DISS. ETH NO. 0000

PROBING DARK MATTER WITH STRONG GRAVITATIONAL LENSING

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

SIMON BIRRER

Master of Science in Physics

ETH Zurich

born on 3rd of February 1988

citizen of SWITZERLAND

accepted on the recommendation of

Prof. Dr. Alexandre Refregier

Dr. Adam Amara

Dr. Sherry Suyu

2016

Abstract

The standard cosmological model is based on Einstein's theory of general relativity, dark matter, dark energy and inflation. To date, the physical origin of several of these ingredients remains a mystery. The aim of this thesis is to test this model by studying the nature of dark matter through its effects on small cosmological scales.

As a first step, we construct a phenomenological description of the relationship between dark matter halos and the normal visible matter in galaxies. This is done by incorporating a prescription of baryonic physics into a dark matter merger tree.

We then probe the physical properties of dark matter on sub-galactic scales using strong gravitational lensing. We develop a new method based on versatile basis sets to accurately model complex lensing and source structures. For the strong lens system RXJ1131-1231 this allows us to reconstruct a reliable high resolution image of the source.

We also develop a new treatment of the mass-sheet degeneracy inherent in gravitational lensing. We use this to revisit the estimate of the Hubble constant from time-delay strong lensing. We find that the Hubble parameter H_0 estimate of the lens system RXJ1131-1231 is strongly sensitive to the dynamical model and the source scale prior and that it is consistent with current cosmic microwave background experiments.

Refining our model further, we develop a new approximation scheme that enables an effective treatment of line-of-sight structures. This allows us to further constrain the connection between dark and visible matter. In addition, it yields constraints on the external shear around strong lensing system to sub-percent level precision.

To probe the small scale structure formation and the nature of dark matter, we analyze the substructure content of strong lenses statistically through extensive forward modeling. We find that the abundance of substructure in the strong lens RXJ1131-1231 is consistent with a

cold dark matter particle scenario. A lower bound on the free streaming mass of a possible dark matter particle can be set to about 2keV.

The precision gain by these new techniques and results shed new light on the distribution of dark matter, its physical nature and its relation to visible matter. These results, combined with the upcoming high resolution instruments, offer great prospects for the high-precision study of dark matter and other sectors of the standard cosmological model with strong lensing.

Zusammenfassung

Das Standardmodel der Kosmologie basiert auf Einsteins Allgemeiner Relativitätstheory, Dunkler Materie, Dunkler Energie und der Inflation. Bis heute bleibt die physikalische Natur mehrerer dieser Bestandteile ein Rätsel. Das Ziel dieser Doktorarbeit ist es, dieses Modell mit Hilfe des starken Gravitationslinseneffekts auf kleinen kosmologischen Skalen zu testen.

In einem ersten Schritt erstellen wir eine phänomenologische Beziehung zwischen Dunkler Materie und der sichtbaren Materie in Galaxien. Dies wird durch das Zusammenfügen von Rezepten zur Beschreibung von baryonischer Physik und der hierarchischen Strukturbildung von Dunkler Materie realisiert.

Wir testen dann die physikalischen Eigenschaften von Dunkler Materie in Sub-galaktischen Skalen mittels dem starken Gravitationslinseneffekt. Wir entwickeln eine neue Methode basierend auf vielseitigen Basisfunktionen zum akkuraten modellieren von komplexen Linsen und Quellenstrukturen. Für die Linse RXJ1131-1231 erlaubt uns dies eine verlässliche, hochauflösende Rekonstruktion der Quelle.

Wir entwickeln auch ein neues Verfahren für die Mass-sheet Degeneration. Wir benutzen dieses zur Reanalyse der Messung der Hubble Konstante basierend auf Zeitverzögerungsmessungen in Gravitationslinsen. Wir finden, dass die Schätzung des Hubble Parameters H_0 für die Linse RXJ1131-1231 stark vom dynamischen Modell der Linse und vom Prior der Lengenskalen in der Quelle abhängt und dass die Messung mit den neuesten Experimente der Kosmischen Hintergrundstrahlung überein stimmt.

Weiter entwickeln wir eine neues Approximationsverfahren das eine effektive Beschreibung von zusätzlichen Effekten entlang des Lichtweges. Dies erlaubt uns weitere Einblicke in die Beziehung zwischen sichtbarer und unsichtbarer Masse. Zusätzlich führt dies zu einer Bestimmung des Scherungseffekts mit sub-prozent Genauigkeit.

