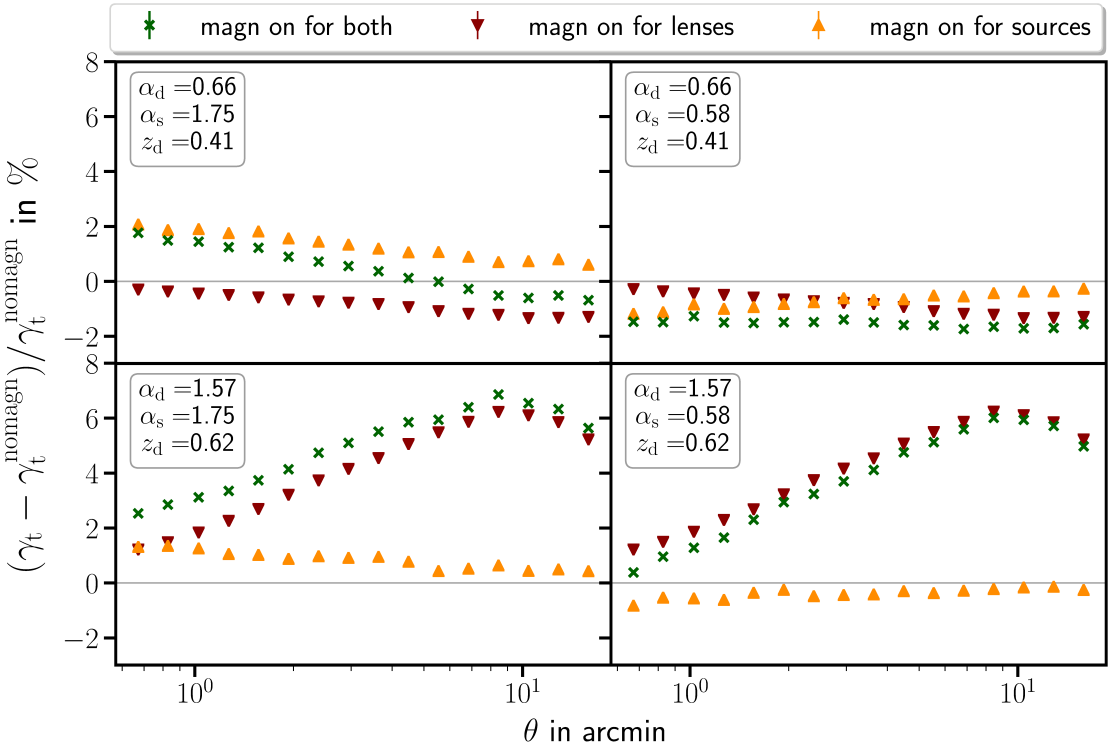


Figure 6.2. – Relative difference between shear profiles with and without magnification. The redshifts for the lenses are $z_d = 0.41$ in the upper panels and $z_d = 0.62$ in the lower panels; the source redshift is kept fixed at $z_s = 0.99$. The limiting magnitude of the lenses is 22 mag in the r-band, and for the sources it is 23 and 25 mag for $\alpha_s = 1.75$ and 0.58, respectively. The red 'downward' triangles indicate shear profiles that only have magnification in the lens galaxy population, while the orange 'upward' triangles show the influence of magnification for source galaxies only. It can be seen that a local slope > 1 of lens or source population leads to an enhanced signal, whereas $\alpha_{d,s} < 1$ causes a reduced signal. The green crosses display a measurement closest to real observations, i.e. where magnification affects both source and lens galaxy populations. A reduction or enhancement depends on both slopes $\alpha_{d,s}$, as well as the redshifts of lenses and sources $z_{d,s}$. In all cases, the shape of the shear profile changes.

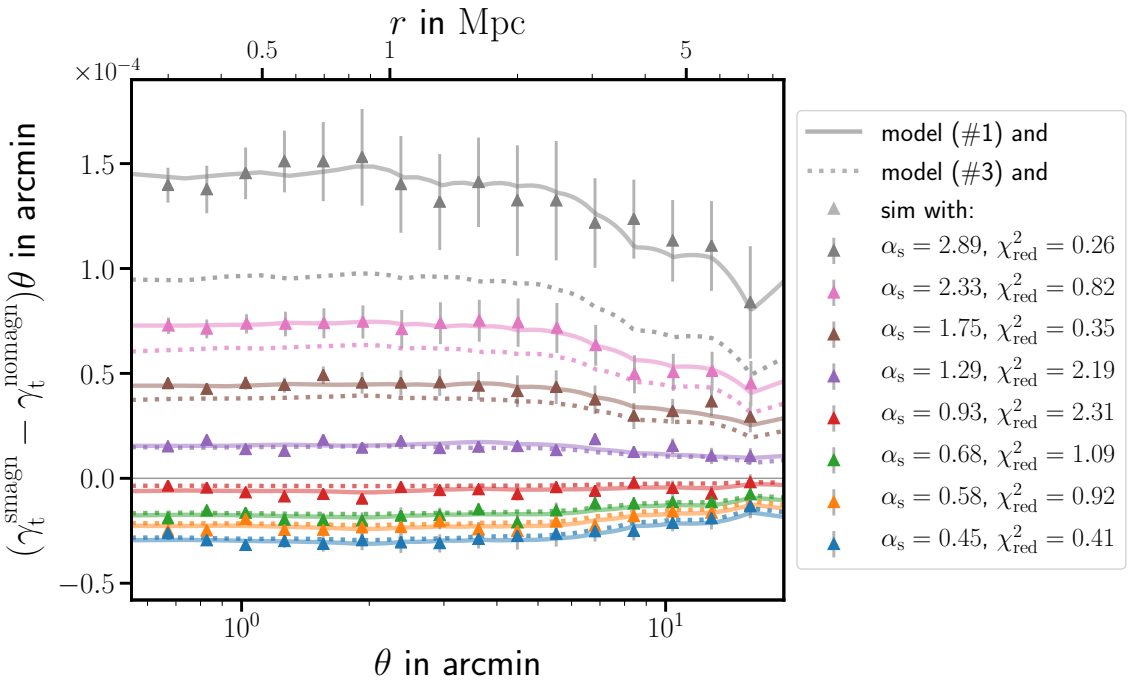


Figure 6.3. – Absolute difference between shear profiles with and without magnification of sources, for different local slopes α_s shown in the legend. The solid lines correspond to the first expression in Eq. (6.16) and dotted lines correspond to its approximation in the third line. The upper scale shows comoving transverse separation, and the redshifts are $z_s = 0.99$ for the sources and $z_d = 0.41$ for the lenses. The brown triangles with $\alpha_s = 1.75$ and the orange ones $\alpha_s = 0.58$ are directly comparable to the orange triangles in the upper panels of Fig. 6.2. We also show the goodness-of-fit parameter χ^2_{red} with 16 degrees of freedom for the solid line.

using the numerical data and show them as solid and dotted lines, respectively. Both models are in good agreement with the numerical data for moderate α_s. However, for very steep $\alpha_\mathrm{s} \gtrsim 2$, the weak lensing approximation $|\mu - 1| \ll 1$ is not sufficient anymore; although large magnifications are rare, they affect the number counts significantly.

We define the mean fractional difference between a shear profile with and without magnification for all bins as

$$\delta_\gamma = \frac{1}{N_\mathrm{bin}} \sum_{i=1}^{N_\mathrm{bin}} \frac{\gamma_{\mathrm{t}i} - \gamma_{\mathrm{t}i}^{\mathrm{nomagn}}}{\gamma_{\mathrm{t}i}^{\mathrm{nomagn}}} \,, \tag{6.20}$$

where we stress that a difference of $\delta_\gamma = 0$ is not necessarily equivalent to an unaltered shear profile. However, we only applied this estimator to the orange 'upward' and red 'downward' triangles seen in Fig. 6.2, which display either a positive or a negative sign for all angular scales investigated.

Results for δ_γ as a function of α_s for constant z_s can be seen in Fig. 6.4. We selected the source galaxies according to Table 6.1a, for which δ_γ is almost linear in α_s, as expected from Eq. (6.16) in the weak lensing approximation, although the slope of the linear relation depends on different z_d (Table 6.1c). The maximum shear difference of 4% is found for the largest α_s.

6.5. Magnification effects on foreground galaxies

In this section, we investigate the influence of magnification on lens galaxies. We follow the structure from the previous section, meaning we obtain and analyse results from the ray-tracing data and then compare those to the analytic estimate presented in Sect. 6.3.1.

6.5.1. Magnification switched on

The galaxies in the Henriques catalogue are affected by magnification by design. Hence, to create a catalogue including magnification, we simply extracted lens positions from galaxies brighter than a magnitude limit $s_\mathrm{lim,d}$ and assigned them to their nearest grid point.

6.5.2. Magnification switched off

To switch off magnification in the mock data, we undid the magnification as follows. As was done for the sources, the magnification-corrected flux s_0 can be easily recovered from the magnification given in the ray-tracing catalogue and the apparent flux of lens galaxies. The galaxy's apparent position, however, is shifted on the sky compared to its unlensed position. Unfortunately, the unlensed position cannot be

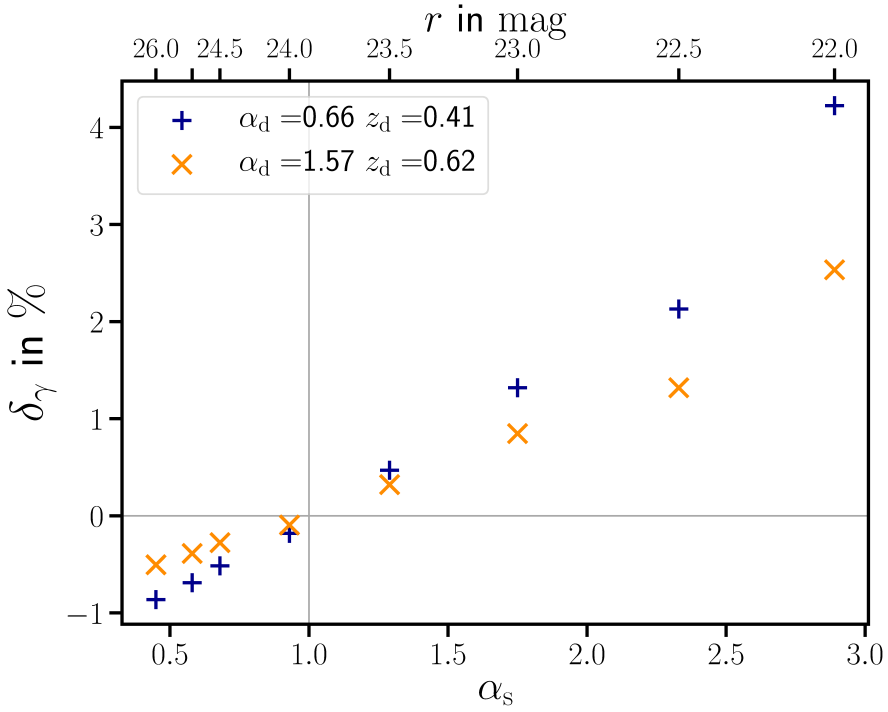


Figure 6.4. – Fractional change of the shear profile by magnification is displayed, where the magnification is only turned on for the source galaxies at $z_\mathrm{s} = 0.99$ (Table 6.1a). At a local slope of $\alpha_\mathrm{s} \approx 1$ the impact of magnification flips its sign. The relation between α_s and δ_γ is almost linear. Different redshifts of the lenses $z_\mathrm{d} = 0.41$ (plus) and 0.62 (cross) affect the impact of the sources' magnification; the r-band limiting magnitude of lenses is fixed to 22 mag.

recovered from the simulated data, and the absolute amount of shifting is of the order of an arcminute and depends on redshift (Chang and Jain, 2014). Hence, we must aim to create a sample of galaxies that is not affected by magnification in a different way.

As was outlined in Sect. 6.2.2, for both local slopes $\alpha_{\mathrm{d,s}} = 1$, the magnification leaves the observed number counts unaffected, and thus the shear profile γ_{t} unbiased. Therefore, we transformed the magnification-corrected flux distribution such that the galaxy counts obey $n_0' \propto {s_0'}^{-1}$. As a reference point, we chose the number of galaxies at limiting magnitude $n_0(> s_{\mathrm{lim,d}}) = n_0'(> s_{\mathrm{lim,d}}')$. This results in the following mapping from observed magnification-corrected flux, s_0, to the transformed flux

$$s_0'(s_0) = s_{\mathrm{lim,d}} \frac{n_0(s_{\mathrm{lim,d}})}{n_0(s_0)} \,. \tag{6.21}$$

In other words, this new flux scale distorts the number counts of lenses in such a way that in the transformed flux system, $\alpha' = 1$, the lens galaxy counts are unaffected by magnification bias (although the individual lens positions are not). We can now calculate the observed transformed number density $n'(> s')$, where $s' = \mu\, s_0'$. To create a lens galaxy sample free from magnification, we again chose only those galaxies that are brighter than the given flux limit $s' > s_{\mathrm{lim,d}}$. This leads to a different selection of galaxies for the original and the transformed number density.

We tested this approach with two consistency checks. The first is based on the fact that in the method described above, the total number of galaxies has to be conserved. This is true for all the lens redshifts used. For a magnitude cut of 22 mag in the r-band, a lens redshift of $z_{\mathrm{d}} = 0.41$, and the full field of view of $64 \times 16\,\mathrm{deg}^2$, 595 348 lenses are found with a cut in observed magnitude, and 595 355 lenses are found with observed transformed magnitude. Compared to the original fluxes, a detailed analysis showed that in the transformed flux system, 917 galaxies became brighter than 22 mag, while 924 galaxies became dimmer, leaving the overall number count almost unchanged. The tiny difference in seven galaxies is due to the fact that the number count function is discretely sampled.

The second consistency check uses a null test: the so-called shear-ratio test (SRT, Jain and Taylor, 2003) which is based on Eq. (6.18),

$$T(\theta; z_{\mathrm{d}}, z_{\mathrm{s}_1}, z_{\mathrm{s}_2}) := \frac{\gamma_{\mathrm{t}}(\theta; z_{\mathrm{d}}, z_{\mathrm{s}_1})}{\gamma_{\mathrm{t}}(\theta; z_{\mathrm{d}}, z_{\mathrm{s}_2})} - \frac{D_{\mathrm{ds}_1}}{D_{\mathrm{s}_1}} \frac{D_{\mathrm{s}_2}}{D_{\mathrm{ds}_2}} \,, \tag{6.22}$$

for which we expect $T(z_{\mathrm{d}}, z_{\mathrm{s}_1}, z_{\mathrm{s}_2}) = 0$ for two source populations at redshifts z_{s_1} and z_{s_2} in the absence of magnification effects. For the test, the location of the same lens galaxies and the shear from two source galaxy populations at different distances $D_{\mathrm{s}_{1,2}}$ were used. A ratio of the tangential shear estimates is equal to the ratio of the corresponding angular-diameter distances, while the lens properties drop out. As shown in Unruh et al. (2019), the SRT is strongly affected by lens magnification.

The impact is stronger for higher lens redshifts and smaller line-of-sight separation of lenses and sources. Therefore, we performed the SRT for lenses selected with and without magnification for two different lens redshifts. We performed the SRT by taking a weighted integral of $T(\theta; z_\mathrm{d}, z_{\mathrm{s}_1}, z_{\mathrm{s}_2})$ over θ from θ_in to θ_out as in Unruh et al. (2019). In the case that includes magnification, we recovered the results from Unruh et al. (2019). For lenses that are selected with corrected magnification, the SRT performs better by a factor of $\gtrsim 100$. We give results for two example redshift combinations in Table 6.2. The corrected SRT still shows a slight scatter due to the statistical noise in the data, coming from the lensing by the large-scale structure in each of the 64 fields, and a bias that arises from shifting lens galaxies to their nearest grid point.

Table 6.2. – Shear-ratio test (6.22) performed for two example cases to demonstrate the removal of the magnification from the lens galaxies.

redshifts			SRT result $T(z_\mathrm{d}, z_{\mathrm{s}_1}, z_{\mathrm{s}_2})$	
z_d	$z_{\mathrm{s}1}$	$z_{\mathrm{s}2}$	magn	no magn
0.41	0.46	0.51	$(5.6 \pm 1.4) \times 10^{-2}$	$(7.0 \pm 13.7) \times 10^{-4}$
0.83	0.91	0.99	$(1.3 \pm 0.2) \times 10^{-1}$	$(-5.1 \pm 10.5) \times 10^{-4}$

6.5.3. Results

The red 'downward' triangles in Fig. 6.2 show the relative impact of magnification on γ_t. Similar to the results given in Sect. 6.4.3, where the magnification of the source galaxies is discussed, the local slope α_d determines whether the shear signal is enhanced or reduced. The upper panels of Fig. 6.2 show that the shear profiles are reduced for a local slope α_d that is smaller than unity at redshift $z_\mathrm{d} = 0.42$, while the lower panels display shear profiles with $\alpha_\mathrm{d} > 1$ at a higher redshift $z_\mathrm{d} = 0.62$, where the reverse effect is observed. The panels with higher z_d show larger magnification effects; relative deviations by up to 7% in a single bin can be seen. In general, the shear signal is more strongly affected at larger separations θ from the lens centre until it reaches a maximum at $\approx 8'$ for $z_\mathrm{d} = 0.62$ and $\approx 10'$ for $z_\mathrm{d} = 0.41$; for even larger separations, the magnification effect becomes relatively weaker.

A comparison of numerical results to the analytic estimate (6.13) can be seen in Fig. 6.5. The absolute difference between shear profiles affected by a magnification of lens galaxies and those unaffected by magnification is plotted. The triangles show numerical results, while lines indicate our analytical model for $2\,[\alpha_\mathrm{d}(z_\mathrm{d}) - 1]\,\gamma_\mathrm{t}^\mathrm{LSS}$. We employed the reduced χ^2-test as an estimator for the goodness of our model and find that all models are in good agreement with the data for the considered angular scales. However, the local slope is not necessarily a sufficiently good quantity for the analytic correction if the local slope α_d becomes very steep, meaning when the

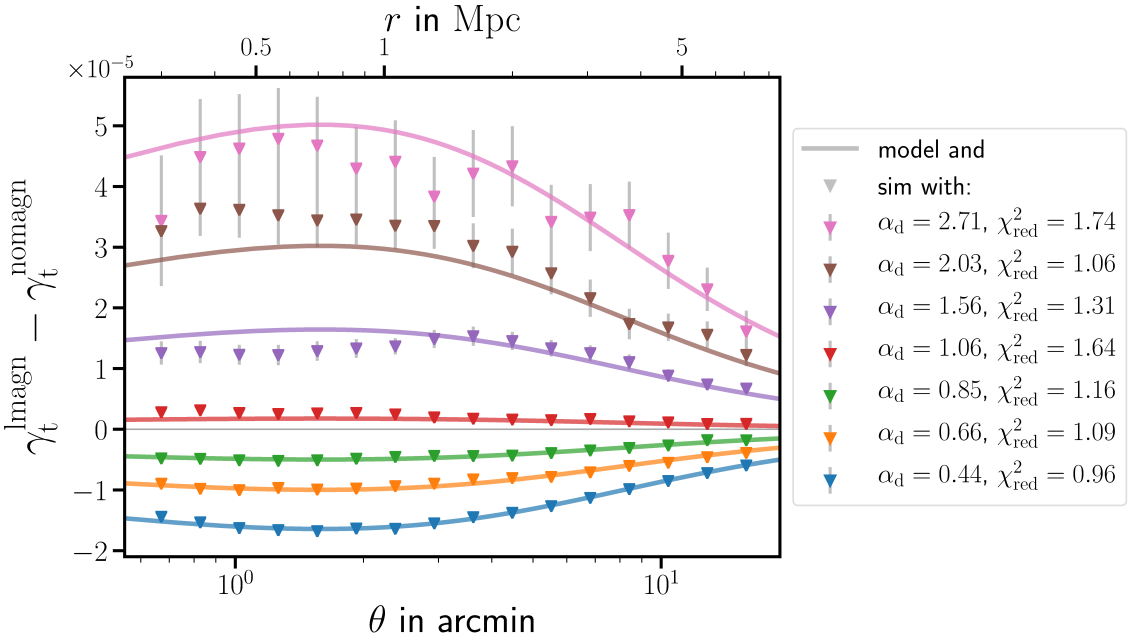


Figure 6.5. – Absolute difference between shear profiles with and without magnification of lenses. The upper scale shows comoving transverse separation and the shear difference is shown for several limiting magnitudes of the lens galaxies, which yield the local slopes α_d shown in the legend. Also, the goodness-of-fit parameter χ^2_{red} with 16 degrees of freedom is indicated in the legend. The redshifts are $z_s = 0.99$ for the sources and $z_d = 0.41$ for the lenses.

luminosity function is not well approximated by a power law anymore (cf. Fig. 6.1). However, the analytic correction still reduces the impact of magnification significantly.

To explore the dependencies of the mean fractional shear difference δ_γ (Eq. 6.20) on α_d, we altered the lens properties according to Table 6.1b. Results can be seen

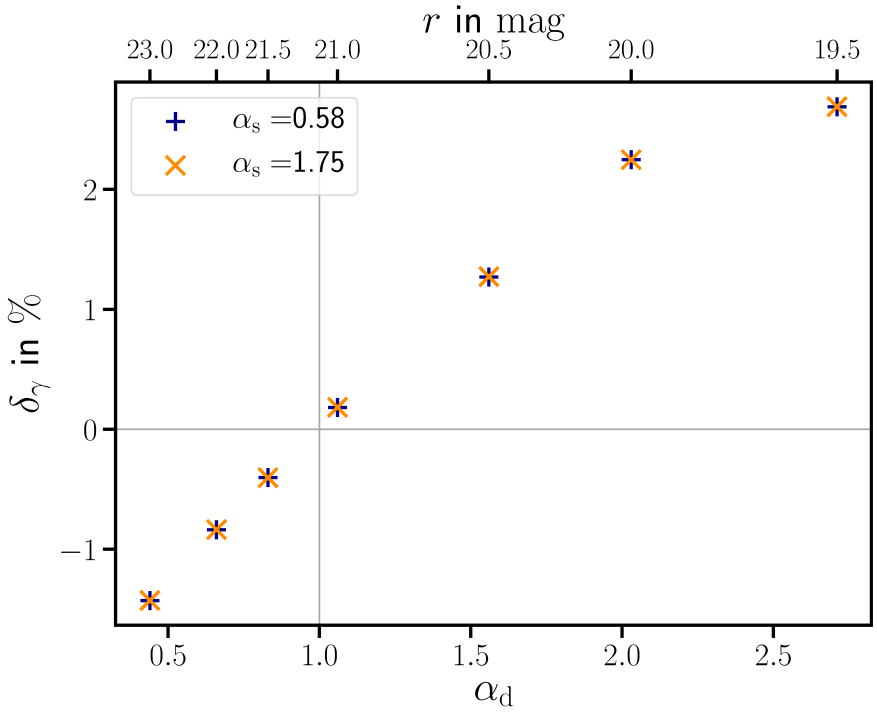


Figure 6.6. – To change the lens properties according to Table 6.1b, we show the behaviour of mean fractional shear difference δ_γ as a function of local slope for lens galaxies α_d and r-band-limiting magnitude. The source and lens redshift is the same for both cases, i.e., $z_\mathrm{d} = 0.41$ and $z_\mathrm{s} = 0.99$. The effect is independent of the sources local slope α_s.

in Fig. 6.6 and show the impact of magnification for constant lens redshift z_d and varying limiting magnitude, for two different local slopes α_s of the sources. The impact on the shear profile is almost exactly the same in both local slopes, with a small residual noise that is present in the data. The dependence of δ_γ on α_d is similar to the one in Fig. 6.4, which shows the magnification for source galaxies only. The shear profile changes by up to 3% in the mean.