Für die Erforschung der Strukturbildung auf kleinen Skalen und der Dunklen Materie analysieren wir die Substruktur in Gravitationslinsen in einem statistischen Verfahren basiert auf extensiven Simulationen. Wir finden, dass die Häufigkeit von Substrukturen in der Linse RXJ1131-1231 konsistent mit dem Scenario der kalten Dunklematerieteilchen ist. Wir können eine untere Schranke auf die freie Ausbreitungslänge und der assoziierten Masse eines Teilchens mit 2keV angeben.

Der Gewinn an Präzission mit den neuen Verfahren und denen daraus hervorgehenden Resultate haben neue Erkenntnisse in die Verteilung von Dunkler Materie, deren Natur und der Beziehung zu der sichtbaren Materie hervorgebracht. Diese Resultate, in Kombination mit den kommenden hoch auflösbaren Instrumenten, ermöglichen gute Aussichten für sehr präzise Studien der Dunklen Materie und anderer Gebiete des Standard Models der Kosmologie mit dem starken Gravitationslinseneffekt.

Contents

Contents

Contents

1 Introduction

The origin and nature of the universe has been pondered upon throughout human history. Remarkable progress in science has led to the development of modern cosmology, which is able to address these fundamental questions in a quantitative and rigorous manner.

Over the last century, the standard model of cosmology has become established and is based on the following pillars: (1) The cosmological principle which states that the universe is statistically homogeneous and isotropic on large scales. (2) General relativity, a theory of gravity that describes the geometry and time-evolution of the universe. (3) Normal matter coming from the standard model of particle physics. (4) Cold dark matter (CDM), an invisible non-interacting form of matter. (5) Dark energy (Λ) that causes the recent acceleration of the universe and (6) Inflation that provides the seeds of the structures in the universe.

In this standard ΛCDM model, the physical nature of several components remains unknown. In particular, the search for the understanding of dark matter and dark energy is one of the most pressing endeavors in fundamental physics.

Evidence for the existence of dark matter and dark energy comes from multiple probes, such as the cosmic microwave background (CMB) temperature anisotropies [1], galaxy clustering [2], weak gravitational lensing [3; 4], supernovae of type Ia (e.g. [5; 6]) and clusters [7]. The smallest scale phenomenological tests come from the Lyman-alpha forest (see e.g., [8; 9]). All these measurements are consistent with a cold and only gravitationally interacting particle.

At sub-galactic scales, there are observational and theoretical challenges in bringing model and data into agreement. Discrepancies have been reported predominantly in the number, phase space densities and density profiles when comparing simulations of dark matter substructure with observations of luminous satellite galaxies in our Milky Way (MW) (see e.g., 10; 11; 12; 13; 14) and a recent review of (15). There also remain open questions about the inner slope in galaxy clusters (e.g. 16).

A potential non-gravitational nature of a dark matter particle may have an effect on structure formation on small scales without having an impact on the well tested large scales. Probing the small scale structure formation and mass distribution may thus provide information beyond the CDM model.

Warm dark matter (WDM), a thermal relic with a free streaming length, has been proposed to solve the discrepancies in the abundance of small scale structure (17; 18). Another approach is a self-interacting dark matter particle (19). The latest constraints from the Lyman-α forest (20) results in a lower bound on the thermal relic mass $m_{TH} = 3.3$keV at the 2σ confidence level. The limits obtained from dwarf galaxy counts disfavors particle masses below $m_{TH} = 2.3$keV at the 2σ confidence level(21; 22).

Advances have also been made in bringing the observations and theory predictions together within the ΛCDM paradigm. On the one hand, more detailed simulations incorporating the effects of baryons result in a suppression of expected luminous structures with significantly changed internal structure. On the other hand, a wealth of new satellite galaxies in our MW have been detected by the Dark Energy Survey (23; 24).

The very dense and dark matter dominated MW satellites provide a laboratory to search for dark matter annihilation signals. In particular, the latest Fermi Large Area Telescope analysis (25; 26) report no detection of gamma-ray excess. If the estimated DM content of these dark matter dominated satellite galaxy candidates is confirmed, they will constrain the annihilation cross section to lie below the thermal relic cross section for DM particles with masses $\lesssim 20$ GeV annihilating via the $b\bar{b}$ or $\tau\bar{\tau}$ channels.

Direct detection experiments have so far not reported a persistent signal detected by several independent experiments (see e.g. 27, for a recent review and references therein). The absence

of direct and indirect detections stimulates further searches. A possible quantification of a non-CDM behavior in the small scale structure formation may lead to a more targeted search in the direct detection experiments.