We further investigated the dependence of the magnification on the values of α_d and z_d for fixed z_s in Fig. 6.7. δ_γ is calculated for various lens redshifts with constant limiting magnitude for the lenses (see Table 6.1c) and the same two values of α_s as before. The figure shows that the sources' local slope does not affect the magnification induced by lens galaxies. The signal is again moderately reduced by 1% for $\alpha_\mathrm{d} < 1$ but is not monotonic anymore. For $\alpha_\mathrm{d} > 1$, the relation deviates

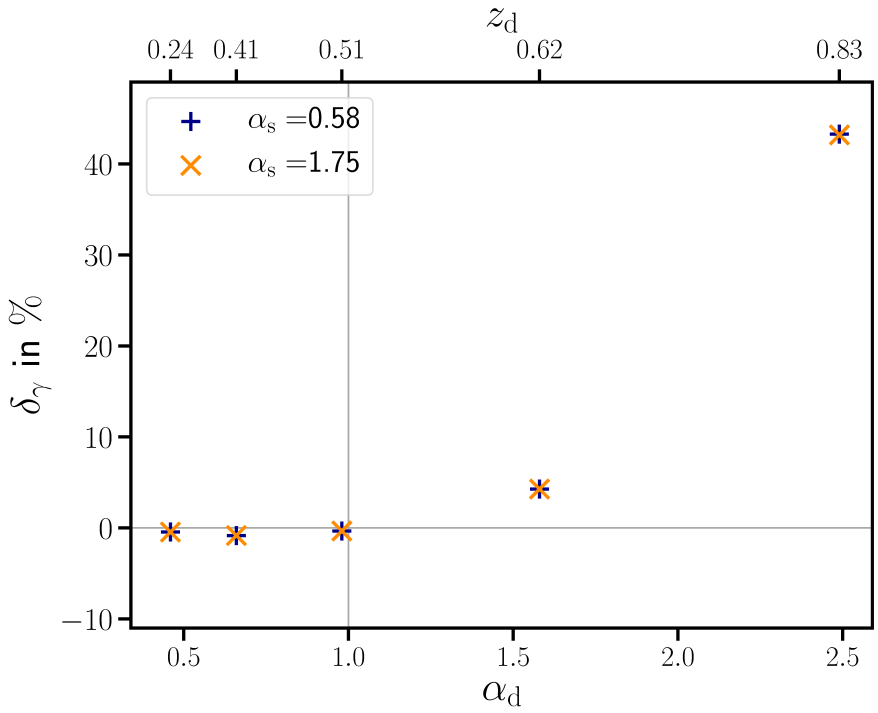


Figure 6.7. – For a magnification of lens galaxies only, the impact on the shear profile is shown as a function of α_d and z_d (Table 6.1c) for fixed flux limit of $r = 22\,\mathrm{mag}$. The redshift z_d is given on the (non-linear) top axis. Again, the effect switches its sign at $\alpha_\mathrm{d} \approx 1$. It then rises up to 45% for larger α_d, while less impact is seen for $\alpha_\mathrm{d} < 1$. Different local slopes of the sources α_s at $z_\mathrm{s} = 0.99$ leave the lens' magnification unaffected.

significantly from a linear one. Magnification is stronger for larger α_d and larger z_d, leading to deviations of up to 45% in the most extreme case.

Lastly, we combined the numerical methods and results that were addressed individually in Sects. 6.4 and 6.5. The green crosses in Fig. 6.2 show numerical results from source and lens galaxy populations that are both affected by magnification, meaning the case of relevance for observational studies of GGL. The fractional change is approximately the sum of the bias of lens galaxies only plus the bias of source galaxies only. Hence, to first order, it is determined by α_d and α_s. The sign of δ_γ cannot easily be predicted if $\alpha_\mathrm{d} < 1$ and $\alpha_\mathrm{s} > 1$, and vice versa. For δ_γ the sign also depends on θ since the change of shear signal per bin behaves differently for magnified lenses and magnified sources.

6.6. Magnification bias in halo-mass estimates

6.6.1. Estimating the mean halo mass of lenses

The GGL signal is sensitive to the surface-mass density around lenses that differs from the average projected cosmological matter density. To infer from this the mean mass of a parent halo that hosts a typical lens galaxy, we used a halo-model prescription to describe the relation between galaxies and matter (Cooray and Sheth, 2002). In this prescription, we expand the galaxy-matter power spectrum at redshift z_d of lenses and at comoving wave number k,

$$P_\mathrm{gm}(k) = P_\mathrm{gm}^\mathrm{1h}(k) + P_\mathrm{gm}^\mathrm{2h}(k) \;, \tag{6.23}$$

in terms of a one-halo term,

$$P_\mathrm{gm}^\mathrm{1h}(k) = \int_0^\infty \frac{\mathrm{d}m\, n(m)\, m}{\Omega_\mathrm{m}\, \rho_\mathrm{crit}\, \bar{n}_\mathrm{d}}\, \tilde{u}_\mathrm{m}(k,m) \left(\langle N_\mathrm{cen}|m\rangle + \langle N_\mathrm{sat}|m\rangle\, \tilde{u}_\mathrm{m}(k,m) \right) , \tag{6.24}$$

and a two-halo term,

$$\begin{aligned} P_\mathrm{gm}^\mathrm{2h}(k) = &\int_0^\infty \frac{\mathrm{d}m\, n(m)\, m\, b_\mathrm{h}(m)}{\Omega_\mathrm{m}\, \rho_\mathrm{crit}}\, \tilde{u}_\mathrm{m}(k,m)\, P_\mathrm{lin}(k) \\ &\times \int_0^\infty \frac{\mathrm{d}m\, n(m)\, b_\mathrm{h}(m)}{\bar{n}_\mathrm{d}} \left(\langle N_\mathrm{cen}|m\rangle + \langle N_\mathrm{sat}|m\rangle\, \tilde{u}_\mathrm{m}(k,m) \right) . \end{aligned} \tag{6.25}$$

In this model, $\tilde{u}_\mathrm{m}(k,m)$ denotes the Fourier transform of a Navarro-Frenk-White (NFW; Navarro et al., 1996) density profile for a virial halo mass m, truncated at the virial radius and normalised to $\tilde{u}_\mathrm{m}(k,m) = 1$ for $k = 0$ (Scoccimarro et al., 2001) for the mass-concentration relation in Bullock et al. (2001). We further denote the mean comoving number density of halos in the mass interval $m_1 \leq m < m_2$ by $\int_{m_1}^{m_2} \mathrm{d}m\, n(m)$ (Sheth and Tormen, 1999), the bias factor of halos of mass m by

$b_\mathrm{h}(m)$ (Scoccimarro et al., 2001), and the linear matter power spectrum by $P_\mathrm{lin}(k)$ (Eisenstein and Hu, 1998). Finally, the mean number density of lenses is

$$\bar{n}_\mathrm{d} = \int_0^\infty \mathrm{d}m\, n(m) \left(\langle N_\mathrm{cen}|m\rangle + \langle N_\mathrm{sat}|m\rangle \right) , \tag{6.26}$$

This version of the halo model assumes a central galaxy at the centre of a halo whenever there are lens galaxies inside the halo, and satellite galaxies with a number density profile equal to the NFW matter density. For the mean number of central and satellite galaxies for a halo mass m, we follow Clampitt et al. (2017) but with central-galaxy fraction $f_\mathrm{cen} \equiv 1$,

$$\langle N_\mathrm{cen}|m\rangle = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log_{10}(m/m_\mathrm{th})}{\sigma_{\log m}}\right)\right] ; \tag{6.27}$$

$$\langle N_\mathrm{sat}|m\rangle = \langle N_\mathrm{cen}|m\rangle \left(\frac{m}{m_1}\right)^\beta , \tag{6.28}$$

where $\boldsymbol{\Theta} = (m_1, m_\mathrm{th}, \sigma_{\log m}, \beta)$ are four model parameters that determine the halo-occupation distribution (HOD) of our lenses, and $\mathrm{erf}(x) = 2\pi^{-1/2}\int_0^x \mathrm{d}t\, \mathrm{e}^{-t^2}$ is the error function. The model parameters have the following meaning: m_th determines at which mass scale $\langle N_\mathrm{cen}|m\rangle = 1/2$; at halo mass m_1 the mean number of satellites equals that of central galaxies; $\sigma_{\log m}$ is the width of the HOD of centrals; and β is the slope of the satellite HOD.

The matter-galaxy cross-power spectrum is related to the mean tangential shear by a Limber projection, which for lenses at z_d and sources at z_s is

$$\gamma_\mathrm{t}(\theta) = \frac{3H_0^2\,\Omega_\mathrm{m}}{2c^2}\,\frac{D_\mathrm{ds}}{D_\mathrm{d}\,D_\mathrm{s}} \int_0^\infty \frac{\mathrm{d}\ell\,\ell}{2\pi}\, \mathrm{J}_2(\ell\theta)\, P_\mathrm{gm}\left(\frac{\ell+1/2}{(1+z_\mathrm{d})\,D_\mathrm{d}}; z_\mathrm{d}\right) . \tag{6.29}$$

We employed a maximum-likelihood estimator (MLE) to infer the mean halo mass of lenses from GGL. For this, we set $\{\gamma_{\mathrm{t},i}(\boldsymbol{\Theta})\,|\,i = 1\ldots N_\mathrm{bin}\}$ as a set of N_bin measurements of the mean tangential shear in our mock data, obtained for different lens-source angular separation bins i; the error covariance estimated from the measurements for bin θ_i and θ_j is $C_{ij,}$ and its inverse $[\mathsf{C}^{-1}]_{ij}$. For the MLE of $\boldsymbol{\Theta}$, we then minimised

$$\chi^2(\boldsymbol{\Theta}) = \sum_{i,j=1}^{N_\mathrm{bin}} \Big(\gamma_{\mathrm{t},i}(\boldsymbol{\Theta}) - \gamma_{\mathrm{t},i}\Big)[\mathsf{C}^{-1}]_{ij}\Big(\gamma_{\mathrm{t},j}(\boldsymbol{\Theta}) - \gamma_{\mathrm{t},j}\Big) , \tag{6.30}$$

with respect to $\boldsymbol{\Theta}$, where $\gamma_{\mathrm{t},i}(\boldsymbol{\Theta})$ is the halo model prediction of $\gamma_\mathrm{t}(\theta|\boldsymbol{\Theta})$ averaged over the size of the ith separation bin that corresponds to $\gamma_{\mathrm{t},i}$. We refer to $\boldsymbol{\Theta}_\mathrm{mle}$ as the parameter set that minimises $\chi^2(\boldsymbol{\Theta})$. Finally, given the MLE $\boldsymbol{\Theta}_\mathrm{mle}$, we obtain the MLE of the mean halo mass by the integral

$$\langle m\rangle_\mathrm{mle} = \int_0^\infty \frac{\mathrm{d}m\, n(m)\, m}{\bar{n}_\mathrm{d}} \left(\langle N_\mathrm{cen}|m\rangle + \langle N_\mathrm{sat}|m\rangle \right) , \tag{6.31}$$

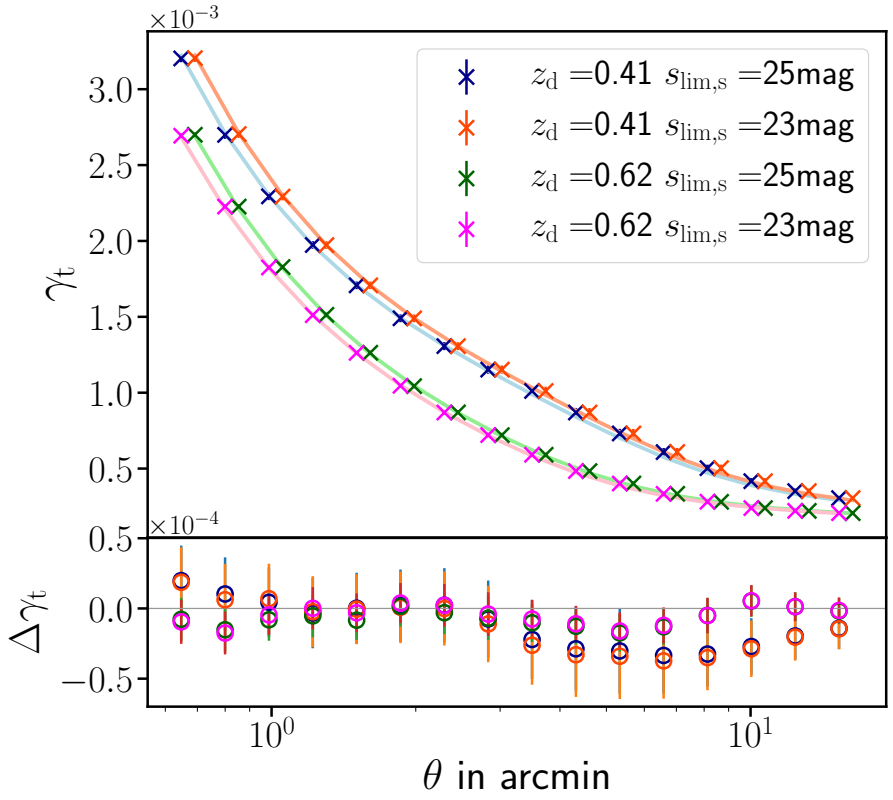


Figure 6.8. – Upper panel: Magnification-corrected shear profiles for two lens redshifts z_d with a r-band magnitude cut of 22 mag and two differently selected source galaxies at redshift $z_\mathrm{s} = 0.99$ are shown (crosses) as well as their best fit according to the halo model (solid line). Lower panel: Absolute difference $\Delta\gamma_\mathrm{t}$ between data and halo model fit (open circles with the same colour). For each lens redshift, source galaxies were chosen with two limiting magnitudes that have the local slopes $\alpha_\mathrm{s} = 0.58$ and 1.75, respectively. As expected, the shear profile does not depend on the choice of source galaxies. The offset between the red/blue points as well as the magenta/green points is for better visibility only.

which has to be evaluated for the HOD of galaxies that is determined by $\boldsymbol{\Theta}_{\rm mle}$.

In Fig. 6.8, we show four examples of tangential shear estimates and their best shear profile fit from the halo model. For the shear profiles, we switched off magnification effects of lens and source galaxies. It can be seen that the fitting procedure works reasonably well. The halo model itself is only an approximation of the inhomogeneous matter distribution in the Universe, and in the presence of our simulated data with almost vanishing Poisson noise, we do not expect the halo model to work perfectly. We further used approximations such as a truncated Navarro-Frenk-White (NFW) model and a specific mass-concentration relation, that certainly further limits the accuracy we can obtain with the halo model. The mean relative difference between the data and the model for the 16 bins is 2.6% for $z_{\rm d} = 0.41$, and 1.3% for $z_{\rm d} = 0.62$. As expected, the shear profiles are almost independent of the flux limit of sources if their redshift is fixed, as can be seen for the red and blue crosses, as well as for the magenta and green crosses. On the other hand, there are three main reasons for the difference between the red/blue and magenta/green shear profiles. The lensing efficiencies $D_{\rm ds}/D_{\rm s}$ are different for different $z_{\rm d}$. Thus, the red/blue shear profile with $z_{\rm d} = 0.41$ has a larger lensing efficiency than the magenta/green one with $z_{\rm d} = 0.62$. Moreover, we observed fixed angular scales, which corresponds to different physical scales. Lastly, the lens galaxy population might evolve between the two redshifts.

Table 6.3 accompanies Fig. 6.8 and lists the fitting parameters, the goodness-of-fit values, and the mass estimates for the different shear profiles. The mean halo mass, in contrast to the shear amplitude, is larger for the high-redshift lenses when the same magnitude limit is applied.

We quantify this bias in halo mass for fixed limiting magnitudes of lenses and sources, which means fixed $\alpha_{\rm d,s}$, and fixed redshifts as follows. Using the halo model as described above, we calculated the best mass estimate from the magnification-corrected shear profile. As could be seen in the previous sections, the relative change of the shear profile is typically of the order of a couple of percent. Therefore, we fixed the scatter in the host halo mass $\sigma_{\log m}$ and the slope of the mean number of satellite galaxies β to their best fit value in the magnification-corrected case. Then, we only fitted the remaining two parameters $m_{\rm th}$ and m_1 to estimate the mass for the three remaining shear profiles, meaning a shear profile with magnification of the sources only turned on, a profile with magnification of the lenses only turned on, and a shear profile with lens and source magnification turned on. Similar to the fractional shear difference (6.20), we define the bias of halo-mass estimates, inferred from $\gamma_{\rm t}$, by

$$\delta_M = \frac{\langle m \rangle_{\rm mle} - \langle m \rangle_{\rm mle}^{\rm nomagn}}{\langle m \rangle_{\rm mle}^{\rm nomagn}} \,. \tag{6.32}$$

Table 6.3. – Fitting results for the halo model, with the mean halo mass $\langle m \rangle_{\rm mle}$, the scatter in host halo mass $\sigma_{\log m}$, the mass scale where 50% of halos host a galaxy $m_{\rm th}$, the normalisation factor for the satellite galaxies m_1 and its slope β, and the goodness-of-fit value $\chi^2_{\rm red}$ with 12 degrees of freedom.

	plot colour			
	blue	red	green	magenta
$z_{\rm d}$	0.41	0.41	0.62	0.62
$\alpha_{\rm d}$	0.67	0.67	1.57	1.57
$\alpha_{\rm s}$	0.58	1.75	0.58	1.75
$\langle m \rangle_{\rm mle}$ in $10^{13}\,M_\odot$	2.45	2.46	2.58	2.60
$\sigma_{\log m}$	0.28	0.28	0.28	0.30
$m_{\rm th}$ in $10^{11}\,M_\odot$	2.24	2.26	4.15	4.19
β	1.08	1.09	1.06	1.05
m_1 in $10^{12}\,M_\odot$	7.70	7.68	11.54	11.92
$\chi^2_{\rm red}$	1.52	1.84	1.10	1.25

Notes. The lens and source redshifts, and local slopes $\alpha_{\rm d,s}$ are chosen as in Fig. 6.8, the source redshift $z_{\rm s}$ is 0.99. The fit values for identical lens redshifts are expected to be very similar.

6.6.2. Numerical results

Figures 6.9 to 6.12 show results for the mass bias δ_M. Figure 6.9 shows the bias for magnified source galaxy counts and magnification-corrected lens counts. The lenses have constant limiting magnitude of 22 mag in the r-band and their redshifts are $z_{\rm d} = 0.41$ and 0.62. Source galaxies at redshift $z_{\rm s} = 0.99$ are selected for several limiting magnitudes (cf. Table 6.1a). The bias is of the same order of magnitude as the corresponding mean fractional difference of the shear, and $\alpha_{\rm s}$ determines whether mass is overestimated or underestimated. The mean halo mass is biased by up to 3.5%.

We then explored the dependencies of the fractional mass bias δ_M on $\alpha_{\rm d}$, while we only considered magnification-corrected sources. Firstly, we fixed the lens redshift to $z_{\rm d} = 0.41$ and altered the lens' limiting magnitude according to Table 6.1b. The result is shown in Fig. 6.10. The mass bias is an almost linear function of $\alpha_{\rm d}$ and shows a similar dependence on $\alpha_{\rm d}$ as the fractional shear difference (cf. Fig. 6.6). The mass is biased up to 5% for the largest $\alpha_{\rm d}$. Lastly, we calculated δ_M for various lens redshifts with constant limiting magnitude for the lenses (see Table 6.1c) and the same two $\alpha_{\rm s}$ as before, which is shown in Fig. 6.11. The mass bias shows a strong redshift dependence, where the bias increases from a couple of percent to a mass overestimate of 55%.

To explore the observationally relevant case, we compared halo-mass estimates with and without magnification for both lenses and sources. Figure 6.12 contains all different $\alpha_{\rm d,s}$-$z_{\rm d}$ combinations with constant $z_{\rm s} = 0.99$ from the Tables 6.1a to

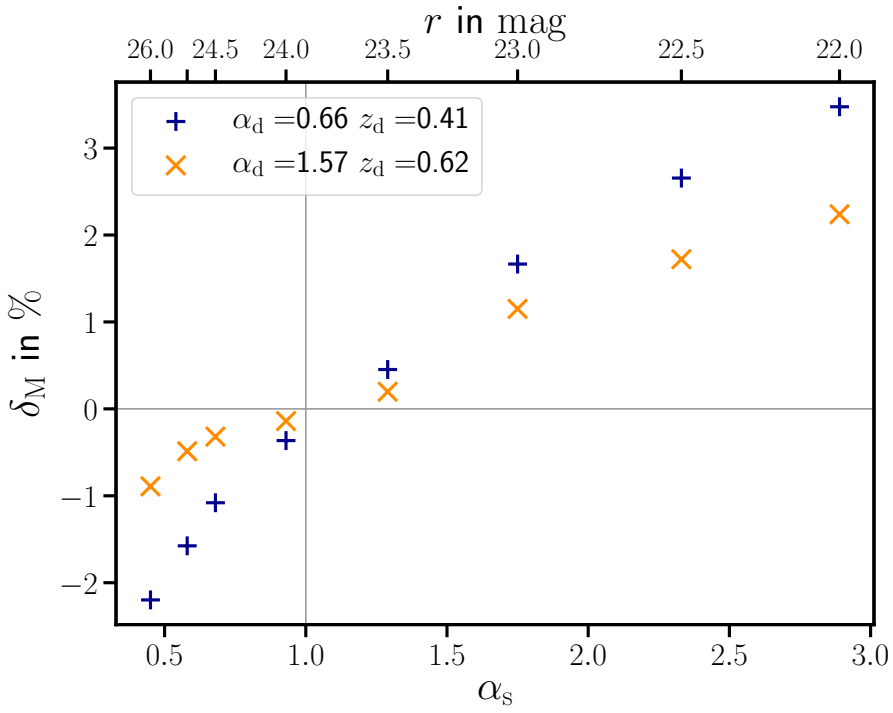


Figure 6.9. – Relative mass bias δ_M (6.32) for magnified source number counts is shown. Source galaxies are at $z_s = 0.99$ and lenses are chosen according to Table 6.1a with redshifts $z_d = 0.41$ (plus) and 0.62 (cross), for fixed limiting magnitude. The mass bias behaves roughly like the fractional shear difference (cf. Fig. 6.4). For local slopes $\alpha_s < 1$, the underestimation of mass is stronger than for the shear profile, while for $\alpha_s > 1$ the overestimate is similar.

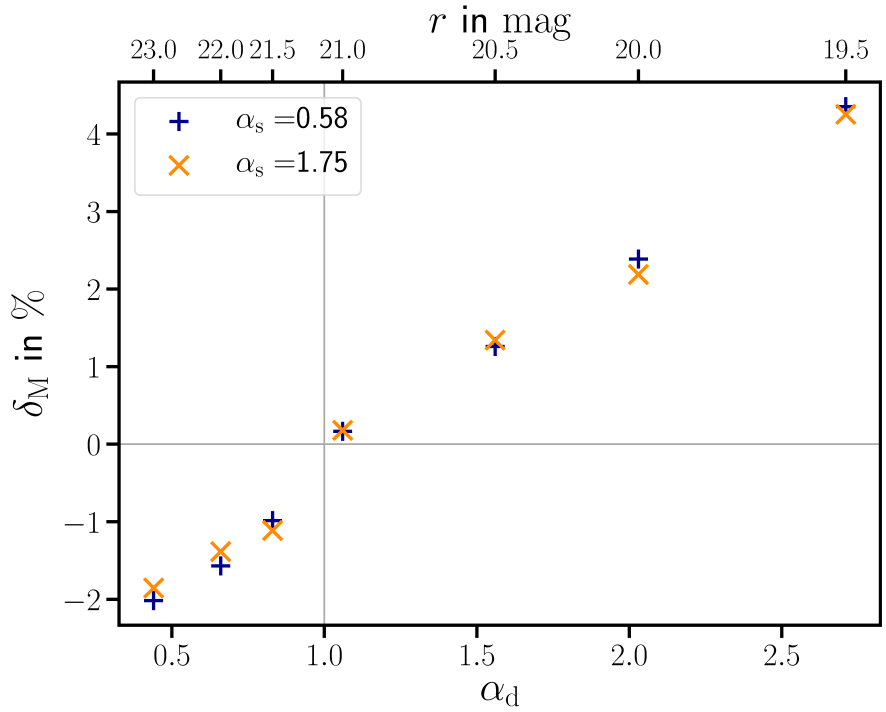


Figure 6.10. – The fractional mass bias δ_M is shown as a function of local slope α_d and r-band limiting magnitude (cf. Table 6.1b). The source redshift is $z_s = 0.99$ and lens redshift is $z_d = 0.41$. For $\alpha_d < 1$, δ_M the mass is biased low, while for $\alpha_d > 1$ the mass is biased high.

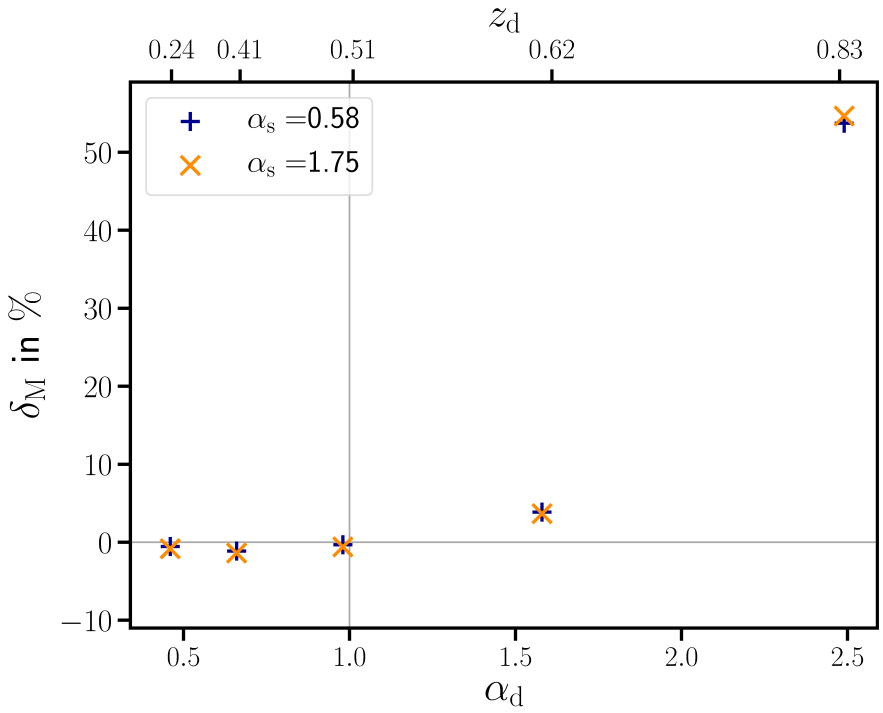


Figure 6.11. – The fractional mass bias δ_M is shown as a function of local slope α_d and lens redshift z_d(cf. Table 6.1c). We keep the r-band limiting magnitude for the lenses and the source redshift $z_\mathrm{s} = 0.99$ constant. The top axis indicates the respective lens redshifts in a non-linear scaling. Following the trend seen in Fig. 6.7, mass is biased low for $\alpha_\mathrm{d} < 1$ and shows large biases for $\alpha_\mathrm{d} > 1$.

6.1c, plus some additional combinations. It shows δ_M as defined in Eq. (6.32) as

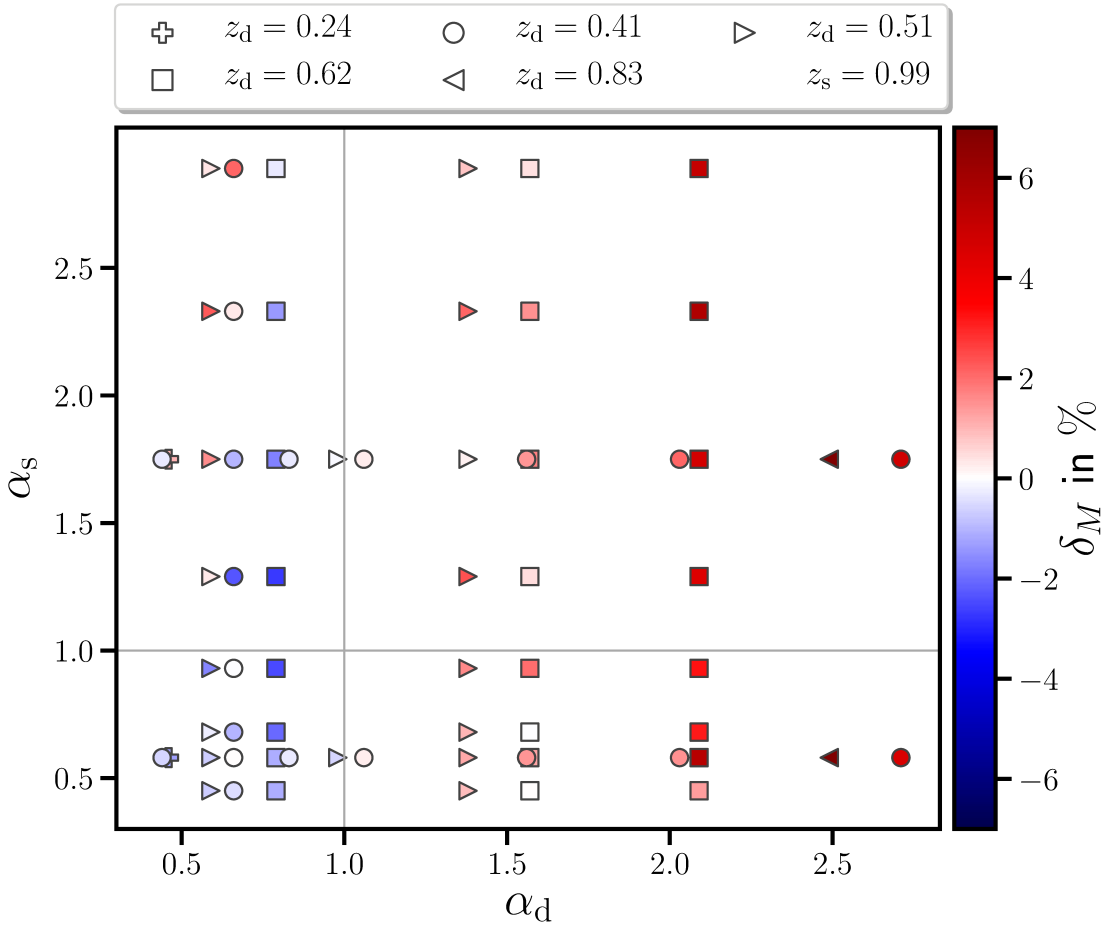


Figure 6.12. – The total magnification bias for halo-mass estimates as a function of limiting magnitudes that yield the local slopes $\alpha_{\mathrm{d,s}}$ and lens redshift z_d (type of symbols). The source redshift is the same in all cases, i.e., $z_\mathrm{s} = 0.99$. The mass difference δ_M is shown in colour code, where blue indicates underestimation, red overestimation, and white is an unbiased result. We cut off the colour bar for the largest deviations of 55% and 58% for $\alpha_\mathrm{d} = 2.41$ and $z_\mathrm{d} = 0.83$ for higher contrast in the colour scale. The plot is roughly divided by the vertical line with $\alpha_\mathrm{d} = 1$ into mass underestimation for $\alpha_\mathrm{d} < 1$ and overestimation $\alpha_\mathrm{d} > 1$.

a function of the local slopes $\alpha_{\mathrm{d,s}}$ on the x- and y-axis. Red values indicate an overestimation of the mass, where we cut off the colour bar for the highest values for better visibility. The blue values show an underestimation, while white values mean no bias in the mass estimate. The upper-right quadrant shows consistently red values since $\alpha_{\mathrm{d,s}} > 1$, while the lower-left quadrant has $\alpha_{\mathrm{d,s}} < 1$ and is consistently blue, as expected from theoretical considerations. In most of the other cases, α_d seems to be the decisive factor for the sign of the mass bias, for $\alpha_\mathrm{d} < 1$ leading to an underestimation and for $\alpha_\mathrm{d} > 1$ leading to an overestimation. The slope α_s

influences the total value of the mass bias. Besides, even if two points are close in the $\alpha_{\rm d}$-$\alpha_{\rm s}$ parameter space, they do not exhibit the same colour, meaning mass bias, due to their different $z_{\rm d}$.

6.7. Discussion & conclusions

In this paper, we quantified the impact of magnification on the GGL signals and halo-mass estimates. Magnification changes the observed galaxy number counts on the sky, which has an impact on the measured tangential shear profiles. It is important to note that magnification affects lens galaxies as well as source galaxies. Analyses of tangential shear profiles, such as estimates of the excess mass density profiles or halo mass, are therefore biased if they ignore lens or source magnification. Our estimates with ray-tracing simulations and synthetic galaxy populations show that the number count slopes of sources and lenses are the most important quantities that determine the relative strength of the bias. While the analytical estimate for the bias on the GGL profile caused by magnification of lenses was known before, to our knowledge the one caused by magnification of the source population has not previously been derived.

How magnification affects tangential shear profiles can be seen in Figs. 6.2 to 6.7, where we varied the local slope of lens and sources galaxies $\alpha_{\rm d,s}$ and the lens redshift $z_{\rm d}$. We studied the magnification effect from lenses and sources individually and compared them to our analytic estimates. The latter are leading-order estimates and describe the numerical results very well in most cases. One of the surprising results, shown in Fig. 6.3, is that the weak lensing approximation for the impact of magnification on source counts can significantly fail if the count slope is steep. Hence, the validity of the commonly used approximation for the number density of sources

$$n_{\rm s}(\boldsymbol{\theta}, > s) \approx n_{\rm s0}(> s)\,[1 + 2(\alpha_{\rm s} - 1)\kappa(\boldsymbol{\theta})]$$

needs to be checked, depending on its application.

Besides confirming that the shear signal is reduced if the slopes $\alpha_{\rm d}$ and $\alpha_{\rm s}$ are < 1 and enhanced for $\alpha_{\rm d,s} > 1$, we find: For fixed redshifts $z_{\rm d,s}$, the change of the tangential shear estimate depends solely on $\alpha_{\rm d,s}$. For fixed source redshifts $z_{\rm s}$, the impact of magnification of sources decreases for larger $z_{\rm d}$ and smaller lens-source separation $D_{\rm ds}$ (cf. Fig. 6.4). Furthermore, the relative importance of magnification of lenses increases with $z_{\rm d}$ and rises sharply as $z_{\rm d}$ approaches $z_{\rm s}$. The relative change from a biased to an unbiased shear profile is a function of angular separation θ, and the mean change is typically a few percent. However, if the redshift differences between sources and lenses become small, the effect can be considerably larger (cf. Fig. 6.7).

For practical applications, the change in the shear signal by magnification is described by the sum of the individual effects of source magnification and the one from

lens magnification. The impact of magnification depends on the galaxy's luminosity function, the limiting magnitudes, and the redshifts $z_{\rm d,s}$, but also the separation θ.

The relative mass bias δ_M on the average mass of a lens parent-halo inherits all trends seen in the shear profiles with magnified lenses and magnified sources (see Figs. 6.9 to 6.12), meaning $\delta_M > 0$ for $\alpha_{\rm d,s} > 0$ and vice versa. Medium-redshift galaxies show a mass bias of less than 10%, and the higher the lens redshift, the stronger the mass bias with up to 58% for $z_{\rm d} = 0.83$ and $z_{\rm s} = 0.99$. The bias δ_M is of the same order of magnitude as the relative change of the shear profile δ_γ. The particular relation, however, between δ_γ and the mass bias δ_M is quite complicated. First and foremost, the amplitude of the measured shear signal determines the underlying halo mass. However, magnification changes the scale dependence of the shear signal. In the halo model, this translates to a different behaviour of the one- and two-halo term, which affects the mass estimate in a highly non-linear way. Another minor effect is that the true mean halo mass of lens galaxies affected by magnification probably differs from the true mean halo mass of unmagnified lenses due to an expected correlation between mass and luminosity. For example, for the highest $z_{\rm d}$ considered in this work, the mean masses differ by 0.16%. Furthermore, we fixed the observed angular scales on the sky, which relate to different physical scales at different redshifts; thus, the relative contribution of the two-halo term on the shear signal grows with redshift, which makes a comparison of mass biases from different lens redshifts more complicated. In general, when we allow for magnification effects both in source and lens galaxies, meaning, the case for real observations, the sign of the mass bias is in most cases determined by the value of $\alpha_{\rm d}$. The only exceptions shown in Fig. 6.12 are two cases, where either the lens redshift is low or the local slope $\alpha_{\rm s}$ is very high; in such cases, the sign of the mass bias is not easily predicted and must be studied case by case.

We also considered a GGL estimate using the flux-limits and redshift distributions as given from the combined data of KiDS+VIKING (Hildebrandt et al., 2020) and GAMA (Driver et al., 2011). KiDS and VIKING are partner surveys that probe the optical and near-infrared sky to obtain high-resolution, wide-field shape and redshift information from galaxies. GAMA is a spectroscopic, flux-limited survey with a partly overlapping footprint in the KiDS+VIKING area. In Appendix 6.8.3, we present a detailed description of the input parameters we use for this estimate. We list the results in Table 6.4. A lens galaxy sample at redshift $z_{\rm d} = 0.21$ shows relative changes in the shear profile due to magnification effects of less than one percent, and relative mass bias of $\approx 3\%$, while GAMA's highest lens sample at $z_{\rm d} = 0.36$ shows that a magnification correction changes the shear profile by $\approx 2\%$ and the relative halo mass bias by 8%. Although our estimate used simplifying assumptions, especially for the source population, for example, no catastrophic outliers in the redshift distribution and no further selection criteria than a cut in magnitude, we conclude that magnification effects must be carefully considered in current and future surveys.

In this paper, we assume that lenses and sources form a flux-limited sample. While this assumption may be a realistic one for lens galaxies (e.g. galaxy redshift surveys frequently start from a flux-limited photometric sample), it is less the case for source galaxies. Sources in weak lensing studies have rather complicated selection criteria, not merely based on flux, but also on size and signal-to-noise ratio, for example. Therefore, our quantitative analysis may not apply directly to observational surveys. Besides, source galaxies typically enter a weak lensing catalogue with a weight that characterises the accuracy of the corresponding shear estimate. We ignored any such weighting scheme in our processing, but it may be relevant, since the weight of an object is also expected to depend on magnitude and size, and is thus affected by magnification.

While the relative amplitude of the bias caused by magnification is modest in most cases, and probably smaller than the uncertainties from shape noise and sample variance in previous surveys, future surveys like *Euclid* or LSST have such improved statistical power that magnification effects must be accounted for in the quantitative analysis of GGL.

6.8. Appendix A: Mock data

6.8.1. Appendix A.1: Millennium Simulation data

To study magnification effects in GGL, we make use of ray-tracing results through the MS (Springel et al., 2005), which is an N-body simulation of 2160^3 dark matter particles. Each particle has a mass of $8.6 \times 10^8\, M_\odot$ that is confined to a cube with side length of $500\, h^{-1}$Mpc and with periodic boundary conditions. The underlying cosmology is a flat ΛCDM model with a matter density parameter of $\Omega_\mathrm{m} = 0.25$, a baryon density parameter of $\Omega_\mathrm{b} = 0.045$, a dimensionless Hubble parameter of $h = 0.73$, a tilt of the primordial power spectrum of $n_\mathrm{s} = 1$, and a variance of matter fluctuations on a scale of $8\, h^{-1}$ Mpc extrapolated from a linear power spectrum of $\sigma_8 = 0.9$. This cosmology is based on combined results of 2dFGRS (Colless et al., 2001) and first-year WMAP data (Spergel et al., 2003).

The ray-tracing results are based on a multiple-lens-plane algorithm in 64 light cones constructed from 37 snapshots between redshifts $z = 0$ to $z = 3.06$, each covering a $4 \times 4\,\mathrm{deg}^2$ field of view. For more information about the ray tracing, the reader is kindly referred to Hilbert et al. (2009). An important aspect of the algorithm is that the galaxy-matter correlation is preserved. The ray-tracing results contain the Jacobians $\mathcal{A}$ on a $N_\mathrm{pix} = 4096^2$ pixel grid, which corresponds to a resolution of 3.5 arcsec per pixel. From this, we calculated shear and magnification on a pixel grid.

We also used a catalogue of galaxies based on a semi-analytic galaxy-formation model by Henriques et al. (2015). This catalogue matches the GGL and galaxy-galaxy-galaxy lensing signal from CFHTLenS (Saghiha et al., 2017). The galaxies

are listed for each redshift snapshot with various properties, for example, (magnified) flux in various filters and positions, which allows for selection according to chosen magnitude limits. However, the galaxy positions are not confined to a grid as is the Jacobi information. Thus, we shifted all selected galaxies to their nearest grid point. Therefore, analyses that are close to the centre of the galaxy suffer from discretisation effects on scales comparable to the pixel size.

6.8.2. Appendix A.2: Obtaining a tangential shear estimate

To extract the shear signal averaged over many lenses, a fast Fourier Transform (FFT) is employed. In order to do so, we first defined lens and source number density on a grid by

$$n_{\mathrm{d,s}}(\boldsymbol{\theta}) = \sum_{i=1}^{N_{\mathrm{d,s}}} \delta_{\mathrm{K}}(\boldsymbol{\theta} - \boldsymbol{\theta}_{\mathrm{d,s}}^{(i)}) \;, \tag{6.33}$$

with δ_{K} being one if $\boldsymbol{\theta} = \boldsymbol{\theta}_{\mathrm{d,s},}^{(i)}$ and zero otherwise. The number of lenses and sources is $N_{\mathrm{d,s}}$, and $\boldsymbol{\theta}_{\mathrm{d,s}}^{(i)}$ are the positions of lenses and sources, respectively. Furthermore, we define the shear field of the sources on the grid by

$$\gamma_{\mathrm{s}}(\boldsymbol{\theta}) = \sum_{i=1}^{N_{\mathrm{s}}} \gamma(\boldsymbol{\theta}, z_{\mathrm{s}})\, \delta_{\mathrm{K}}(\boldsymbol{\theta} - \boldsymbol{\theta}_{\mathrm{s}}^{(i)}) \;. \tag{6.34}$$

Then, the tangential shear estimator (6.11) can be expressed as

$$\hat{\gamma}_{\mathrm{t}}(\theta) = -\Re \frac{\sum_{\boldsymbol{\theta}'} \Delta(\theta, |\boldsymbol{\theta}'|)\, \boldsymbol{\theta}'^{*}/\boldsymbol{\theta}' \;\sum_{\boldsymbol{\theta}''} n_{\mathrm{d}}(\boldsymbol{\theta}'')\, \gamma_{\mathrm{s}}(\boldsymbol{\theta}'' + \boldsymbol{\theta}')}{\sum_{\boldsymbol{\theta}'} \Delta(\theta, |\boldsymbol{\theta}'|) \;\sum_{\boldsymbol{\theta}''} n_{\mathrm{d}}(\boldsymbol{\theta}'')\, n_{\mathrm{s}}(\boldsymbol{\theta}'' + \boldsymbol{\theta}')} \;, \tag{6.35}$$

where the sums over $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}''$ extend over the whole grid. The equivalence of Eqs. (6.35) and (6.11) can be verified by inserting the definitions of $n_{\mathrm{d,s}}$ and γ_{s} into the former. The sum over $\boldsymbol{\theta}''$ in the denominator calculates for each θ' the number of lens-source pairs, ensuring that (6.35) is not affected by a potential masking or inhomogeneous survey areas. Further, the sums over $\boldsymbol{\theta}''$ in the numerator and denominator of the GGL estimator (6.35) are convolutions. Thus, the convolution theorem can be applied to these sums. If $\mathcal{F}\{f\}$ is the Fourier transform of a function f and $\mathcal{F}^{-1}$ the inverse Fourier transform, then we can rewrite the estimator (6.35) as

$$\hat{\gamma}_{\mathrm{t}}(\theta) = -\Re \frac{\sum_{\boldsymbol{\theta}'} \Delta(\theta, |\boldsymbol{\theta}'|)\, \boldsymbol{\theta}'^{*}/\boldsymbol{\theta}' \;\mathcal{F}^{-1}\left\{\mathcal{F}\left\{n_{\mathrm{d}}\right\} \mathcal{F}\left\{\gamma_{\mathrm{s}}\right\}\right\}(\boldsymbol{\theta}')}{\sum_{\boldsymbol{\theta}'} \Delta(\theta, |\boldsymbol{\theta}'|) \;\mathcal{F}^{-1}\left\{\mathcal{F}\left\{n_{\mathrm{d}}\right\} \mathcal{F}\left\{n_{\mathrm{s}}\right\}\right\}(\boldsymbol{\theta}')} \;, \tag{6.36}$$

which can be readily solved by an FFT method. For this, we employed routines from the FFTW library by Frigo and Johnson (2012) in our code.