Gravitational lensing, the effect of light propagation through perturbed space-time, is a direct probe of the underlying energy distribution and its geometry. Gravitational lensing provides a probe of the invisible universe that can test the predictions of ΛCDM.

The aim of this thesis is to probe structure at small cosmological scales to constrain potential non-gravitational properties of dark matter with gravitational lensing. For this purpose, we develop new methods to infer lensing properties through forward modeling.

The thesis is structured as follows: In this chapter 1, we introduce the basic theoretical concepts of ΛCDM, structure formation and the theory of gravitational lensing. Chapter 2 presents an empirical model that links dark matter structure formation with galaxy evolution. In Chapter 3, we review the current state of lens modeling and image reconstruction and introduce a new framework involving basis sets. In Chapter 4, we present a method to infer the impact of line-of-sight structure of strong gravitational lenses and combine the techniques of chapter 3 and 2 on data. Chapter 5 probes the effect of dark energy with time delay cosmography. In particular, the effect of an important degeneracy in lensing on the cosmological inference is explored. Chapter 6 presents a new statistical approach to quantify the substructure content in strong gravitational lenses that is able to probe the effect of different dark matter models at small cosmological scales. In Chapter 7, we summarize the thesis, draw conclusions and point on implications and prospects for future work.

Most of the material in this thesis has been assembled from published work (28; 29; 30) and submitted work (31), and we give the corresponding references at the beginning of each chapter. The work presented in chapter 6 will appear in a similar form in a future publication.

1.1   Homogeneous and Isotropic Universe

The theoretical foundations of ΛCDM are well known and are covered in detail in many standard cosmological textbooks such as (32; 33). In this section, we discuss the dynamics of

a homogeneous and isotropic universe governed by Einstein's Theory of General Relativity (GR) and introduce some notation. We mainly follow (32) and give further references where needed.

GR relates the geometry of space-time to its energy content through the *Einstein field equations*

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.1}$$

where $G_{\mu\nu}$ is the *Einstein tensor*, which is a functional of the *metric tensor* $g_{\mu\nu}$, $T_{\mu\nu}$ is the Stress-energy tensor, G is Newton's gravitational constant and c is the speed of light.

The cosmological principle states that, on large scales, the universe is statistically homogeneous and isotropic. The most general metric that describes such an universe, and allows for a time dependence, is the Friedmann-Lemaître-Robertson-Walker (FLRW) universe (34; 35; 36), given by

$$ds^2 = g_{\mu\nu}^{\text{FLRW}} dx^\mu dx^\nu = c^2 dt^2 - a(t)^2 d\Sigma^2 \tag{1.2}$$

where Σ denotes coordinates over a 3-dimensional space of uniform curvature and $a(t)$ is a time-dependent *scale factor* of the spacial part of the metric.

In hyperspherical or curvature-normalized coordinates, the spatial part of the metric can be written as

$$d\Sigma^2 = d\chi^2 + S_k(\chi)^2 d\Omega^2 \tag{1.3}$$

where χ is the time-independent *co-moving* radial distance and $d\Omega^2$ is the angular part of the metric

$$d\Omega^2 = d\theta^2 + \sin^2\theta \, d\phi^2 \tag{1.4}$$

and

$$S_k(\chi) = \begin{cases} \sqrt{k}^{-1} \sin(\chi\sqrt{k}), & k > 0 \\ \chi, & k = 0 \\ \sqrt{|k|}^{-1} \sinh(\chi\sqrt{|k|}), & k < 0, \end{cases} \tag{1.5}$$

where k specifies the curvature of space (i.e. $k > 0$ is a closed, $k = 0$ flat and $k < 0$ an open universe). The metric is invariant under transformations that rescale a and r such that $\frac{a}{r}$ remains constant. Conventionally, one uses this freedom to fix the scale factor at present time to $a_0 = 1$.

Substituting the FLRW metric into Einstein's equations (1.1), one obtains the Friedmann equations which describe the evolution of the scale factor a for a universe containing an ideal fluid with energy density ρ, pressure p (in units of density, i.e. $/c^2$) and curvature k:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \tag{1.6}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p). \tag{1.7}$$

The fluids are often described by their equation of state parameter w through

$$p = w\rho. \tag{1.8}$$

The conservation of the energy-momentum tensor directly implies that the energy density of a homogeneous fluid with equation of state parameter w scales with the scale factor a as

$$\rho \propto a^{-3(1+w)}. \tag{1.9}$$

Relativistic fluids such as photons have an equation of state parameter of $w = \frac{1}{3}$ and hence their energy density falls off as $\rho_R \propto a^{-4}$. Non-relativistic fluids such as dark matter obey $w = 0$ and therefore dilute at a slower rate with $\rho_M \propto a^{-3}$.