An FFT implicitly assumes periodic boundary conditions, which introduces a bias to the averaged shear data and, thus, must be mitigated for. We can restrict the selection of lenses to the inner $(4^\circ - 2\theta_{\rm out})^2$ of the field, where $\theta_{\rm out}$ is the maximum separation from the lens that we considered. However, for a $\theta_{\rm out} \approx 20'$, we already lose approximately 30% of the lens galaxies. Alternatively, we employed a zero-padding method in which we increased the FFT-area to $(4^\circ + \theta_{\rm out})^2$ and filled the added space with zeros. In this case, we used all available lenses with the cost of slightly increased computational time and the gain of a less-noisy shear profile. For the whole $64 \times 16\,{\rm deg}^2$, our FFT-based code needs a CPU time of 823 s, independent of the number of sources and lenses. We compared the performance of our method to the publicly available `athena` tree-code (Kilbinger et al., 2014). In contrast to the FFT method, the computation time of `athena` is enhanced with the number of lens-source pairs. We adjusted the settings to our survey parameters while leaving the parameter that sets the accuracy of the tree-code, meaning the open-angle threshold, to its pre-set value. For 10^7 sources and lenses at $z_{\rm d} = 0.41$ with limiting magnitudes 19.5, 22, and 29 mag, `athena` performs with a CPU time of 889 s, 1167 s, and 2043 s, respectively.

6.8.3. Appendix A.3: Estimating the impact of the magnification bias on a KiDS+VIKING+GAMA-like survey

We considered GAMA-like lens galaxies with narrow redshift distributions at $z_{\rm d} = 0.21$ and $z_{\rm d} = 0.36$ with a flux limit of 19.8 mag. Using simulated data from the semi-analytic galaxy-formation model by Henriques et al. (2015), we obtain local slopes of $\alpha_{\rm d} = 0.85$ and $\alpha_{\rm d} = 2.11$, respectively. To keep the statistical error low, we still consider the data from the whole simulated area of $64 \times 16\,{\rm deg}^2$. To mimic the KiDS+VIKING-like source population, we matched the last three bins of the best-estimated redshift distribution to Millennium data from $0.51 \leq z_{\rm s,Mil} < 1.28$. We then calculated a weighted shear map from the simulated data, the final source galaxy distribution has a local slope of $\alpha_{\rm s} = 0.51$ for a limiting magnitude of 25 mag. We repeated our analysis from Sects. 6.4, 6.5, and 6.6, and obtain the results listed in Table 6.4. The results are quantitatively comparable to the results from Figs. 6.4 and 6.6.

A KiDS+VIKING+GAMA-like survey is moderately affected. We stress that this estimate has been made with simplified assumptions, for example, there are no catastrophic outliers in the redshift distribution and no selection criteria for source galaxies other than a magnitude cut.

Table 6.4. – Impact of the magnification bias on a KiDS+VIKING+GAMA-like survey.

z_{d}	0.21	0.36
α_{d}	0.85	2.11
α_{s}	0.51	0.51
$\delta_\gamma{}^{\mathrm{d}}$	-0.09%	1.67%
$\delta_\gamma{}^{\mathrm{s}}$	-0.88%	-0.94%
δ_M	-2.84%	-8.26%

Notes. The superscript 'd' indicates the magnification bias caused only by the lens galaxies, while 's' stands for the source galaxies.

CHAPTER 7

BARYON ACOUSTIC OSCILLATIONS IN GALAXY-GALAXY LENSING

BAO are frozen-in density perturbations that are imprinted on the CMB and today's three-dimensional matter distribution. They are detected in the LSS of galaxy distributions using measurements of correlation functions as well as power spectra. The BAO scale serves as a standard ruler, and thus, enables us to measure the expansion history of the Universe. For this reason, the Dark Energy Task Force (Albrecht et al., 2006) acknowledged BAO measurements as one of the four most promising ways to unveil the nature of dark energy. As of yet, BAO have not been detected with weak lensing methods. The aim of this chapter is to use weak lensing simulations to constrain BAO in GGL.

7.1. State of the art

In the pre-recombination era, photons were tightly coupled to electrons, while the electrons were coupled to the baryons by Coulomb interactions. They formed a baryon-photon fluid that propagated through the early Universe with sound speed c_s. The photon pressure balanced the in-fall of baryons into the potential wells of the dark matter, forming acoustic oscillations (for a review see, e.g., Bassett and Hlozek, 2010, and references therein). In the era of recombination $z_\mathrm{rec} \approx 1090$, the photons decouple from baryons and diffuse away. Thus, the propagation of the sound waves comes to a halt shortly after at the so-called baryon drag epoch $z_\mathrm{drag} \approx 1020$. While the majority of the dragged baryons fall back into the potential well of the dark matter distribution, a part of the dark matter particles follows the baryon density field. Since the subsequent formation of galaxies follows the underlying matter

density field, the imprint of this density perturbation is still visible in the late-time LSS.

The time a sound wave can travel is from the end of inflation to z_{drag}. Thus, the physical size of the sound horizon can be calculated as

$$r_{\mathrm{drag}} = \int_{z_{\mathrm{drag}}}^{\infty} \mathrm{d}z \, \frac{c_{\mathrm{s}}(z)}{H(z)} \, . \tag{7.1}$$

Using CMB observations, r_{drag} is obtained within 0.2% uncertainty (*Planck* Collaboration et al., 2018a). Currently, the best estimate for r_{drag} is $(147.38 \pm 0.25)\,\mathrm{Mpc}$ comoving which is known as the standard ruler of the Universe. By either measuring the angular BAO scale or its line-of-sight distance in the LSS of the late-time Universe, we infer cosmological parameters from the distance-redshift relation (see Eq. 2.17, and for a practical application, e.g., Ross et al., 2017).

The first detection of BAO was made by observing the galaxy distribution; Cole et al. (2005) studied the BAO imprint on the power spectrum within the 2dFGRS and Eisenstein et al. (2005) used galaxy data from SDSS I and II to detect BAO in the correlation function. The most recent results estimate the BAO scale with percent-level accuracy (Alam et al., 2017; Ata et al., 2018, with BOSS galaxies and eBOSS quasars, respectively). Galaxy surveys yield cosmological constraints up to a redshift of approximately two; an estimate at $z \approx 2.3$ has recently been done by Bautista et al. (2017) using the Lyman α forest measured in BOSS quasars. Planned spectroscopic surveys will span an even wider and deeper survey volume and improve the measurement accuracy by an order of magnitude. Upcoming surveys are performed by the Dark Energy Spectroscopic Instrument (DESI, starting the survey in summer 2020, `www.desi.lbl.gov`), the 4-metre Multi-Object Spectroscopic Telescope (4MOST, `www.4most.eu`), the Prime Focus Spectrograph (PFS, `pfs.ipmu.jp`) and the space-based *Euclid* satellite.

To eventually reach sub-percent accuracy on the angular BAO scale, we have to control systematic uncertainties extremely well. With the characteristic scale being about $\approx 100\,h^{-1}\mathrm{Mpc}$, we are in the linear regime of structure growth and it is comparably easy to construct accurate models. On the downside, we only have a low number of modes available at these large scales. Therefore, the impact of sample variance or ultimately cosmic variance on the error budget must be carefully taken account since noise fluctuations affect the shape of the BAO peak (Ruggeri and Blake, 2019). So-called redshift-space distortions (RSD) are a major source of error for the BAO scale along the line-of-sight. The RSD are caused by the peculiar motion of galaxies v_{pec} which slightly alters the observed redshift, while the projected distance on the sky remains unaffected. If $v_{\mathrm{pec},z}$ is the peculiar velocity component that is aligned with the line-of-sight, we can write the three-dimensional

redshift-space coordinates $\boldsymbol{s}$ approximately as

$$\begin{aligned} s_{x,y} &= r_{x,y} = \frac{z\,c}{H(z)}\,\theta_{x,y}\ , \\ s_z &= r_z + \frac{v_{\mathrm{pec},z}}{H(z)}\ , \end{aligned} \tag{7.2}$$

where $\boldsymbol{r}$ are the proper coordinates. The RSD cause the observed correlation function to be slightly anisotropic, which subsequently shifts the observed BAO scale.

BAO are imprinted on the matter distribution of the Universe. We use, however, galaxies as tracers of the distribution who are known to be biased. Especially the brightest galaxies, e.g., the luminous red galaxies (LRG) and quasars, are highly biased tracers but can be observed out to high redshifts. The scale-independent linear bias factor b (2.51) accounts for the difference of matter tracers and matter distribution to leading order. Nevertheless, to infer cosmological parameters accurately, we have to measure the scale-dependency of the galaxy bias precisely, which is known to be notoriously difficult. Although structure growth on large scales is fairly well described with linear perturbation theory, the desired percent or even sub-percent level accuracy warrants attention to non-linear effects as any scale-dependent bias induces a shift in the BAO peak position and changes the shape of the peak. Biases that are known to affect small scales are the satellite distribution bias, the galaxy assembly bias and the velocity bias for central and satellite galaxies. The satellite distribution bias refers to a non-Poissonian distribution of satellite galaxies, and the assembly bias accounts for differences in the formation of galaxies for equal halo masses due to environmental effects. Lastly, the velocity bias accounts for differences in the distribution of galaxy velocities compared to that of dark matter (sub)halos. They are also potential candidates to affect large scales, which is currently investigated (see, e.g., Duan and Eisenstein, 2019, and references therein).

Another potential source of such a scale-dependent bias that is not captured by linear theory is proposed by Tseliakhovich and Hirata (2010) and discussed more recently by Blazek et al. (2016). Shortly after recombination, the baryon sound waves moved supersonically compared to the velocity field of the dark matter particles. The matter particles fell into the potential wells, but the supersonic streaming velocity of the baryons led to suppression in the galaxy formation rate due to a less effective gas accretion and cooling rate. Thus, the galaxy distribution is not necessarily matched to the dark matter halo distribution by a linear factor, even on large scales. This potentially introduces a relative offset of the characteristic BAO scale between these distributions.

Currently, a method called `density field reconstruction` is used to alleviate issues related to non-linear structure growth and RSD (Eisenstein et al., 2007) which has been proven to obtain the desired accuracy until now (Duan and Eisenstein, 2019). It has therefore been integrated into the standard analyses routine (for a recent example of handling observational systematics see, e.g., Ross et al., 2017).

By using Lagrangian perturbation theory, the reconstruction method aims to reduce anisotropies in the clustering and takes gravitational bulk flows on large-scales into account. Recently, many approaches have been made to improve the standard reconstruction method (see, e.g., Schmittfull et al., 2017; Seljak et al., 2017; Hada and Eisenstein, 2018; Shi et al., 2018; Zhu et al., 2018; Wang and Pen, 2019).

A complementary approach is to use different tracers for the matter structures to measure the BAO scale. The inference of the BAO scale from the Lyman α forest has already been obtained, and potential future candidates for detecting BAO are intensity mappings with the 21 cm line of hydrogen (Chang et al., 2008) and cosmic voids which refer to three-dimensional troughs in the density field (Zhao et al., 2020). A multi-tracer analysis reduces the impact of systematic errors and thus, yields tighter constraints on cosmological parameters (Abramo and Leonard, 2013). Besides, higher-order n-point statistics also have the advantage to capture non-Gaussian information in the LSS statistics. Recent results were obtained by de Carvalho et al. (2020) who used the three-point galaxy correlation function to measure the BAO scale. In 1983, Davis and Peebles proposed to perform the BAO analysis in a projected space to alleviate RSD and use a model-independent approach that does not rely on a fiducial cosmological model to calculate the line-of-sight distance. In this case, the true distance $\sqrt{r_\mathrm{p}^2 + l^2}$ between objects is integrated over the line-of-sight distance l, while the projected distance r_p is measured as the angular scale. McBride et al. (2011) performed such an analysis for three-point correlation functions which showed promising results to improve the control over systematic uncertainties.

To circumvent the galaxy bias entirely, we need to obtain the BAO scale directly from the matter distribution. This will be achieved by a successful detection of BAO in cosmic shear, which is a two-point correlation function of lensing shear as a function of angular scale. Cosmic shear is known to be sensitive to the geometry of space and the structure growth history; it measures the strength of gravitational clustering in the projected matter distribution. However, the lensing kernel along the line-of-sight is broad, which smears out the imprint of BAO on the correlation function. Therefore, galaxies with available shape information need to be finely binned in redshift. For most galaxies, however, only photo-z estimates are available which show large uncertainties compared to spec-z estimates. Further, due to the presence of shape noise, a large number of galaxies with accurate shape and redshift information is needed. For these reasons, a detection of BAO in cosmic shear is challenging even for future surveys. Grassi and Schäfer (2014) calculated the detection significance for future tomographic cosmic shear surveys that bin galaxies using spin-spherical harmonics and spherical Bessel functions (Heavens, 2003). Deciding parameters include shape noise, photo-z errors, median redshifts, the average number of galaxies per area, and the sky coverage. They conclude that LSST as well as *Euclid* will be able to detect the first three and four BAO wiggles in the power spectrum with high significance, respectively. This result has been confirmed

and further quantified in Bernardeau et al. (2020). Recently, Ding et al. (2019) proposed to only use galaxies with available spec-z information, which is referred to as kinematic weak lensing. They forecast a 3σ detection of the BAO scale in the power spectrum with a precision of 4% for a hypothetical stage IV-like kinematic weak lensing survey.

The aim of BAO surveys is a statistically significant, unbiased constraint on the peak position with high precision. Any improvement in the precision level using multiple tracers or detection techniques, as summarised above, is as effective as further enlarging the survey area. Accuracy, however, is only improved by comparing results from different tracers and techniques. In general, a joint analysis is not only advantageous but also necessary.

To complete the picture of measuring the BAO scale in the galaxy distribution as well as shear-shear correlation functions, we can analyse the galaxy-shear cross-correlation which is known as the GGL signal. The dependence of the galaxy-matter correlation function on the galaxy bias is linear, whereas the galaxy-galaxy correlation function depends quadratically on the bias. Also, RSD and detection systematics affect GGL estimates differently. Finally, GGL estimates can be performed with the combined data from cosmic shear and spectroscopic surveys. Therefore, a complementary detection of the BAO scale in GGL can only improve future measurements with no additional observing time. Jeong et al. (2009) considered an LSST-like survey and three fiducial lens samples at $z_\mathrm{d} = 0.3, 0.5$, and 0.8 with one million lenses each; they concluded with an analytical model that χ^2-difference of a model with and without BAO is $\Delta\chi^2 = 3.2$. The covariance matrix is dominated by cosmic variance and is further influenced by the number of lenses per redshift slice, shape noise, and the number density of background galaxies. By adding more redshift slices, the S/N increases roughly by the square root of the lens number count, $\sqrt{N_\mathrm{d}}$.

Aside from the analytical ascertainment that a BAO detection in GGL is possible but challenging, no further research has been made (to my best knowledge). Therefore, I attempt in the following to use artificial galaxy catalogues and ray-tracing results to obtain a forecast for detecting the BAO signal in GGL. A summary of this still ongoing project is given in the following sections.

7.2. Methods

We use ray-tracing results from the MXXL and the semi-analytical galaxy formation model (SAM) from Henriques et al. (2015) as well as artificial lens and shear catalogues from SLICS, which are explained in Sect. 4.1 in detail. The tangential shear signal is estimated as given in Eq. (4.20), which is calculated by the FFT method explained in Sect. 4.3. To alleviate the impact from the FFT's periodic boundary conditions, we employ a zero-padding method. For lens redshifts at $z_\mathrm{d} = 0.4$, 1, and

1.4, the BAO signal is expected on angular scales around 3.7°, 1.7°, and 1.4°, respectively. We calculate the shear signal in 30 linearly spaced bins from $R_{\rm in} = 50\,h^{-1}$Mpc to $R_{\rm out} = 200\,h^{-1}$Mpc in comoving coordinates, and also in angular coordinates with 40 bins for the whole simulated field, with $\theta_{\rm out} = 12°$ for the MXXL and $\theta_{\rm out} = 10°$ for the SLICS. The lens galaxy sample for both simulations is restricted to bright, elliptical galaxies with several limiting magnitudes in the r-band. We project their position using the NGP method onto a grid with $N_{\rm pix} = 4096^2$.

The shear catalogues for each $12 \times 12\,\mathrm{deg}^2$-field of the MXXL are set up identical to the MS with its $4 \times 4\,\mathrm{deg}^2$-fields. Thus, we obtain shear estimates and their errors in the same way as explained in Sect. 6.8.2, where we explicitly set $n_{\rm s}(\boldsymbol{\theta})$ to unity for the whole field. The total sky coverage of the MXXL is $2304\,\mathrm{deg}^2$, which is equivalent to 6% of the whole sky and 15% of the Euclid footprint.

The SLICS provide catalogues per line-of-sight with galaxies and their properties of all redshifts. This applies to lens as well as source galaxy catalogues. Thus, we have to make two adjustments compared to MXXL. Firstly, shear information is not available for the whole simulated grid, but only for the source galaxies with positions that need to be projected onto the grid. Hence, we treat the source galaxies the same as the lens galaxies, with the difference that we average shear values per pixel instead of stacking the number of lenses per pixel. Secondly, we use all available sources with redshifts $z_{\rm s} > z_{\rm d}$, such that we obtain a comparable S/N as for the MXXL. The critical surface mass density $\Sigma_{\rm crit}$ is calculated using the effective source redshift $z_{\rm s,eff}$ of the source galaxies' redshift distribution $p(z_{\rm s})$. In this case, the lensing efficiency becomes

$$\frac{D_{\rm ds}}{D_{\rm s}} := \frac{D(z_{\rm d}, z_{\rm s,eff})}{D(z_{\rm s,eff})} = \int_{z_{\rm d}}^{\infty} \mathrm{d}z_{\rm s}\, p(z_{\rm s}) \frac{D(z_{\rm d}, z_{\rm s})}{D(z_{\rm s})}\,, \tag{7.3}$$

where $p(z_{\rm s})$ is normalised to unity. From the available mock data, we use for the source galaxies the deep LSST-like catalogues with mean redshifts $\bar{z}_{\rm s} = 1.0$ and 1.8 which correspond to the sixth and ninth tomographic bin, respectively. For the lens galaxies, we use the GAMA-like catalogues. For this combination, we have 632 lines-of-sight with each $10 \times 10\,\mathrm{deg}^2$ at our disposal. Thus, the total sky coverage is $63\,200\,\mathrm{deg}^2$ or 1.5-times the full sky.

7.3. Results

In the following, the expected BAO feature in GGL estimates is analytically explored. Tangential shear and convergence signals are then compared to numerical results from the MXXL and SLICS. The SLICS cover a huge simulated area while providing only relatively low lens redshifts. On the other hand, catalogues from the MXXL provide very deep lens galaxy information with a substantially lower area. Unfortunately, both simulations cannot detect BAO in GGL as is shown in this section. The discussion of the results follows in the Sect. 7.4.

7.3.1. Analytical estimate

We start by predicting the expected signal analytically with simplified assumptions. Linear theory is used and we ignore observational systematics as discussed in the previous section, in particular, RSD, photo-z errors, and shape noise. The linear matter power spectrum (2.42) is calculated using the Python package `COLOSSUS` (Diemer, 2018) which implements a model for transfer function $T(k)$ from Eisenstein and Hu (1998) with or without baryons as a function cosmological parameters. Fig. 7.1a shows the corresponding spatial linear matter-matter correlation function $\xi_{\rm 3D}$, as a function of comoving distance R and of angular scale θ. The cosmology is set to the one from MXXL and the redshift for all panels is $z_{\rm BAO} = 0.4$. At a scale of roughly $110\,h^{-1}$Mpc, the BAO imprint is clearly seen. The projected, two-dimensional correlation function is calculated as

$$\xi_{\rm 2D}(r_{\rm p}; z) = \int_{-\Delta l}^{\Delta l} {\rm d}l\; \xi_{\rm 3D}\left[\sqrt{r_{\rm p}^2 + l^2}; z(l)\right] \; . \tag{7.4}$$

Using a slice of thickness $2\Delta l$ centred on the redshift $z_{\rm BAO}$, we approximate $z(l) \approx z_{\rm BAO}$ and show the result in Fig. 7.1b. The projected correlation function is negative for separations $R \gtrsim 150\,h^{-1}$Mpc comoving and the BAO feature is smeared out compared to the signal in $\xi_{\rm 3D}$. For a fixed redshift, the smearing is stronger for thicker slices Δl. The ray-tracing algorithms rely on the multiple-lens-plane approximation as was described in Sect. 4.1. The bin width for projecting the matter onto two-dimensional mass sheets is $\Delta l = 50\,h^{-1}$Mpc for the MXXL and $\Delta l = 125\,h^{-1}$Mpc for SLICS. Both cases are visualised in Fig. 7.1b.

The two lower panels of Fig. 7.1 show the tangential shear $\gamma_{\rm t}$ (panel c) and the convergence κ (panel d). They are calculated as

$$\gamma_{\rm t}(R) = \frac{b\,\rho_0}{\Sigma_{\rm crit}(z_{\rm d}, z_{\rm s})} \int \frac{{\rm d}k}{2\pi}\; k\,P(k, z_{\rm d})\,J_2(kR) \; , \tag{7.5}$$

$$\kappa(R) = \frac{b\,\rho_0}{\Sigma_{\rm crit}(z_{\rm d}, z_{\rm s})} \int \frac{{\rm d}k}{2\pi}\; k\,P(k, z_{\rm d})\,J_0(kR) \; , \tag{7.6}$$

where $J_{0,2}$ are the zeroth and second Bessel function of the first kind. The integral is solved using a Hankel transform method introduced by Ogata (2005). We use a source redshift of $z_{\rm s} = 1$, a lens redshift $z_{\rm d} = z_{\rm BAO}$, as well as an absent linear galaxy bias, i.e. $b = 1$. The BAO feature in the tangential shear signal has been dubbed 'shoulder' in the literature and shows the least relative distinction compared to a zero-baryon model. The convergence in the fourth panel is, like the projected correlation function, not necessarily a positive function but nevertheless approaches zero in the limit of infinite separation.