Finally, a cosmological constant can be described as an ideal fluid with $w = -1$ and a constant

energy density ρ_Λ.

As a ΛCDM universe contains all of the aforementioned species, the possible solutions for the scale factor are parametrized by the energy density of each of the components at the present time. For convenience, the energy densities are often given in units of the critical density ρ_c which is simply given by solving equation (1.6) for ρ in a flat universe with $k = 0$:

$$\rho_c = \frac{3H^2}{8\pi G}.\tag{1.10}$$

The density parameter Ω is then defined in terms to the energy density ρ variable as:

$$\Omega \equiv \frac{\rho}{\rho_c}.\tag{1.11}$$

For a species X with equation of state parameter w_X, one can now parametrize its energy density ρ_X by its present day value at $a = 1$ and write it as a function of scale factor as:

$$\rho_X(a) = \rho_X(a = 1)a^{-3(1+w_X)}.\tag{1.12}$$

Denoting the *Hubble parameter* today as H_0, one can rewrite the first Friedmann equation (1.6) for ΛCDM as:

$$\frac{H^2}{H_0^2} = \Omega_M a^{-3} + \Omega_R a^{-4} + \Omega_\Lambda + \Omega_k a^{-2}\tag{1.13}$$

where the curvature terms is written as a curvature "density" $\Omega_k = 1 - \Omega$.

An important consequence of a time dependent FLRW metric is that light propagation changes the observed wavelength λ_{obs} relative to the emitted wavelength λ_{emit}. Light travels along null-geodesics, i.e $ds^2 = 0$. The FLRW metric implies that

$$\frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} = \frac{a_{\text{obs}}}{a_{\text{emit}}}.\tag{1.14}$$

With the convention of $a_{\text{obs}} = 1$, one can rewrite Equation 1.14 as

$$\lambda_{obs} = \frac{\lambda_{\text{emit}}}{a} \equiv (1+z)\lambda_{\text{emit}}, \tag{1.15}$$

where the definition of redshift $z = \frac{1}{a} - 1$ ranging from 0 today to infinity at $a = 0$ is introduced. The effect of the cosmological redshift leads to an observable distance-redshift relation.

The co-moving radial distance χ which appears in the FLRW metric (1.3) of a source at scale factor a that is observable today at $t = t_0$ can be expressed as:

$$\chi = \int_{t(a)}^{t_0} \frac{dt'}{a(t')} = \int_a^1 \frac{da'}{a'^2 H(a')}. \tag{1.16}$$

The first Friedmann equation (1.6) was used to transform the integral over time into an integral over the scale factor.

Cosmological distances are not directly observable but have direct consequences on the angular size and the brightness of a distant source. The *transverse co-moving distance T* is defined such that two object at equal redshift z separated by an angel $\delta\theta$ have a radial co-moving separation $d = \delta\theta T$. In a flat universe, the radial and transverse co-moving distances are identical. In a general FLRW metric, from the spacial part of the metric (Equation 1.5), T is given by $S_k(\chi)$.

The *angular diameter distance d_A* relates the angular separation $\delta\theta$ between two objects at the same redshift to their physical separation. The angular diameter distance can be expressed in terms of the transverse co-moving distance as

$$d_A(z) = \frac{T(z)}{1+z}. \tag{1.17}$$

Geometrical probes of cosmology that make use of a known physical size and determine the angular diameter distance (or relations thereof) are known as *Standard Rulers* . The most powerful probe within this class is the CMB (1). Further probes are the baryonic acoustic oscillation (BAO) signal in the galaxy clustering (37) and time-delay cosmography (see e.g. 38, for latest results).

The *Luminosity distance* d_L is the distance derived from flux measurements F relative to a (known) intrinsic luminosity L of an object. The distance d_L is defined such that $4\pi F d_L{}^2 = L$ is satisfied. The Luminosity distance can be related to the transverse comoving distance as

$$d_L(z) = (1 + z)\, T(z). \tag{1.18}$$

Geometrical probes of cosmology that make use of a known brightness of a physical object to determine the luminosity distance are known as *Standard Candles*. One prominent example are supernovae of type Ia (e.g. 5; 6). Cepheid variable stars recently became a competing measurement of the late time expansion (see 39).

The maximum distance light can travel radially since $t = 0$ defines the *co-moving horizon* or *conformal time* η:

$$\eta(t) = \int_0^t \frac{dt'}{a(t')}. \tag{1.19}$$

This distance defines the physical extent of our observable universe.