7.3.2. Numerical estimates

We calculate the GGL signal from the SLICS simulation for their highest GAMA-like lens redshifts. The lens galaxy type is restricted to elliptical galaxies with

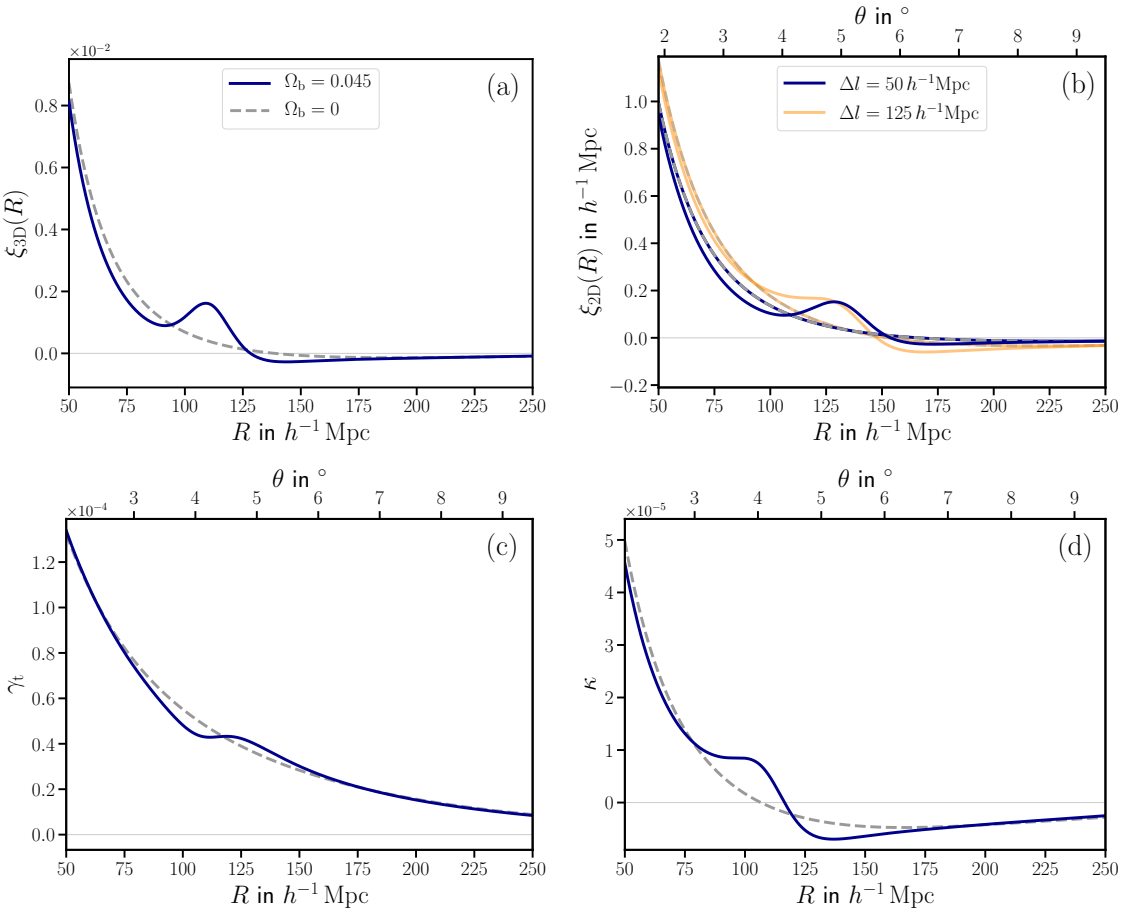


Figure 7.1. – The four panels show various correlation functions as a function of comoving separation and angular scale. A linear perturbation theory model that includes baryons is shown in solid lines, while dashed lines show correlation functions in the absence of baryons. The cosmology for all panels is the one from the MXXL. In the panel a, the spatial correlation function $\xi_{3D}(R = \sqrt{r_p^2 + l^2})$ is shown for a redshift of $z_{BAO} = 0.4$. Projecting the spatial correlation function in a slice of $2\Delta l$ to two dimensions, $\xi_{2D}(R = r_p)$, results in the panel b, where the blue curve is matched to ray tracing through the MXXL and the orange curve to SLICS. The lower panels show the tangential shear γ_t (panel c) and the convergence κ (panel d) for an unbiased lens sample at $z_{BAO} = 0.4$ and $z_s = 1$ as a function of projected distance.

magnitudes $m_\mathrm{d} < 19\,\mathrm{mag}$ and $m_\mathrm{d} < 20\,\mathrm{mag}$, we chose two redshift bins for each limiting magnitude with $0.4 < z_\mathrm{d} \leq 0.45$ and $0.45 < z_\mathrm{d} \leq 0.5$. The angular BAO scale changes over the redshift bin by $\Delta\theta = 0.36°$ and $\Delta\theta = 0.29°$, respectively. In Fig. 7.2, we show the results of the four different combinations, where the tangential shear γ_t is depicted as a function of comoving distance R and angular scale θ. We

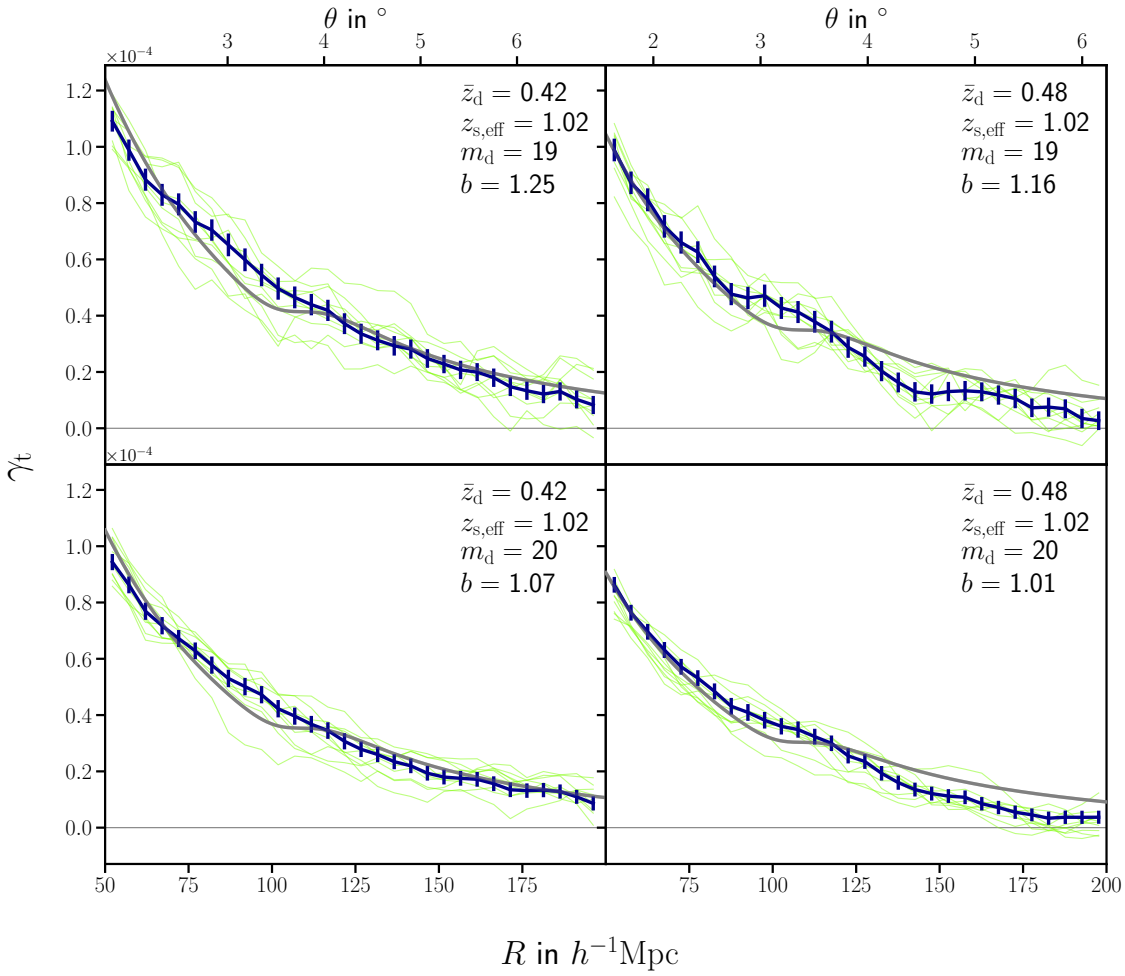


Figure 7.2. – Tangential shear signals γ_t estimated from bright, elliptical lenses in a thin redshift bin from SLICS are shown. The solid, blue lines indicate the shear estimates with Jackknife errors. Each green line only uses randomly selected $15\,000\,\mathrm{deg}^2$ of simulated data for the tangential shear estimate, which is equivalent to a Euclid-like sky coverage. The grey lines show the best-fitting analytical model, where the galaxy bias b is fitted, and its results are indicated in the legend. In all four panels the fitting error is $\Delta b < 0.03$. The legend also indicates the mean lens redshifts $\bar{z}_\mathrm{d}$, the effective source redshifts $z_\mathrm{s,eff}$, and the lenses' limiting magnitudes m_d. All curves are shown as a function of comoving redshift and angular scale.

further list the number of lenses and sources in Table 7.1. The mean redshift $\bar{z}_\mathrm{d}$ is calculated by averaging the redshifts of each galaxy that fit the criteria; and the effective source redshift is calculated according to Eq. (7.3). We obtain the

Table 7.1. – We list the mean number of lenses and sources per line-of-sight that are used from SLICS in Fig. 7.2, and from MXXL in Fig. 7.3. The total number of lines-of-sight for SLICS is 632 and for MXXL 16. For both simulations, the number of lenses is given as a function of limiting magnitude in the r-band and redshift. While the lens redshifts are similar by design, their limiting magnitudes are chosen such that the resulting S/N of the tangential shear signal is roughly comparable.

Fig. 7.2 with SLICS			Fig. 7.3 with MXXL		
$\bar{N}_\mathrm{d}$			$\bar{N}_\mathrm{d}$		
	$\bar{z}_\mathrm{d} = 0.42$	$\bar{z}_\mathrm{d} = 0.48$		$\bar{z}_\mathrm{d} = 0.41$	$\bar{z}_\mathrm{d} = 0.47$
$m_\mathrm{d} = 19$	1256	819	$m_\mathrm{d} = 21$	11 834	7783
$m_\mathrm{d} = 20$	5077	3954	$m_\mathrm{d} = 22$	37 059	33 876
$\bar{N}_\mathrm{s}$	936 005		$\bar{N}_\mathrm{s}$	16 777 216	

galaxy bias b by fitting the analytical model to the data with a piecewise third-order polynomial interpolation, commonly referred to as a cubic spline interpolation (Press et al., 1992a). The error of the galaxy bias is < 0.03 for all cases shown in this section. Lenses selected with a lower magnitude limit have a higher galaxy bias, as expected. However, lenses at a higher mean redshift show a slightly lower bias which is not understood. The solid, blue lines indicate the simulated data with Jackknife errors, and the best fitting analytical estimate is shown in grey. Moreover, we plot in green the result of randomly drawing ten times 150 lines-of-sight which is equivalent to a Euclid-like sky coverage. Although, data and simplified theory show overall a good match, no BAO feature is visible 'by eye' in either of the panels. Furthermore, the error bars indicate that only one redshift bin is not sufficient to constrain the signal even in the absence of shape noise. For a Euclid-like survey, this problem is even more severe.

Before averaging the signal over the two redshift bins, we show comparable results for the MXXL in Fig. 7.3. We again choose elliptical galaxies and match the lens and source redshifts to the previous case. Due to the smaller simulated area and less width of the redshift bin, we chose fainter galaxies with magnitudes $m_\mathrm{d} < 21\,\mathrm{mag}$ and $m_\mathrm{d} < 22\,\mathrm{mag}$ to achieve comparable S/N ratios. The explicit numbers of lenses and sources are listed in Table 7.1. Again, the tangential shear is shown as a function of comoving distance and angular scale. The data are shown as a solid, blue line and the theory in grey. Light-blue curves depict the average curve when leaving out one line-of-sight, which depicts the variance in the data using a Jackknife approach. The fitted galaxy bias is higher for brighter lens galaxies and those residing at a higher redshift. As before, no BAO are easily distinguished in the MXXL data curves and the relative difference of the BAO feature is smaller than the variance of the data. Repeating the plotting procedure but with the convergence instead of the tangential shear leads to very similar results in terms of noisiness.

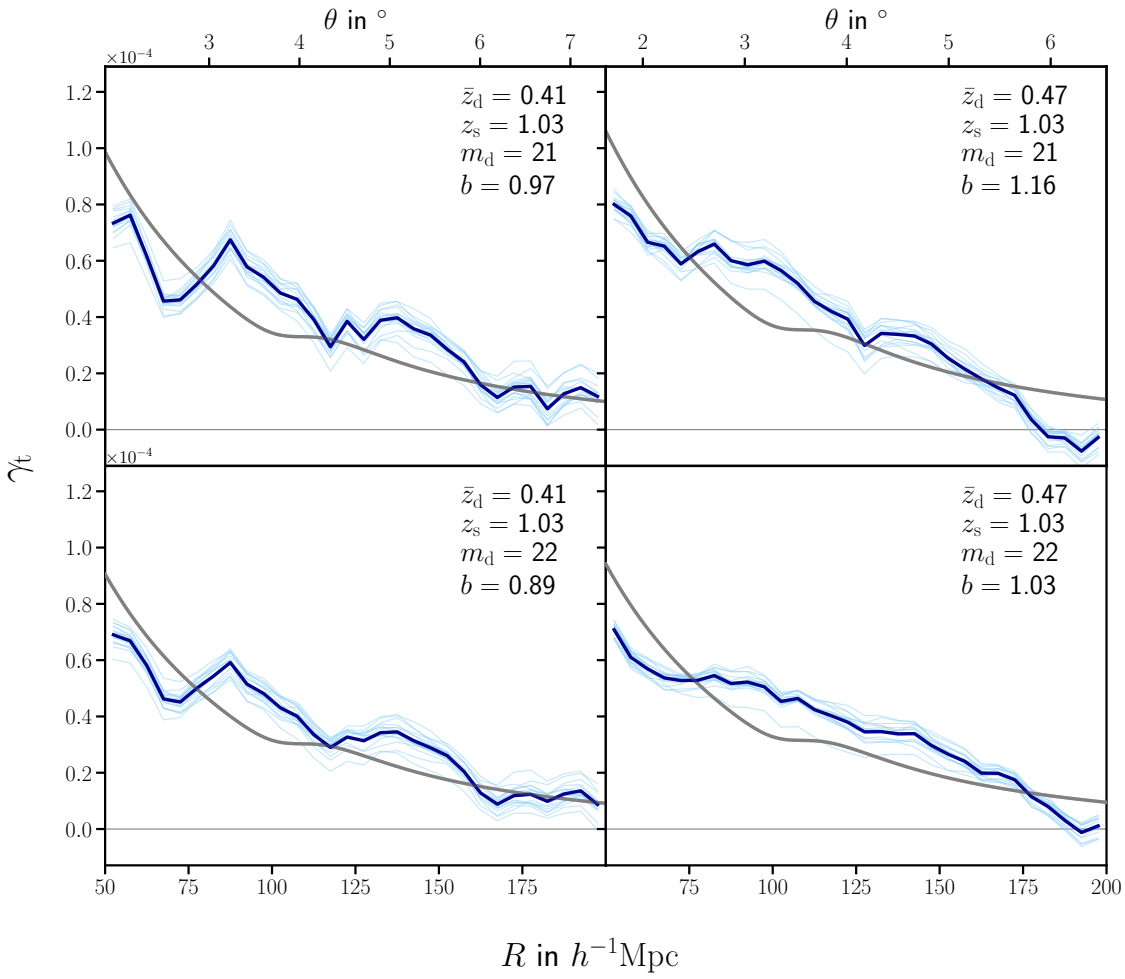


Figure 7.3. – Tangential shear signals γ_t estimated from bright, elliptical lenses in a thin redshift bin from MXXL are shown. Lines and legends are set up as in Fig. 7.2. Additionally, the light-blue lines show the average signal of 15 from the 16 lines-of-sight, thus, indicating the Jackknife variance. The source redshift is fixed to $z_s = 1.03$.

Therefore, we proceed to average the data over redshift bins. We calculate a weighted average with weights N_d per each lens magnitude limit. For SLICS, the result is presented in Fig. 7.4. However, no BAO feature can be detected. We

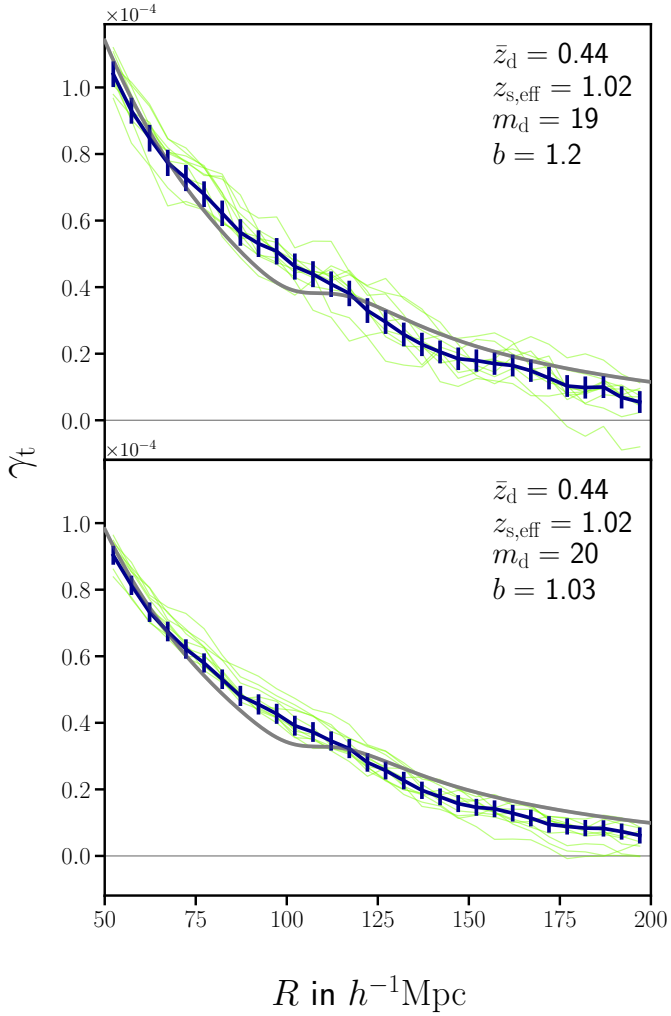
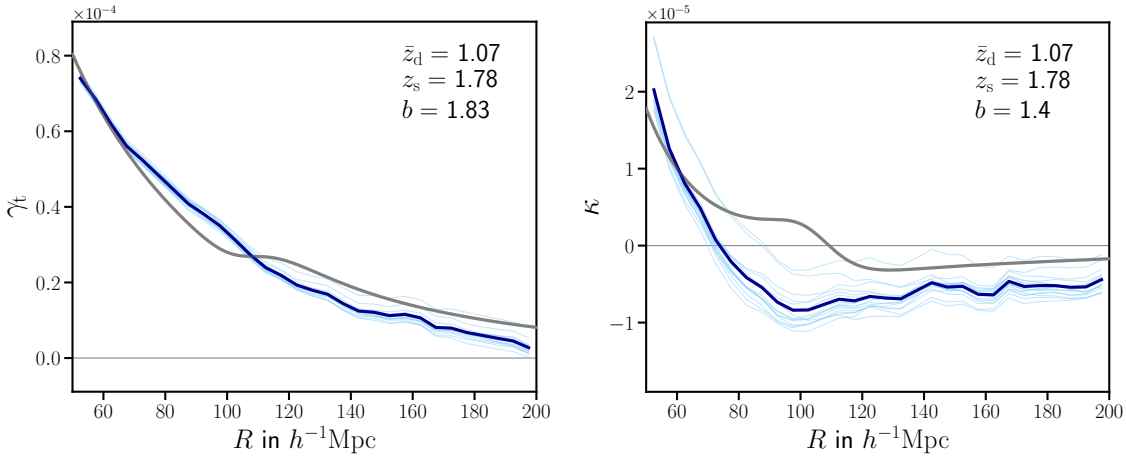


Figure 7.4. – Tangential shear signals γ_t estimated from bright, elliptical lenses in two stacked redshift bins from SLICS are shown. Lines and legends in this figure are set up as in Fig. 7.2. The data curve is a weighted average of the lensing signals presented in Fig. 7.2 per each limiting magnitude.

repeat the analysis for the MXXL using tangential shear and convergence, where we average over four bins in redshift with $z_\mathrm{d} = 0.50$, 0.47, 0.44, and 0.42 as well as constant source redshift. However, the reduction in the scatter of the data is also not sufficient to exhibit the BAO feature and we omit to show the results here.

Instead, we use 38 lens redshift bins in the range $0.24 < z_\mathrm{d} \leq 1.5$ and change the magnitude limit in each redshift bin to only obtain the brightest objects per redshift. This is achieved by analysing the cumulative number density as a function

of magnitude. We choose the limiting magnitude such that it corresponds to the magnitude where the Schechter-like function changes its exponential behaviour to a linear one. The exact choice of the limiting magnitude is not crucial as long as bright and therefore highly-biased galaxies are chosen. We fix the source redshift to $z_s = 1.78$ for all lenses. The results for the tangential shear and the convergence are depicted in Fig. 7.5. However, also, in this case, no BAO feature is visible. Surpris-



(a) tangential shear signal γ_t **(b)** convergence signal κ

Figure 7.5. – Lensing signals estimated from 6 698 255 bright, elliptical lenses in stacked redshift bins from MXXL are shown. Lines and legends are set up exactly as in Fig. 7.3. The lenses are binned in 38 redshift bins from $0.24 < z_d \leq 1.5$ with a fixed source redshift $z_s = 1.78$. The magnitude limit m_d is chosen fainter with increasing lens redshift (bin).

ingly, the data even suggest the apparent absence of BAO features since now the variance of the data is sufficiently low, especially for the tangential shear data in Fig. 7.5a. A theoretical model with zero baryons fits the data slightly better. The zero-baryon model fit leaves the bias factor almost unchanged, which still overestimates the signal on large scales. The convergence curve in Fig. 7.5b also lacks any hint for BAO. A zero-baryon model fit yields a similar galaxy bias factor while still underpredicting small scales and overpredicting large scales.

The tangential shear as well as the convergence are calculated by subtracting the signal around random points as has been described Sect. 4.2. To investigate the form of the signal shown in Fig. 7.2 to 7.5, we show the lensing signals around random points for tangential shear $\gamma_{t,rand}$ and convergence κ_{rand} estimates in Fig. 7.6 as a function of angular scale for all the source redshifts from the previous figures. For consistency, we show error bars obtained by a Jackknife method for the SLICS data and leave-one-out curves for MXXL data, although they depict the same statistical property. The $\gamma_{t,rand}$ and κ_{rand} estimates are independent of lens redshift. Both tangential shear curves show larger variance for larger angular scales, and at around 4° the curves exhibit a positive signal. The convergence shows a variance that is

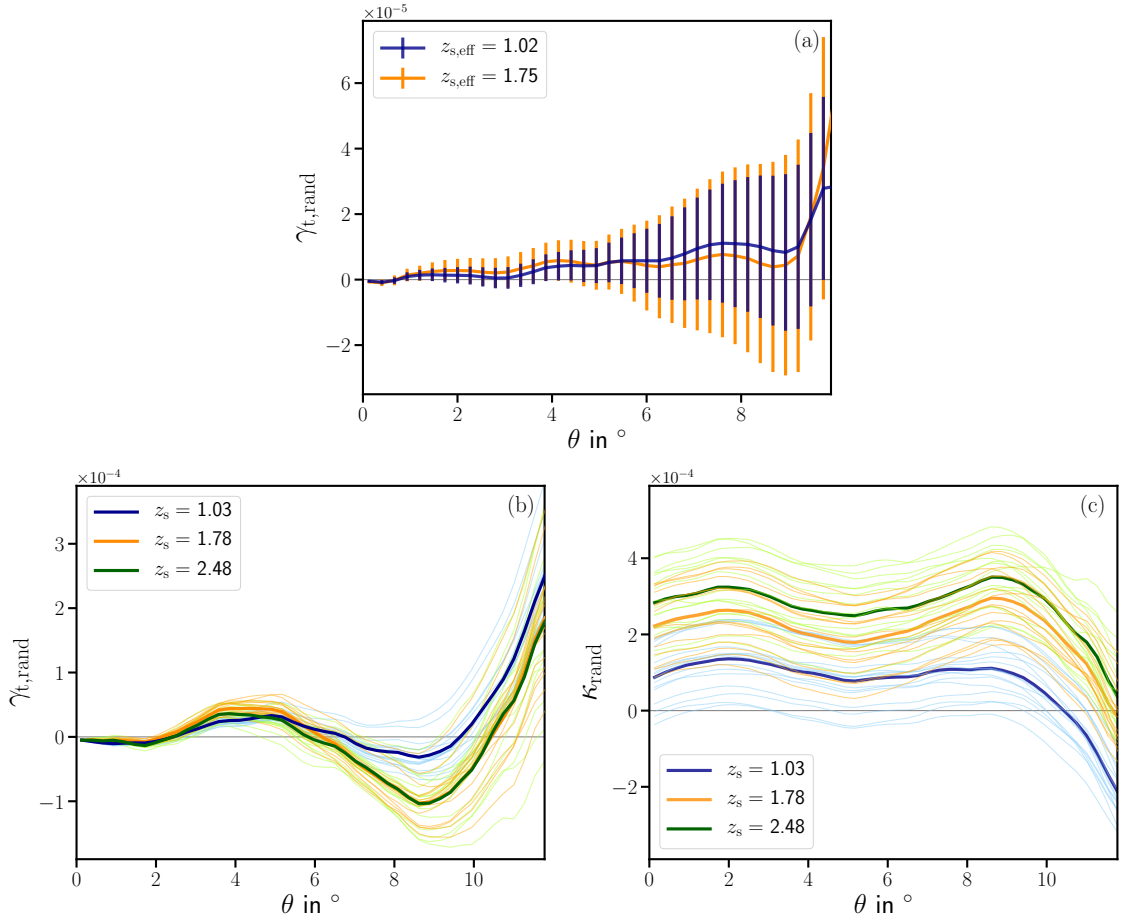


Figure 7.6. – Instead of the location of lens galaxies, random positions are used and the lensing signals are estimated. In panel a results from SLICS are shown. For two effective source redshifts, the tangential shears with Jackknife errors are depicted. The blue curve is the signal around random positions that is subtracted to obtain the Figs. 7.2 and 7.4. In panel b and c results from the MXXL for three different source redshifts are shown; panel b shows the tangential shear signal around random positions and in panel c the convergence. The thin lines indicate the Jackknife variance for each source redshift as in Fig. 7.3 and 7.5. The blue line in panel b shows the signal that is subtracted from the data in Fig. 7.3, the orange lines in panel b and c are subtracted from the left and right panel of Fig. 7.5, respectively, and the green line shows an example for an even higher source redshift. The average values of the convergence fields from panel c are $(0.87 \pm 2.00) \times 10^{-4}$ for $z_s = 1.03$, $(2.08 \pm 2.30) \times 10^{-4}$ for $z_s = 1.78$, and $(2.57 \pm 2.29) \times 10^{-4}$ for $z_s = 2.48$.

almost independent of angular scale; its amplitude has a sharp decrease in amplitude at roughly 9° independent of redshift.

7.4. Discussion & outlook

In the previous section, we showed results for measuring the galaxy-matter cross-correlations for a few selected lens redshifts and limiting magnitudes from simulated data. We aimed for the detection of BAO in mock data, its comparison to the analytical model, and ultimately a forecast for a stage IV survey. The FFT-based code enables us to analyse large quantities of data fast and accurately. Unfortunately, neither the MXXL nor the SLICS exhibit the BAO signature. Results illustrated in Figs. 7.2 to 7.5 clearly indicate the absence of BAO in the ray-tracing simulations. Possible reasons for this are listed in the following. Beyond that, we highlight necessary improvements for the method towards a robust forecast for a stage-IV-like survey.

Choosing GGL estimators, i.e. the tangential shear and the convergence, to detect the BAO signal is certainly valid as was motivated in Sect. 7.1 and illustrated in Fig. 7.1. Besides, the N-body simulations used in this thesis include BAO in their initial power spectra. In general, the size of such a simulation must be considerably larger than the scale of the signal we aim to detect. A periodic box of size $500\,h^{-1}$Mpc is as large as the SLICS cube and six times smaller than the one from MXXL. Such a box retains the BAO feature to redshift zero only with a large scatter around the mean signal due to the small number of available modes which was shown in Springel et al. (2005) in their Fig. 6 for the MS. The MXXL uses the same cosmology as the MS, and Harnois-Déraps and van Waerbeke (2015) report that they base their initial power spectrum for SLICS on `CAMB`[1] (Lewis et al., 2000) with $\Omega_\mathrm{b} > 0$.

Next, we shift our attention to the ray-tracing simulations. The multiple-lens-plane approximation projects matter in thin redshift slices onto a series of two-dimensional mass sheets. This leads to smearing the BAO peak as a function of slice thickness (cmp., e.g., Fig. 7.1b), and therefore a dilution of the BAO signal. Since SLICS uses a bin width that is 2.5-times thicker than the one from MXXL, a detection of the BAO signal in SLICS data is more challenging.[2] From the technical aspect of the simulations, the detection significance is governed by the size of the simulated area, the size of individual lines-of-sight, as well as the number and maximum redshift of the simulated lens and source galaxies. SLICS provides

[1] which stands for 'Code for Anisotropies in the Microwave Background' and can be visited under camb.info

[2] This problem can be entirely circumvented by taking the full advantage of the spatial and time resolution of the N-body simulation at the expense of computational power. The `RAY-RAMSES` code (Barreira et al., 2016) provides such ray-tracing simulation. It has not been used in this thesis since its data products have not been made publicly available as of yet.

a sky coverage which is 27 times larger than the one from MXXL, while individual lines-of-sight are 31% smaller compared to MXXL. The number of lenses per given redshift is comparable, however, the SLICS provide lens galaxies as deep as $z_\mathrm{d} \approx 0.5$ and the MXXL as deep as $z_\mathrm{d} \approx 3$.

Figure 7.6c shows that the convergence measured around random positions has a similar dependence on the angular scale regardless of redshift. Since angular scales translate to different comoving scales for different redshifts, potential problems with the box size of the N-body simulations would impact the convergence at different angular scales. However, the convergence measured around random points must be compatible with zero, which is violated especially for the highest source redshift bin. This hints to problems with the ray-tracing method, which must be investigated further.

In any case, the advantages and disadvantages of SLICS and MXXL stack in a way that makes a BAO detection impossible. The major drawback of the MXXL is most probably the lack of overall sky coverage. We exclude the size of an individual line-of-sight as the main problem. This assumption is based on the notion that otherwise high-redshift lenses with $1 \leq z_\mathrm{d} < 1.5$ and a corresponding BAO scale of $1.3^\circ \lesssim \theta \lesssim 1.7^\circ$ should exhibit a clear BAO signal. Unfortunately, it is highly unlikely that additional data for the MXXL will be provided. The SLICS would benefit most from deeper lens galaxy data. A new lens catalogue is planned for the near future that contains LRG with a maximum redshift of $z_\mathrm{d} \approx 0.8$. If even the LRG do not exhibit a BAO signal, the projection bin width Δl is probably chosen too wide or the N-body box size too small, and SLICS cannot be used to constrain BAO.

The next immediate measures in this project are: waiting for new SLICS catalogue and sending an inquiry to obtain catalogues from `RAY-RAMSES`. In the time-frame of this doctoral thesis, the project 'BAO in GGL' was stopped at this point. Nonetheless, we suggest here several improvements of the analysis assuming that BAO will be eventually found in simulated or observed data. First of all, the assumption of a scale-independent, linear galaxy bias b must be treated with caution. The non-linearity becomes less important if the fitting range is kept in a small range around the BAO scale. In this case, the size of the fitting range must be chosen with care such that the galaxy bias is not entirely governed by the statistical uncertainty of the lensing signal. The galaxy bias is further important when averaging the lensing signal over many redshift bins. This is due to the slight dependence of the BAO scale on the galaxy bias which dilutes the BAO signal additionally. The impact can be estimated by choosing lens galaxy samples that have limiting magnitudes corresponding to a constant galaxy bias. In general, we must choose narrow redshift bins for the lens galaxies, since the angular BAO scale is a function of redshift.

Furthermore, the full covariance of the data must be used when lens redshifts are stacked to forecast reliable error estimates. Non-diagonal terms in the covariance matrix arise, e.g., from the strong correlation of the signal from neighbouring angular

bins. So far, only the variance has been estimated by a Jackknife method, which underestimates the total error. Finally, shape noise must be included in the analysis. Although shape noise is a sub-dominant source of noise compared to the sample variance, it also leads to an increased error budget. It further adds another source of correlated errors, i.e. off-diagonal terms in the covariance matrix, since the same source galaxies are used for different lens galaxies.

Ultimately, we want to observe the BAO signal in stage IV surveys. The number of lens-source pairs that are needed for the detection is huge and conventional tree-code algorithms are computationally expensive. The FFT-based method presented here is fast, precise, and easily applicable to any data catalogue format. This application is not only limited to the galaxy-shear correlation function but also the shear-shear correlation function can be analysed with this FFT-based algorithm. Cosmological constraints from simultaneously obtaining the BAO from these two in addition to the galaxy-galaxy correlation function will enable us to obtain cosmological parameters with unprecedented precision and accuracy.

CHAPTER 8

SUMMARY

This thesis is part of a worldwide effort to study the Universe using weak gravitational lensing. The focus lies on GGL which enables us to obtain precise mass estimates of the dark matter halos the galaxies reside in. Thus, GGL is crucial to understand the distribution of the elusive dark matter and to constrain cosmological parameters. Upcoming surveys will provide extensive data sets with unprecedented statistical accuracy. The data analysis depends on a sound theoretical model as well as capable numerical tools to deal with large data sets; this work contributes to both. The following chapter provides a brief summary of my main projects and concludes with an outlook.

`ShearCo` – fast extraction of shear estimates

In GGL, we measure the galaxy-matter correlation function in terms of the tangential shear around lens galaxies. This correlation function is calculated as the convolution of the lens density and a background shear field. Using the convolution theorem, an efficient FFT method can be used to calculate the tangential shear. In the framework of this doctoral thesis, the numerical tool `ShearCo` was developed and made publicly available under `github.com/SandraUnruh/shearco`. This tool is geared towards applications in GGL, although cosmic shear analyses are also possible. The computational time only depends on the size of the regular grid the galaxies need to be projected to, while the number of lens-source pairs is irrelevant. This is particularly useful for the extensive data sets expected from upcoming surveys.

The numerical accuracy is only limited by the angular resolution of the regular grid. Masks and patchy survey footprints are handled by default. For GGL analyses, several modules were added which can be activated in a configuration file. The

modules mainly estimate, mitigate or model weak lensing magnification effects. Details about the implementation and the validation of the code are found in Chap. 4, and three applications are given in Chap. 5, 6 and 7.

Magnification bias in the shear-ratio test

In weak lensing surveys, unbiased galaxy shape and redshift estimates are indispensable for cosmological data analysis. The SRT is a null-test that probes for systematics in both estimates. It is a geometrical probe which is almost independent of cosmology. Using one lens and two source populations, the SRT equates the ratio of tangential shear values to the ratio of their corresponding lensing efficiencies. Chapter 5 shows that this is only true if weak lensing magnification can be neglected. Magnification by the LSS between us and the lenses changes the number count of lens galaxies, and at the same time, induces a shear on the background galaxies. Thus, the bias is stronger with increasing lens redshift because more LSS can be found between us and the lenses. For a fixed lens redshift, the bias is stronger the closer the sources reside to the lenses. The magnification bias further depends on the angular scale on which the tangential shear is measured as well as the selection criteria for the lens galaxies.

Using simulations, we show that an SRT with flux-limited lens galaxies at redshift $z_\mathrm{d} = 0.8$ deviates up to 9σ from zero. We conclude that neglecting magnification effects in the SRT is unsuitable for ongoing and future weak lensing surveys. Unfortunately, it is very challenging to observe the magnification effects directly. In Chap. 5 we provide a mitigation strategy that solely relies on already present observational data. The mitigation reduces the bias by a factor of ~ 100 and, at the same time, reduces the total uncertainties. This results in a deviation of typically $< 1\sigma$. The numerical implementation is available within `ShearCo` as one of the options to calculate magnification-corrected shear estimates.

Magnification effects in galaxy-galaxy lensing

Chapter 6 generalises and extends the findings from Chap. 5 for GGL. We provide a leading-order analytical description for the magnification effects by the LSS between us and the lenses. Furthermore, we present a new method to select a lens galaxy sample that is unbiased by magnification effects. The method is particularly useful for weak lensing simulations since the magnification of the lenses must be known. Additionally, we explore the magnification effects that the matter associated with the lens galaxies induces on the source galaxies. For the first time, we provide an analytical approximation of this effect which is given by the third-order cross-correlation of the lenses' number density, and the shear and convergence as experienced by background galaxies. We further provide a numerical method to cal-

culate magnification-affected and unaffected source catalogues.[1] The magnification effect of the source galaxies is typically $\lesssim 1\%$ of the measured GGL signal.

We conclude our analysis by quantifying the magnification bias on halo mass estimates. The mean mass profile of halos around galaxies can be inferred from tangential shear estimates using a halo model prescription. We then estimate the bias in shear and halo mass for a KiDS+VIKING+GAMA-like survey. For lens galaxies at redshift $z_\mathrm{d} = 0.36$ and source galaxies with mean redshift $\bar{z}_\mathrm{s} = 0.79$, the shear profile is changed by 2% and the mass is biased by 8%. Thus, magnification effects by source and lens galaxies must be carefully taken into account for ongoing surveys, and the statistical power of future weak lensing surveys certainly warrants correction for this effect.

Baryon acoustic oscillations in galaxy-galaxy lensing

BAO are frozen-in density fluctuations in the LSS that were generated in the early Universe. They have been detected in the distribution of galaxies and are, among others, also observable in weak lensing correlation functions. Upcoming weak lensing surveys will provide the statistical power to measure the BAO signal in tomographic cosmic shear and GGL surveys. Chapter 7 aims to detect the BAO in simulated weak lensing catalogues. The BAO signal is first modelled analytically and then estimated from numerical data. However, various problems with the ray-tracing simulations were discovered that prohibit detection of BAO in the lensing simulations. Nonetheless, the analysis pipeline has been set-up successfully and improved data catalogues can be analysed on the time-scale of an hour. In the framework of the doctoral thesis, the project was stopped here, but it will be continued in the future with the overarching objective of measuring BAO in upcoming lensing surveys for the first time.

Outlook

We live in exciting times. The standard model of cosmology has been proven to beautifully explain observations for a wide range of time scales and physical scales. Now, the physical origin of dark matter and dark energy challenges the scientific community. Weak gravitational lensing surveys will soon reach the accuracy to shed light on the dark sector. The satellites *Euclid* and WFIRST[2], and the ground-based LSST will provide redshift and shape information of several billion galaxies up to a redshift of $z \approx 3$ with an unprecedented and partly overlapping sky coverage. The improved statistics require tighter control over observational systematics and accurate theoretical predictions. Only then will a joint analysis of shear-shear,

[1] The mentioned numerical methods and numerical implementations of the analytical approximations are made publicly available with `ShearCo`.

[2] While writing this chapter, WFIRST has been renamed as the Nancy Grace Roman Space Telescope, in honour of 'the Mother of Hubble', Nancy Roman.

galaxy-shear and galaxy-galaxy correlations yield results with the necessary accuracy and precision to explore dark matter and dark energy. The scientific community is putting many resources in understanding all the nuances in the data analysis. This work contributes to the effort.

APPENDIX A

BREAK-DOWN OF THE THIN-LENS APPROXIMATION

A common assumption in GGL is that the mass concentration relevant for lensing is localised, i.e., the extension of lenses along the line-of-sight is small compared to the distances between observer, lens, and source. This is referred to as the TLA. The TLA breaks down when the LSS as an inhomogeneous, extended mass distribution is taken into account. Since the statistical properties of the LSS are a powerful probe for cosmology, the continuous deflection and distortion are studied in cosmic shear analyses (for a recent review see, e.g., Dodelson, 2017). In GGL the extent of the lens, which is determined by its physical size and its correlation to the surrounding LSS, is typically neglected. In the following, we drop the assumption of the TLA and calculate the change of the tangential shear signal as well as its implications for the SRT analytically for an SIS model.

A.1. Analytical treatment for an SIS

The convergence (4.8) in a flat Universe is given as

$$\kappa(\theta, \chi_\mathrm{s}) = \frac{3H_0^2\Omega_\mathrm{m}}{2c^2} \int_0^{\chi_\mathrm{s}} \mathrm{d}\chi \, \frac{\chi(\chi_\mathrm{s} - \chi)}{\chi_\mathrm{s}} \, \frac{\delta(\chi\theta, \chi)}{a(\chi)} \,, \tag{A.1}$$

where χ_s is the comoving distance to the source and κ is somtimes referred to as the effective (dimensionless) surface mass density. The expression (A.1) is valid in the Born approximation and neglects lens-lens coupling. The fractional density contrast is

$$\delta = \frac{\Delta\rho}{\bar{\rho}} = \frac{8\pi G}{3H_0^2} \frac{a^3}{\Omega_\mathrm{m}} \Delta\rho \,, \tag{A.2}$$

and the surface mass density is $\Delta\rho\,\mathrm{d}D_{\mathrm{prop}}$ in a small proper radial distance interval $\mathrm{d}D_{\mathrm{prop}} = c\,\mathrm{d}t = a\,\mathrm{d}\chi$. The TLA is equivalent of assuming that the main contribution to the integral in (4.8) comes from $\chi \approx \chi_{\mathrm{d}}$, where χ_{d} is the comoving distance from the centre of the lens to us. Thus, we can replace all χ's by χ_{d}, and extend the integration limits to infinity which simplifies the distance ratio in (4.8) to $\chi(\chi_{\mathrm{s}} - \chi)/\chi_{\mathrm{s}} = \chi_{\mathrm{d}}(\chi_{\mathrm{s}} - \chi_{\mathrm{d}})/\chi_{\mathrm{s}} = D_{\mathrm{d}}D_{\mathrm{ds}}/(a\,D_{\mathrm{s}})$. We insert this result and (A.2) into (4.8) and recover the convergence (3.7) derived in Sect. 3.1.2.

To quantitatively test the implications of the TLA, we employ an SIS model whose radial mass profile behaves as $\rho = C\,r^{-2}$ (cf. Sect. 3.2). Using the SIS profile in comoving coordinates superposed on the homogeneous background, we find

$$\Delta\rho = \frac{C}{a^2\,[(\chi - \chi_{\mathrm{d}})^2 + (\chi\theta)^2]}\,, \tag{A.3}$$

where the two terms in the denominator account for the separation along and transversal to the line-of-sight from the lens centre. For consistency, we choose the constant C to be

$$C = \frac{\chi_{\mathrm{s}}\,\theta_{\mathrm{E}}}{2\pi\,(\chi_{\mathrm{s}} - \chi_{\mathrm{d}})}\,, \tag{A.4}$$

and we obtain

$$\kappa(\theta, \chi_{\mathrm{s}}) = \frac{\theta_{\mathrm{E}}}{2\pi\,(\chi_{\mathrm{s}} - \chi_{\mathrm{d}})} \int_0^{\chi_{\mathrm{s}}} \mathrm{d}\chi\; \frac{\chi(\chi_{\mathrm{s}} - \chi)}{(\chi - \chi_{\mathrm{d}})^2 + (\chi\theta)^2}\,. \tag{A.5}$$

We employ the TLA and recover the first expression in Eq. (3.18)

$$\kappa_0(\theta, \chi_{\mathrm{s}}) = \frac{\theta_{\mathrm{E}}\chi_{\mathrm{d}}}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\chi\; \frac{1}{(\chi - \chi_{\mathrm{d}})^2 + (\chi_{\mathrm{d}}\theta)^2} = \frac{\theta_{\mathrm{E}}}{2\theta}\,. \tag{A.6}$$

The full integral in Eq. (A.5) can be solved analytically, but the resulting expression is not very insightful. However, we can restrict our attention to small angles. The fractional difference between κ and κ_0 becomes

$$\begin{aligned}\frac{\kappa_0 - \kappa}{\kappa_0} &= \left[2\chi_{\mathrm{s}} + (2\chi_{\mathrm{d}} - \chi_{\mathrm{s}})\,\ln\left(\frac{\chi_{\mathrm{s}} - \chi_{\mathrm{d}}}{\chi_{\mathrm{d}}}\right)\right] \frac{\theta}{(\chi_{\mathrm{s}} - \chi_{\mathrm{d}})\pi} \\ &\quad + \frac{\chi_{\mathrm{s}} - 3\chi_{\mathrm{d}}}{\chi_{\mathrm{s}} - \chi_{\mathrm{d}}}\theta^2 + \mathcal{O}\left(\theta^3\right) \\ &= r_1(\chi_{\mathrm{s}})\theta + r_2(\chi_{\mathrm{s}})\theta^2 + \mathcal{O}\left(\theta^3\right)\,,\end{aligned} \tag{A.7}$$

where in the last step we defined the coefficients $r_i(\chi_{\mathrm{s}})$. Making use of (A.6), we then find for the convergence

$$\kappa(\theta, \chi_{\mathrm{s}}) = \frac{\theta_{\mathrm{E}}}{2\theta}\left[1 - r_1(\chi_{\mathrm{s}})\theta - r_2(\chi_{\mathrm{s}})\theta^2 + \mathcal{O}\left(\theta^3\right)\right]\,. \tag{A.8}$$

Hence, by dropping the TLA, the convergence attains additional terms, the lowest one of which corresponding to a constant. However, a constant surface mass density does not yield any shear and this term is irrelevant for weak lensing studies. The tangential shear becomes

$$\gamma_{\mathrm{t}}^{\mathrm{SIS}}(\theta, \chi_{\mathrm{s}}) = \frac{\theta_{\mathrm{E}}}{2\theta}\left[1 + \frac{1}{3} r_2(\chi_{\mathrm{s}})\theta^2 + \mathcal{O}\left(\theta^3\right)\right] . \tag{A.9}$$

For an order of magnitude estimate, we can consider a lens galaxy at comoving redshift $\chi_{\mathrm{d}} = 1000\,\mathrm{Mpc}$ and a source population at $\chi_{\mathrm{s}} = 1100\,\mathrm{Mpc}$. For a fairly large angle $\theta = 1^\circ$ the deviation is of the order of 10^{-5}; smaller angles lead to even smaller deviations.

A.2. Effects on the shear-ratio test

The SRT relates shear measurements corresponding to same lens, which is just given by a ratio of Einstein radii for an SIS. However, according to the considerations above, the effective lens and its redshift are slightly different for every lens-source pair and, thus, we introduce a systematic bias in (5.12) if we do not account for this effect. Since the effective lens redshift is a bit lower than the redshift where the actual lens galaxy resides, we underpredict the ratio of lensing efficiencies $\beta(z_{\mathrm{d}}, z_2)/\beta(z_{\mathrm{d}}, z_1)$ for $z_1 > z_2$, leading to a positive bias if we choose the null test as in Eq. (5.12). We thus consider the difference

$$\frac{\gamma_{\mathrm{t}}^{\mathrm{SIS}}(\theta, \chi_{\mathrm{s}2})}{\gamma_{\mathrm{t}}^{\mathrm{SIS}}(\theta, \chi_{\mathrm{s}1})} - \frac{\chi_{\mathrm{s}2} - \chi_{\mathrm{d}}}{\chi_{\mathrm{s}2}} \frac{\chi_{\mathrm{s}1}}{\chi_{\mathrm{s}1} - \chi_{\mathrm{d}}} = \frac{2\chi_{\mathrm{s}1}\left(\chi_{\mathrm{s}2} - \chi_{\mathrm{s}1}\right)\chi_{\mathrm{d}}}{3\chi_{\mathrm{s}2}(\chi_{\mathrm{s}1} - \chi_{\mathrm{d}})^2}\theta^2 . \tag{A.10}$$

An order of magnitude estimate as before with a second source population at $\chi_{\mathrm{s}2} = 1200\,\mathrm{Mpc}$, yields for $\theta = 1^\circ$ a deviation of 10^{-3}, where again smaller angles lead to smaller deviations.

We now compare (A.10) to the results of the null test (5.12) in Sect. 5. The same shear estimator $\bar{\gamma}_{\mathrm{t}}$ (5.7) is used as well as the same weighting function $w(\theta; \theta_{\mathrm{in}}, \theta_{\mathrm{out}})$

$$\int_{\theta_{\mathrm{in}}}^{\theta_{\mathrm{out}}} \mathrm{d}\theta\, \frac{1}{\theta}\gamma_{\mathrm{t}}^{\mathrm{SIS}}(\theta, \chi_{\mathrm{s}}) = \frac{\chi_{\mathrm{s}} - 3\chi_{\mathrm{d}}}{3(\chi_{\mathrm{s}} - \chi_{\mathrm{d}})}\left(\theta_{\mathrm{out}} - \theta_{\mathrm{in}}\right) + \frac{1}{\theta_{\mathrm{in}}} - \frac{1}{\theta_{\mathrm{out}}} . \tag{A.11}$$

Also, the redshifts are taken to match those considered in Sect. 5. The result can be seen in Fig. A.1; the deviation from zero is of the order 10^{-6}. The bias increases with lens redshift z_{d}, likewise, a smaller separation along the line-of-sight leads to a larger bias. Compared to the results from Fig. 5.1 the bias induced by assuming the TLA shows a similar behaviour with redshift but is six orders of magnitude smaller.

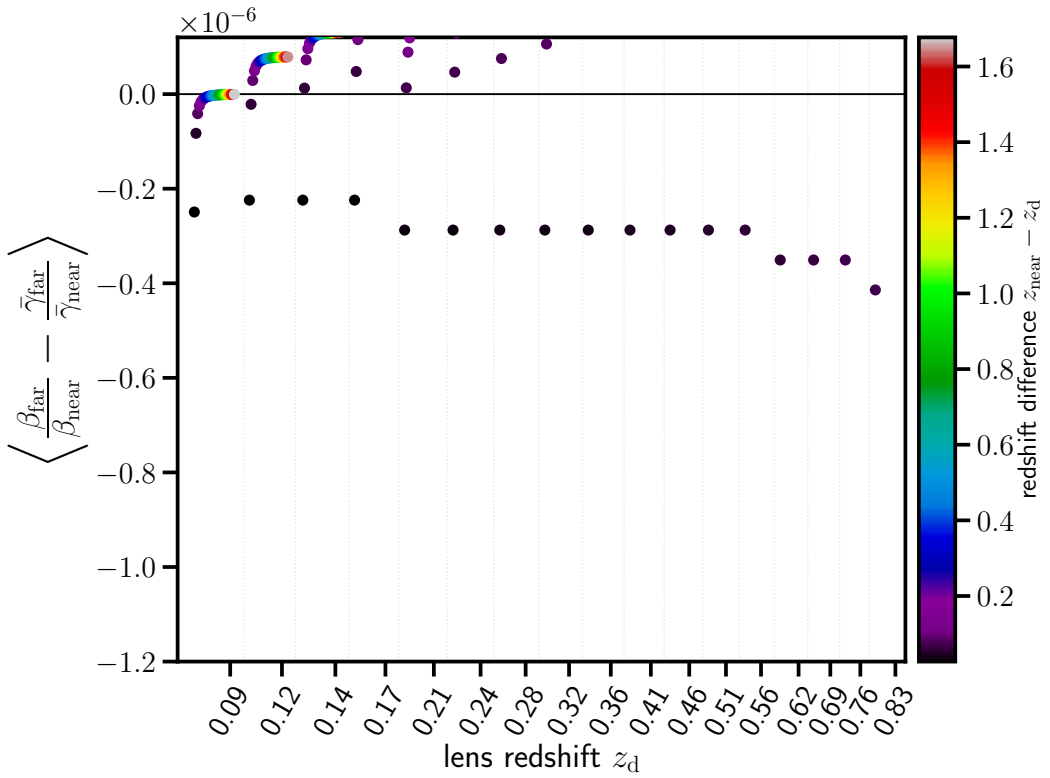


Figure A.1. – Shear ratio test of two populations of source galaxies that reside farther or nearer to the lens galaxy, i.e. $z_{\rm far} > z_{\rm near}$. The dotted lines separate shear-ratios from different lens galaxies, while different source redshifts are highlighted by colour code, each colour represents a combination of two sources at $z_{\rm s}$ and $z_{\rm s+1}$. On the y-axis a deviation from zero can be seen that arises if the TLA is removed.

A.3. Conclusions

The breakdown of the TLA is a function of χ_d, χ_s, and the shape of the extended mass distribution ρ. It also depends on the transverse distance to the lens as well as slightly on the cosmology through the distance-redshift relation. However, the systematic that arises from employing the TLA is a sub-dominant effect for all current purposes, which includes stage IV experiments.

We further note that an SIS is only a valid choice to describe the 1-halo term of the lensing signal. The 2-halo term usually has a shallower profile than an SIS and roughly scales with $\rho \propto r^{-1.8}$. It is certainly the dominant contribution to the lensing signal at comoving scales larger than 2 Mpc. However, for the SRT, the nominator and denominator in (A.10) are affected alike and the difference to an SIS is only marginal. We omit further investigations into this due to the current irrelevance of this effect.

BIBLIOGRAPHY

Abbott, T. M. C., Abdalla, F. B., Alarcon, A., et al. (2018). Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing. *Phys. Rev. D*, 98:043526.

Abramo, L. R. and Leonard, K. E. (2013). Why multitracer surveys beat cosmic variance. *MNRAS*, 432:318–326.

Ahmad, Q. R., Allen, R. C., Andersen, T. C., et al. (2001). Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301.

Aihara, H., Arimoto, N., Armstrong, R., et al. (2018). The Hyper Suprime-Cam SSP Survey: Overview and survey design. *PASJ*, 70:S4.

Alam, S., Ata, M., Bailey, S., et al. (2017). The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *MNRAS*, 470:2617–2652.

Albrecht, A., Bernstein, G., Cahn, R., et al. (2006). Report of the Dark Energy Task Force. *arXiv e-prints*, astro–ph/0609591.

Alcock, C., Allsman, R. A., Alves, D. R., et al. (2000). The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations. *ApJ*, 542:281–307.

Amon, A., Heymans, C., Klaes, D., et al. (2018). KiDS-i-800: comparing weak gravitational lensing measurements from same-sky surveys. *MNRAS*, 477:4285–4307.

Angulo, R. E., Springel, V., White, S. D. M., et al. (2012). Scaling relations for galaxy clusters in the Millennium-XXL simulation. *MNRAS*, 426:2046–2062.

Asgari, M., Heymans, C., Blake, C., et al. (2017). Revisiting CFHTLenS cosmic shear: optimal E/B-mode decomposition using COSEBIs and compressed COSEBIs. *MNRAS*, 464:1676–1692.

Ata, M., Baumgarten, F., Bautista, J., et al. (2018). The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: first measurement of baryon acoustic oscillations between redshift 0.8 and 2.2. *MNRAS*, 473:4773–4794.

Bardeen, J. M., Bond, J. R., Kaiser, N., and Szalay, A. S. (1986). The Statistics of Peaks of Gaussian Random Fields. *ApJ*, 304:15–61.

Barreira, A., Llinares, C., Bose, S., and Li, B. (2016). `RAY-RAMSES`: a code for ray tracing on the fly in N-body simulations. *J. Cosmology Astropart. Phys.*, 2016:001.

Bartelmann, M. (1996). Arcs from a universal dark-matter halo profile. *A&A*, 313:697–702.

Bartelmann, M. and Schneider, P. (2001). Weak gravitational lensing. *Phys. Rep.*, 340:291–472.

Bartelmann, M., Viola, M., Melchior, P., and Schäfer, B. M. (2012). Calibration biases in measurements of weak lensing. *A&A*, 547:A98.

Bartolo, N., Komatsu, E., Matarrese, S., and Riotto, A. (2004). Non-Gaussianity from inflation: theory and observations. *Phys. Rep.*, 402:103–266.

Bassett, B. and Hlozek, R. (2010). *Dark Energy: Observational and Theoretical Approaches*, chapter 'Baryon acoustic oscillations', pp. 246. Cambridge University Press.

Bautista, J. E., Busca, N. G., Guy, J., et al. (2017). Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with SDSS DR12 Lyα-Forests. *A&A*, 603:A12.

Bernardeau, F., Nishimichi, T., and Taruya, A. (2020). Observing Baryonic Acoustic Oscillations in tomographic cosmic shear surveys. *arXiv e-prints*, arXiv:2004.03201.

Bernardi, M., Sheth, R. K., Annis, J., et al. (2003). Early-Type Galaxies in the Sloan Digital Sky Survey. III. The Fundamental Plane. *AJ*, 125:1866–1881.

Bernstein, G. and Jain, B. (2004). Dark Energy Constraints from Weak-Lensing Cross-Correlation Cosmography. *ApJ*, 600:17–25.

Bertin, G. and Lombardi, M. (2006). Looking at the Fundamental Plane through a Gravitational Lens. *ApJ*, 648:L17–L20.

Bertone, G., Hooper, D., and Silk, J. (2005). Particle dark matter: evidence candidates and constraints. *Phys. Rep.*, 405:279–390.

Betoule, M., Kessler, R., Guy, J., et al. (2014). Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *A&A*, 568:A22.

Blake, C. A., Abdalla, F. B., Bridle, S. L., and Rawlings, S. (2004). Cosmology with the SKA. *New A Rev.*, 48:1063–1077.

Blandford, R. and Narayan, R. (1986). Fermat's Principle, Caustics, and the Classification of Gravitational Lens Images. *ApJ*, 310:568–582.

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. (2017). Sloan Digital Sky Survey IV: Mapping the Milky Way Nearby Galaxies and the Distant Universe. *AJ*, 154:28–62.

Blazek, J. A., McEwen, J. E., and Hirata, C. M. (2016). Streaming Velocities and the Baryon Acoustic Oscillation Scale. *Phys. Rev. Lett.*, 116:121303.

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. (2015). Reionization After *Planck*: The Derived Growth of the Cosmic Ionizing Emissivity Now Matches the Growth of the Galaxy UV Luminosity Density. *ApJ*, 811:140–158.

Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. (2001). Profiles of dark haloes: evolution scatter and environment. *MNRAS*, 321:559–575.

Cai, R.-G., Guo, Z.-K., and Yang, T. (2016). Null test of the cosmic curvature using $H(z)$ and supernovae data. *Phys. Rev. D*, 93:043517.

Capozzi, F., Di Valentino, E., Lisi, E., et al. (2017). Global constraints on absolute neutrino masses and their ordering. *Phys. Rev. D*, 95:096014.

Chang, C. and Jain, B. (2014). Delensing galaxy surveys. *MNRAS*, 443:102–110.

Chang, T.-C., Pen, U.-L., Peterson, J. B., and McDonald, P. (2008). Baryon Acoustic Oscillation Intensity Mapping of Dark Energy. *Phys. Rev. Lett.*, 100:091303.

Chiu, I., Dietrich, J. P., Mohr, J., et al. (2016). Detection of enhancement in number densities of background galaxies due to magnification by massive galaxy clusters. *MNRAS*, 457:3050–3065.

Chua, K. T. E., Pillepich, A., Vogelsberger, M., and Hernquist, L. (2019). Shape of dark matter haloes in the Illustris simulation: effects of baryons. *MNRAS*, 484:476–493.

Clampitt, J., Sánchez, C., Kwan, J., et al. (2017). Galaxy-galaxy lensing in the Dark Energy Survey Science Verification data. *MNRAS*, 465:4204–4218.

Cole, S., Percival, W. J., Peacock, J. A., et al. (2005). The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *MNRAS*, 362:505–534.

Colless, M., Dalton, G., Maddox, S., et al. (2001). The 2dF Galaxy Redshift Survey: spectra and redshifts. *MNRAS*, 328:1039–1063.

Cooray, A. and Sheth, R. (2002). Halo models of large scale structure. *Phys. Rep.*, 372:1–129.

Davis, M., Efstathiou, G., Frenk, C. S., and White, S. D. M. (1985). The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292:371–394.

Davis, M. and Peebles, P. J. E. (1983). A survey of galaxy redshifts. V. The two-point position and velocity correlations. *ApJ*, 267:465–482.

Dawson, W. A., Schneider, M. D., Tyson, J. A., and Jee, M. J. (2016). The Ellipticity Distribution of Ambiguously Blended Objects. *ApJ*, 816:11–21.

de Carvalho, E., Bernui, A., Xavier, H. S., and Novaes, C. P. (2020). Baryon acoustic oscillations signature in the three-point angular correlation function from the SDSS-DR12 quasar survey. *MNRAS*, 492:4469–4476.

de Graaff, A., Cai, Y.-C., Heymans, C., and Peacock, J. A. (2019). Probing the missing baryons with the Sunyaev-Zeldovich effect from filaments. *A&A*, 624:A48.

de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., et al. (2017). The third data release of the Kilo-Degree Survey and associated data products. *A&A*, 604:A134.

Diemer, B. (2018). `colossus`: a python toolkit for cosmology, large-scale structure, and dark matter halos. *ApJS*, 239:35–48.

Ding, Z., Seo, H.-J., Huff, E., Saito, S., and Clowe, D. (2019). Detecting baryon acoustic oscillations in dark matter from kinematic weak lensing surveys. *MNRAS*, 487:253–267.

Dodelson, S. (2003). *Modern cosmology*. Academic Press.

Dodelson, S. (2017). *Gravitational Lensing*. Cambridge University Press.

Domínguez, A., Wojtak, R., Finke, J., et al. (2019). A New Measurement of the Hubble Constant and Matter Content of the Universe Using Extragalactic Background Light γ-Ray Attenuation. *ApJ*, 885:137–144.

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. (2011). Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, 413:971–995.

Duan, Y. and Eisenstein, D. (2019). Tests of acoustic scale shifts in halo-based mock galaxy catalogues. *MNRAS*, 490:2718–2731.

Einstein, A. (1916). Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354:769–822.

Eisenstein, D. J. and Hu, W. (1998). Baryonic Features in the Matter Transfer Function. *ApJ*, 496:605–614.

Eisenstein, D. J., Seo, H.-J., Sirko, E., and Spergel, D. N. (2007). Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak. *ApJ*, 664:675–679.

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. (2005). Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *ApJ*, 633:560–574.

Etherington, I. M. H. (1933). On the Definition of Distance in General Relativity. *Philosophical Magazine*, 15:761–773.

Freese, K. (2009). Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars. *EAS Publications Series*, 36:113–126.

Frigo, M. and Johnson, S. G. (2012). `fftw`: fastest fourier transform in the west. ascl:1201.015.

Fukuda, Y., Hayakawa, T., Ichihara, E., et al. (1998). Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81:1562–1567.

Fukugita, M. and Peebles, P. J. E. (2004). The Cosmic Energy Inventory. *ApJ*, 616:643–668.

Garcia-Fernandez, M., Sanchez, E., Sevilla-Noarbe, I., et al. (2018). Weak lensing magnification in the Dark Energy Survey Science Verification data. *MNRAS*, 476:1071–1085.

Gerbino, M. and Lattanzi, M. (2017). Status of neutrino properties and future prospects – Cosmological and astrophysical constraints. *Frontiers in Physics*, 5:70.

Grassi, A. and Schäfer, B. M. (2014). Detecting baryon acoustic oscillations by 3D weak lensing. *MNRAS*, 437:2632–2641.

Gurvich, A. and Mandelbaum, R. (2016). The impact of correlated noise on galaxy shape estimation for weak lensing. *MNRAS*, 457:3522–3534.

Guyonnet, A., Astier, P., Antilogus, P., Regnault, N., and Doherty, P. (2015). Evidence for self-interaction of charge distribution in charge-coupled devices. *A&A*, 575:A41.

Hada, R. and Eisenstein, D. J. (2018). An iterative reconstruction of cosmological initial density fields. *MNRAS*, 478:1866–1874.

Harnois-Déraps, J., Amon, A., Choi, A., et al. (2018). Cosmological simulations for combined-probe analyses: covariance and neighbour-exclusion bias. *MNRAS*, 481:1337–1367.

Harnois-Déraps, J., Tröster, T., Chisari, N. E., et al. (2017). KiDS-450: tomographic cross-correlation of galaxy shear with *Planck* lensing. *MNRAS*, 471:1619–1633.

Harnois-Déraps, J. and van Waerbeke, L. (2015). Simulations of weak gravitational lensing – II. Including finite support effects in cosmic shear covariance matrices. *MNRAS*, 450:2857–2873.

Harrison, E. R. (1970). Fluctuations at the Threshold of Classical Cosmology. *Phys. Rev. D*, 1:2726–2730.

Hartlap, J. (2009). *Ray-Tracing Simulations of Weak Gravitational Lensing*. PhD thesis, AIfA – Argelander-Institut für Astronomie, Rheinische Friedrich-Wilhelms Universität Bonn, Germany, `http://hss.ulb.uni-bonn.de/2009/1796/1796.htm`.

Heavens, A. (2003). 3D weak lensing. *MNRAS*, 343:1327–1334.

Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. (2015). Galaxy formation in the *Planck* cosmology – I. Matching the observed evolution of star formation rates colours and stellar masses. *MNRAS*, 451:2663–2680.

Heymans, C., Van Waerbeke, L., Miller, L., et al. (2012). CFHTLenS: the Canada-France-Hawaii Telescope Lensing Survey. *MNRAS*, 427:146–166.

Hilbert, S., Barreira, A., Fabbian, G., et al. (2020). The accuracy of weak lensing simulations. *MNRAS*, 493:305–319.

Hilbert, S., Hartlap, J., White, S. D. M., and Schneider, P. (2009). Ray-tracing through the Millennium Simulation: Born corrections and lens-lens coupling in cosmic shear and galaxy-galaxy lensing. *A&A*, 499:31–43.

Hilbert, S., Xu, D., Schneider, P., et al. (2017). Intrinsic alignments of galaxies in the Illustris simulation. *MNRAS*, 468:790–823.

Hildebrandt, H., Arnouts, S., Capak, P., et al. (2010). PHAT: PHoto-z Accuracy Testing. *A&A*, 523:A31.

Hildebrandt, H., Köhlinger, F., van den Busch, J. L., et al. (2020). KiDS+VIKING-450: Cosmic shear tomography with optical and infrared data. *A&A*, 633:A69.

Hildebrandt, H., Viola, M., Heymans, C., et al. (2017). KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. *MNRAS*, 465:1454–1498.

Hirata, C. M. and Seljak, U. (2003). Reconstruction of lensing from the cosmic microwave background polarization. *Phys. Rev. D*, 68:083002.

Hockney, R. W. and Eastwood, J. W. (1988). *Computer simulation using particles*. Routledge.

Hoekstra, H. (2013). Weak gravitational lensing. *arXiv e-prints*, arXiv:1312.5981.

Hotokezaka, K., Nakar, E., Gottlieb, O., et al. (2019). A Hubble constant measurement from superluminal motion of the jet in GW170817. *Nature Astronomy*, 3:940–944.

Huff, E. M. and Graves, G. J. (2011). Magnificent Magnification: Exploiting the Other Half of the Lensing Signal. *arXiv e-prints*, arXiv:1111.1070.

Hui, L., Gaztañaga, E., and Loverde, M. (2007). Anisotropic magnification distortion of the 3D galaxy correlation. I. Real space. *Phys. Rev. D*, 76:103502.

Huterer, D. and Shafer, D. L. (2018). Dark energy two decades after: observables probes consistency tests. *Reports on Progress in Physics*, 81:016901.

Illingworth, G. D., Magee, D., Oesch, P. A., et al. (2013). The HST eXtreme Deep Field (XDF): Combining All ACS and WFC3/IR Data on the HUDF Region into the Deepest Field Ever. *ApJS*, 209:6–19.

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. (2019). LSST: From Science Drivers to Reference Design and Anticipated Data Products. *ApJ*, 873:111–155.

Jain, B. and Taylor, A. (2003). Cross-Correlation Tomography: Measuring Dark Energy Evolution with Weak Lensing. *Phys. Rev. Lett.*, 91:141302.

Jenkins, A., Frenk, C. S., White, S. D. M., et al. (2001). The mass function of dark matter haloes. *MNRAS*, 321:372–384.

Jeong, D., Komatsu, E., and Jain, B. (2009). Galaxy-CMB and galaxy-galaxy lensing on large scales: Sensitivity to primordial non-Gaussianity. *Phys. Rev. D*, 80:123527.

Joachimi, B., Cacciato, M., Kitching, T. D., et al. (2015). Galaxy Alignments: An Overview. *Space Sci. Rev.*, 193:1–65.

Joudaki, S., Blake, C., Johnson, A., et al. (2018). KiDS-450 + 2dFLenS: Cosmological parameter constraints from weak gravitational lensing tomography and overlapping redshift-space galaxy clustering. *MNRAS*, 474:4894–4924.

Kaiser, N., Squires, G., and Broadhurst, T. (1995). A Method for Weak Lensing Observations. *ApJ*, 449:460–475.

Kannawadi, A., Shapiro, C. A., Mandelbaum, R., et al. (2016). The Impact of Interpixel Capacitance in CMOS Detectors on PSF Shapes and Implications for WFIRST. *PASP*, 128:095001.

Kilbinger, M. (2015). Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78:086901.

Kilbinger, M., Bonnett, C., and Coupon, J. (2014). `athena`: tree code for second-order correlation functions. ascl:1402.026.

Kilbinger, M., Heymans, C., Asgari, M., et al. (2017). Precision calculations of the cosmic shear power spectrum projection. *MNRAS*, 472:2126–2141.

Kitching, T. D., Heavens, A. F., Taylor, A. N., et al. (2007). Cosmological constraints from COMBO-17 using 3D weak lensing. *MNRAS*, 376:771–778.

Knebe, A., Knollmann, S. R., Muldrew, S. I., et al. (2011). Haloes gone MAD: The Halo-Finder Comparison Project. *MNRAS*, 415:2293–2318.

Laureijs, R., Amiaux, J., Arduini, S., et al. (2011). Euclid Definition Study Report. *arXiv e-prints*, arXiv:1110.3193.

Lewis, A., Challinor, A., and Lasenby, A. (2000). Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ*, 538:473–476.

Li, X.-D., Park, C., Sabiu, C. G., et al. (2017). Cosmological Constraints from the Redshift Dependence of the Volume Effect Using the Galaxy 2-point Correlation Function across the Line of Sight. *ApJ*, 844:91–102.

Macaulay, E., Nichol, R. C., Bacon, D., et al. (2019). First cosmological results using Type Ia supernovae from the Dark Energy Survey: measurement of the Hubble constant. *MNRAS*, 486:2184–2196.

Mandelbaum, R. (2018). Weak Lensing for Precision Cosmology. *ARA&A*, 56:393–433.

Martel, H., Premadi, P., and Matzner, R. (2002). Light Propagation in Inhomogeneous Universes. III. Distributions of Image Separations. *ApJ*, 570:17–32.

McBride, C. K., Connolly, A. J., Gardner, J. P., et al. (2011). Three-point Correlation Functions of SDSS Galaxies: Constraining Galaxy-mass Bias. *ApJ*, 739:85–106.

Melchior, P. and Viola, M. (2012). Means of confusion: how pixel noise affects shear estimates for weak gravitational lensing. *MNRAS*, 424:2757–2769.

Moessner, R. and Jain, B. (1998). Angular cross-correlation of galaxies: a probe of gravitational lensing by large-scale structure. *MNRAS*, 294:L18–L24.

Navarro, J. F., Frenk, C. S., and White, S. D. M. (1996). The Structure of Cold Dark Matter Halos. *ApJ*, 462:563–575.

Navarro, J. F., Frenk, C. S., and White, S. D. M. (1997). A Universal Density Profile from Hierarchical Clustering. *ApJ*, 490:493–508.

Nicastro, F., Kaastra, J., Krongold, Y., et al. (2018). Observations of the missing baryons in the warm-hot intergalactic medium. *Nature*, 558:406–409.

Niemi, S.-M., Kitching, T. D., and Cropper, M. (2015). On weak lensing shape noise. *MNRAS*, 454:1221–1230.

Niikura, H., Takada, M., Yasuda, N., et al. (2019). Microlensing constraints on primordial black holes with Subaru/HSC Andromeda observations. *Nature Astronomy*, 3:524–534.

Ogata, H. (2005). A Numerical Integration Formula Based on the Bessel Functions. *Publ. RIMS*, 41:949–970.

Onions, J., Knebe, A., Pearce, F. R., et al. (2012). Subhaloes going Notts: the subhalo-finder comparison project. *MNRAS*, 423:1200–1214.

Peebles, P. J. E. (1970). Structure of the Coma Cluster of Galaxies. *AJ*, 75:13–20.

Peebles, P. J. E. (1993). *Principles of Physical Cosmology*. Princeton University Press.

Perlmutter, S., Aldering, G., Goldhaber, G., et al. (1999). Measurements of Ω and Λ from 42 High-Redshift Supernovae. *ApJ*, 517:565–586.

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. (2016). *Planck* 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13.

Planck Collaboration, Aghanim, N., Akrami, Y., et al. (2018a). *Planck* 2018 results. VI. Cosmological parameters. *arXiv e-prints*, arXiv:1807.06209.

Planck Collaboration, Akrami, Y., Arroja, F., et al. (2018b). *Planck* 2018 results. I. Overview and the cosmological legacy of *Planck*. *arXiv e-prints*, arXiv:1807.06205.

Plazas, A. A., Shapiro, C., Kannawadi, A., et al. (2016). The Effect of Detector Nonlinearity on WFIRST PSF Profiles for Weak Gravitational Lensing Measurements. *PASP*, 128:104001.

Prat, J., Sánchez, C., Fang, Y., et al. (2018). Dark Energy Survey year 1 results: Galaxy-galaxy lensing. *Phys. Rev. D*, 98:042005.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992a). *Numerical recipes in C. The art of scientific computing*. Cambridge University Press.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992b). *Numerical recipes in FORTRAN. The art of scientific computing*. Cambridge University Press.

Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., and Scolnic, D. (2019). Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM. *ApJ*, 876:85–98.

Riess, A. G., Filippenko, A. V., Challis, P., et al. (1998). Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ*, 116:1009–1038.

Robertson, H. P. (1935). Kinematics and World-Structure. *ApJ*, 82:284–301.

Ross, A. J., Beutler, F., Chuang, C.-H., et al. (2017). The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: observational systematics and baryon acoustic oscillations in the correlation function. *MNRAS*, 464:1168–1191.

Ruggeri, R. and Blake, C. (2019). How accurately can we measure the baryon acoustic oscillation feature? *arXiv e-prints*, arXiv:1909.13011.

Ryan, J., Chen, Y., and Ratra, B. (2019). Baryon acoustic oscillation Hubble parameter and angular size measurement constraints on the Hubble constant dark energy dynamics and spatial curvature. *MNRAS*, 488:3844–3856.

Saghiha, H., Simon, P., Schneider, P., and Hilbert, S. (2017). Confronting semi-analytic galaxy models with galaxy-matter correlations observed by CFHTLenS. *A&A*, 601:A98.

Schmittfull, M., Baldauf, T., and Zaldarriaga, M. (2017). Iterative initial condition reconstruction. *Phys. Rev. D*, 96:023505.

Schneider, A., Stoira, N., Refregier, A., et al. (2020). Baryonic effects for weak lensing. Part I. Power spectrum and covariance matrix. *J. Cosmology Astropart. Phys.*, 2020:019.

Schneider, P. (1996). Detection of dark matter concentrations via weak gravitational lensing. *MNRAS*, 283:837–853.

Schneider, P. (2003). Gravitational lensing as a probe of structure. *arXiv e-prints*, astro–ph/0306465.

Schneider, P. (2015). *Extragalactic Astronomy and Cosmology: An Introduction*. Springer Verlag.

Schneider, P. (2016). Generalized shear-ratio tests: A new relation between cosmological distances and a diagnostic for a redshift-dependent multiplicative bias in shear measurements. *A&A*, 592:L6.

Schneider, P., Ehlers, J., and Falco, E. E. (1992). *Gravitational Lenses*. Springer Verlag.

Schneider, P., Eifler, T., and Krause, E. (2010). COSEBIs: Extracting the full E-/B-mode information from cosmic shear correlation functions. *A&A*, 520:A116.

Schneider, P. and Watts, P. (2005). Galaxy-galaxy-galaxy lensing: Third-order correlations between the galaxy and mass distributions in the Universe. *A&A*, 432:783–795.

Schumann, M. (2019). Direct detection of WIMP dark matter: concepts and status. *Journal of Physics G Nuclear Physics*, 46:103003.

Scoccimarro, R., Sheth, R. K., Hui, L., and Jain, B. (2001). How Many Galaxies Fit in a Halo? Constraints on Galaxy Formation Efficiency from Spatial Clustering. *ApJ*, 546:20–34.

Scolnic, D. M., Jones, D. O., Rest, A., et al. (2018). The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample. *ApJ*, 859:101–129.

Scranton, R., Ménard, B., Richards, G. T., et al. (2005). Detection of Cosmic Magnification with the Sloan Digital Sky Survey. *ApJ*, 633:589–602.

Seljak, U., Aslanyan, G., Feng, Y., and Modi, C. (2017). Towards optimal extraction of cosmological information from nonlinear data. *J. Cosmology Astropart. Phys.*, 2017:009.

Sellentin, E. and Heavens, A. F. (2017). Quantifying lost information due to covariance matrix estimation in parameter inference. *MNRAS*, 464:4658–4665.

Semboloni, E., Hoekstra, H., Schaye, J., van Daalen, M. P., and McCarthy, I. G. (2011). Quantifying the effect of baryon physics on weak lensing tomography. *MNRAS*, 417:2020–2035.

Sheth, R. K. and Tormen, G. (1999). Large-scale bias and the peak background split. *MNRAS*, 308:119–126.

Shi, Y., Cautun, M., and Li, B. (2018). New method for initial density reconstruction. *Phys. Rev. D*, 97:023505.

Silk, J. (2013). Unleashing Positive Feedback: Linking the Rates of Star Formation Supermassive Black Hole Accretion and Outflows in Distant Galaxies. *ApJ*, 772:112–122.

Simon, P., Erben, T., Schneider, P., et al. (2013). CFHTLenS: higher order galaxy-mass correlations probed by galaxy-galaxy-galaxy lensing. *MNRAS*, 430:2476–2498.

Simon, P. and Hilbert, S. (2018). Scale dependence of galaxy biasing investigated by weak gravitational lensing: An assessment using semi-analytic galaxies and simulated lensing data. *A&A*, 613:A15.

Simon, P., Watts, P., Schneider, P., et al. (2008). First detection of galaxy-galaxy-galaxy lensing in RCS. A new tool for studying the matter environment of galaxy pairs. *A&A*, 479:655–667.

Singh, S., Mandelbaum, R., Seljak, U., Slosar, A., and Vazquez Gonzalez, J. (2017). Galaxy-galaxy lensing estimators and their covariance properties. *MNRAS*, 471:3827–3844.

Spergel, D. N., Verde, L., Peiris, H. V., et al. (2003). First-Year Wilkinson Microwave Anisotropy Probe WMAP Observations: Determination of Cosmological Parameters. *ApJS*, 148:175–194.

Springel, V., Frenk, C. S., and White, S. D. M. (2006). The large-scale structure of the Universe. *Nature*, 440:1137–1144.

Springel, V., White, S. D. M., Jenkins, A., et al. (2005). Simulations of the formation evolution and clustering of galaxies and quasars. *Nature*, 435:629–636.

Springel, V., White, S. D. M., Tormen, G., and Kauffmann, G. (2001). Populating a cluster of galaxies – I. Results at $z = 0$. *MNRAS*, 328:726–750.

Squires, G. and Kaiser, N. (1996). Unbiased Cluster Lens Reconstruction. *ApJ*, 473:65–80.

Stoughton, C., Lupton, R. H., Bernardi, M., et al. (2002). Sloan Digital Sky Survey: Early Data Release. *AJ*, 123:485–548.

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., and Oguri, M. (2012). Revising the Halofit Model for the Nonlinear Matter Power Spectrum. *ApJ*, 761:152–162.

Tanimura, H., Hinshaw, G., McCarthy, I. G., et al. (2019). A search for warm/hot gas filaments between pairs of SDSS Luminous Red Galaxies. *MNRAS*, 483:223–234.

Taylor, A. N., Kitching, T. D., Bacon, D. J., and Heavens, A. F. (2007). Probing dark energy with the shear-ratio geometric test. *MNRAS*, 374:1377–1403.

Taylor, J. E., Massey, R. J., Leauthaud, A., et al. (2012). Measuring the Geometry of the Universe from Weak Gravitational Lensing behind Galaxy Groups in the HST COSMOS Survey. *ApJ*, 749:127–139.

Tegmark, M., Blanton, M. R., Strauss, M. A., et al. (2004). The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey. *ApJ*, 606:702–740.

Thiele, L., Duncan, C. A. J., and Alonso, D. (2020). Disentangling magnification in combined shear-clustering analyses. *MNRAS*, 491:1746–1758.

Troxel, M. A. and Ishak, M. (2015). The intrinsic alignment of galaxies and its impact on weak gravitational lensing in an era of precision cosmology. *Phys. Rep.*, 558:1–59.

Troxel, M. A., MacCrann, N., Zuntz, J., et al. (2018). Dark Energy Survey Year 1 results: Cosmological constraints from cosmic shear. *Phys. Rev. D*, 98:043528.

Tseliakhovich, D. and Hirata, C. (2010). Relative velocity of dark matter and baryonic fluids and the formation of the first structures. *Phys. Rev. D*, 82:083520.

Tyson, J. A., Valdes, F., and Wenk, R. A. (1990). Detection of Systematic Gravitational Lens Galaxy Image Alignments: Mapping Dark Matter in Galaxy Clusters. *ApJ*, 349:L1.

Unruh, S., Schneider, P., and Hilbert, S. (2019). Magnification bias in the shear-ratio test: a viable mitigation strategy. *A&A*, 623:A94.

van Daalen, M. P., Schaye, J., McCarthy, I. G., Booth, C. M., and Dalla Vecchia, C. (2014). The impact of baryonic processes on the two-point correlation functions of galaxies subhaloes and matter. *MNRAS*, 440:2997–3010.

van Uitert, E., Joachimi, B., Joudaki, S., et al. (2018). KiDS+GAMA: cosmology constraints from a joint analysis of cosmic shear galaxy-galaxy lensing and angular clustering. *MNRAS*, 476:4662–4689.

Vázquez, J. A., Padilla, L. E., and Matos, T. (2018). Inflationary Cosmology: From Theory to Observations. *arXiv e-prints*, arXiv:1810.09934.

Vogelsberger, M., Genel, S., Springel, V., et al. (2014). Properties of galaxies reproduced by a hydrodynamic simulation. *Nature*, 509:177–182.

Voigt, L. M. and Bridle, S. L. (2010). Limitations of model-fitting methods for lensing shear estimation. *MNRAS*, 404:458–467.

Walker, A. G. (1937). On Milnes Theory of World-Structure. *Proceedings of the London Mathematical Society*, 42:90–127.

Wang, X. and Pen, U.-L. (2019). Understanding the Reconstruction of the Biased Tracer. *ApJ*, 870:116–125.

Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. (2019). H0LiCOW XIII. A 2.4% measurement of H_0 from lensed quasars: 5.3σ tension between early and late-Universe probes. *arXiv e-prints*, arXiv:1907.04869.

Wright, A. H., Hildebrandt, H., van den Busch, J. L., and Heymans, C. (2020). Photometric redshift calibration with self-organising maps. *A&A*, 637:A100.

Wright, C. O. and Brainerd, T. G. (1999). Gravitational Lensing by NFW Halos. *arXiv e-prints*, astro–ph/9908213.

Zahid, H. J., Geller, M. J., Fabricant, D. G., and Hwang, H. S. (2016). The Scaling of Stellar Mass and Central Stellar Velocity Dispersion for Quiescent Galaxies at $z < 0.7$. *ApJ*, 832:203–217.

Zeldovich, Y. B. (1972). A hypothesis unifying the structure and the entropy of the Universe. *MNRAS*, 160:1–3.

Zhao, C., Chuang, C.-H., Kitaura, F.-S., et al. (2020). Improving baryon acoustic oscillation measurement with the combination of cosmic voids and galaxies. *MNRAS*, 491:4554–4572.

Zhu, H.-M., Yu, Y., and Pen, U.-L. (2018). Nonlinear reconstruction of redshift space distortions. *Phys. Rev. D*, 97:043502.

Ziour, R. and Hui, L. (2008). Magnification bias corrections to galaxy-lensing cross-correlations. *Phys. Rev. D*, 78:123517.

Zuntz, J., Sheldon, E., Samuroff, S., et al. (2018). Dark Energy Survey Year 1 results: weak lensing shape catalogues. *MNRAS*, 481:1149–1182.

Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae. *ApJ*, 86:217–246.

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The completion of my PhD project took me several years. During this time, many people made big and small contributions to my work – whether related to the science itself or my mental well-being. I will forever be grateful that you took this journey with me.

The greatest contribution to this thesis was undoubtedly from my supervisor Prof. Dr. Peter Schneider. I thank you wholeheartedly for the countless hours of discussion in your office, for all the stupid and not so stupid questions you answered, and for your guidance through the academic life. About 10 years went by from attending my first lecture with you in my 2nd Bachelor semester to the end of my PhD project. It was quite an adventure and I am very happy that you stuck with me. I further want to thank Prof. Dr. Cristiano Porciani for being the second reviewer of my thesis. You also agreed to be part of my thesis advisory committee meetings (TAC), where you as the only non-German proved to know the Promotionsordnung by heart. The other members of TAC where Prof. Dr. Peter Schneider, Dr. Patrick Simon, and Dr. Norbert Wex – thank you for ordering my thoughts and my project.

During the past three years, I met a lot of people in the Argelander-Institute for Astronomy (AIfA). Some of them made my work not only productive but also fun. I will always be indebted to everyone who took the time to discuss life, the Universe and everything. Thanks to Bastian Orthen I probably have more coffee running through my veins than blood. I can only thank you for listening to all my ramblings about the codes that decided to not work (which was certainly not my fault). Also, Sven Heydenreich had the patience to listen to the details of my project, especially when the codes were acting up again. Thanks to you, some of the coffee has been replaced with black tea, and our discussions (and bets) made my life so much easier (not that I managed to win a lot of bets). In the last year, we also met weekly with Pierre Burger and Rongchuan Zhao to discuss our progress together, this helped more than all the coffee and black tea combined. Thank you Pierre, although we started off by you calling me in my well-deserved vacation, I had so much fun discussing projects inside and outside academia in the past two years.

I also want to express my sincere thanks to Dr. Jens Erler. Not only are you the living proof that night owls are as productive as these day-time-lovers, you also filled the AIfA-life with fun, food and expertise. You and Sven Heydenreich proofread most of my thesis in your free time, without your help I could not be as happy with the final result. Right behind Jens' neatly ordered desk, I found one of my best friends, Ana Mikler. We must have spent hundreds of hours working together and I cannot really put in words how much your company helped me, especially in the last few months, in completing this thesis. The WebCoffee, we established in the self-quarantine time, with you, Jens and Maude Charmetant was always fun and inexplicably void of actual coffee. Thanks to Kim Werner, tutoring with you was fun, but sitting on the empty desk in your office and discussing cosmology, science and life were even better. Also thank you, Sabine Derdau, you not only simplified the bureaucratic academia for me, you also made it hard to leave your office without a good conversation.

I want to thank Fatimah Raihan for laughing with me; Dr. Joey Martinez for the best barbecues; Dr. Angus Wright for your upside-down humour; David (Dato) Alkhanishvili for not killing me when I ate those dumplings with a fork; Marcus Keil for organizing the best game nights; Hannah Zohren for agreeing that it is never too warm for a blanket; David Ohse for surprise cake; Cosima Eibensteiner and Jakob den Brok for combining European politics with coffee; Victoria Yankelevich for years of fun; Sandra Martin for long discussions; Joseph Kuruvilla for all the laughs; Béibhinn Whelan for adding Irish tea to my coffee addiction; and Melanie Hampel for motivation from the other side of the globe. Thanks to my great office mates Zeinab Shafiee, Zahra Sheikhbahaee and Rongchuan Zhao for the greenest office in the institute. In September 2018, I helped to found the monthly outreach event 'Astronomy on Tap Bonn' and every single event was awesome. Therefore, I want to thank Dr. Eleni Vardoulaki, Laila Linke, Dr. Enrico Garaldi, Ana Mikler, Dr. Aarti Nagarajan, Toma Badescu, Manali Jeste, Maude Charmetant, Felix Plötzl, Kevin Harrington and the Fiddlers Pub team. Special thanks to Diana Scognamiglio for not minding our regular office invasions, I consider you part of the AoT team.

I further want to acknowledge my financial support. Thanks to the *Studienstiftung des deutschen Volkes* I received a full scholarship that let me fully focus on this project. I also received support from the *Bonn Cologne Graduate School* (BCGS) within their H2 program, as well as the *International Max Planck Research School* (IMPRS). It is thanks to the support from these two institutions that I have been able to attend so many different workshops and conferences in Bonn and abroad. A special thank you goes to Dr. Rainer Mauersberger and Dr. Simone Pott for not only leading the IMPRS as a science institution but also caring about the physical and mental well-being of all of us.

In my project I coded mainly in `C`, `C++` and `Python`. Thank you `stackoverflow` and `stackexchange` for providing solutions to all kinds of issues. Also, the creators of the Python module `astropy` receive my special thanks. The grammar of my

articles and this thesis received a quality boost from the free internet tools: `linguee`, `deepl` and `thesaurus`.

Diese Arbeit wäre nicht möglich gewesen ohne die Unterstützung meiner Familie, die mich mit Liebe, Geduld und Humor durch die letzten Jahre begleitet haben. Den meisten Dank verdient mein Freund Jens Derksen, der mit grenzenloser Energie und Enthusiamus unsere gemeinsame Zeit gestaltet. Dank unseren wunderbaren Kindern Verena und Leonard ist es jeder Tag voll mit Kreativität, Kuscheln und Chaos. Danke auch an meine Eltern Simone und Ronald Unruh für lange Abende, Motivation im Alltag und frühzeitige Planungen von überraschenden Jahrestagen wie Ostern, Weihnachten und Geburtstagen. Danke an meine kleine Schwester Franziska Unruh, dass du die Ferien am liebsten bei uns verbringst und uns regelmäßig dein Zimmer (fast) freiwillig überlässt. Danke an meine Großeltern Ingrid und Manfred Frobenius, die auch noch mit 80+ regelmäßig mit dem Zug nach Bonn kommen um mich zu besuchen und mir etwas Arbeit abzunehmen. Ich bin ganz sicher, dass euch die Deutsche Bahn eines Tages ohne Zugausfälle oder Umleitungen wieder nach Hause bringen wird. Danke auch an meine Großeltern Waltraud und Hermann Unruh aus Hennigsdorf (oder sollte ich besser Lindow sagen?). Ihr gebt mir den Rückzugsort, den ich jeden Sommer gebraucht habe und auch brauchen werde. Danke an meine ganze Familie für die Liebe und den Rückhalt. Die letzten Jahre waren spannend, kurzweilig und faszinierend, nicht zuletzte dank der verrückten und lustigen Ideen von Lena Schmoll, Alexandra Philos, Fabian Amrath, Maren Lutz, Nico Lutz, Dominique Neumann, Paul Konieczny, Cora Neikes, Norman Lorenz, Mareike Schell, Tobias Isken, Yannik Völzke, Dr. Sascha Agne und Sahra Frank.

Although I could go further and tell more stories, I will stop here and look forward to the journey ahead with all of you. If I had the choice, I would start all over again. My final thoughts go to my family and friends who passed away during my studies; Julia Hampel, Hermann Unruh und Hansi Dietrich – you will be missed.