1.2 Cosmological structure formation

The previous section discussed the homogeneous universe. The universe today is however inhomogeneous on small scales as observed on the scale of galaxy clusters, galaxies and down to scales of stars and planets. These non-linear structure emerged from an originally smooth linear density field. We will first revise the linear perturbation theory approach to structure formation and then discuss formalisms to capture the non-linear formation of structure.

1.2.1 Linear growth of structure

Density fluctuations can be described as a relative density contrast

$$\delta(\vec{x}) \equiv \frac{\rho(\vec{x}) - \bar\rho}{\bar\rho}, \tag{1.20}$$

where $\rho(\vec{x})$ is the density at position $\vec{x}$ and $\bar{\rho}$ is the mean density of the universe. A convenient description is the density contrast in Fourier space $\delta(\vec{k})$

$$\delta(\vec{k}, \eta) = \frac{1}{\pi^{3/2}} \int \delta(\vec{x}, \eta) e^{-\vec{k}\cdot\vec{x}} d\vec{x}, \tag{1.21}$$

where $\vec{k}$ is the wave vector and η is the conformal time. The Einstein-Boltzmann equations describe the evolution of the Fourier modes in linear perturbation theory. They describe the interaction of small density and pressure perturbations of different species and their evolution, growth and decay. This is beyond the scope of this introduction and can be read in standard textbooks, such as (32). The solution of the time dependent density fluctuations $\delta(k, \eta)$ at late times, in the matter or dark energy dominated era, can approximately be separated into a k and η dependent function

$$\delta(k, \eta) \propto k^2 \Phi_p(k) T(k) D(a), \tag{1.22}$$

where $\Phi_p(k)$ is the gravitational potential at the initial condition set by inflation, $T(k)$ is the *transfer function*, and $D(a)$ is the *growth function* quantifying the time dependence of the evolving density field.

The initial conditions $\Phi_p(k)$ are required to predict any form of structure in the universe. They are thought to be produced during inflation. Inflation itself does not predict specific values of $\Phi_p(k_i)$. A common prediction of inflationary models is a Gaussian random field with a 2pt function

$$\langle \Phi_p(k_i) \Phi_p(k_i') \rangle = (2\pi)^3 P_\Phi(k) \delta^3(k_i - k_i'), \tag{1.23}$$

where the primordial power spectrum P_Φ can be parametrized as

$$\frac{k^3}{2\pi^2} P_\Phi(k) = A_s \left(\frac{k}{k_0} \right)^{n_s - 1}. \tag{1.24}$$

A_s is the amplitude of the primordial power spectrum at the reference scale k_0 and n_s is the

power law slope. A_s and n_s are free parameters that have to be inferred from observations.

The transfer function $T(k)$ describes how scales evolve through horizon crossing. Models outside the horizon ($\eta k \ll 1$) stay mostly frozen and are set by the initial conditions.

During matter domination, all sub-horizon modes grow with the same rate. The linear growth function is given by (see e.g. (32) for derivation)

$$D(a) \propto H(a) \int_0^a \frac{da'}{(a' H(a'))^3}. \tag{1.25}$$

The matter power spectrum predicted by linear perturbation theory is then given by

$$\langle \delta(k_i, \eta) \delta(k_i', \eta) \rangle = (2\pi)^3 P(k, a) \delta^3(k_i - k_i'), \tag{1.26}$$

where $P(k, a)$ is given by combining equations (1.22) and (1.26):

$$P(k, a) = \frac{2\pi^2}{k_0} A_s \left(\frac{k}{k_0} \right)^{n_s} [T(k) D(a)]^2. \tag{1.27}$$

The most prominent probes of the large scale structure in the universe at late times are galaxy clustering (2) and weak gravitational lensing (3; 4).

As δ grows over time and the variance in the density field increases accordingly, at one point, the structure becomes non-linear and collapsed over-dense objects form, which are called halos. These objects are the seeds for galaxies and clusters and lead to a hierarchical scenario for the formation of structure.

In the following, we revise the basic theory that enables a description of non-linear structures.

1.2.2 Spherical collapse model

General non-linear solutions of structure formation can not be described analytically and require extensive numerical methods and calculations. There are however specific configurations that have analytic solutions. One of them is the spherical collapse. The spherical collapse model, although it relies on simplistic assumptions, has been proven to provide a reasonable

good qualitative prediction of structure formation. In fact, extensions of its simplest form can give reliable quantitative predictions without relying on numerical N-body simulations. The solution of the non-linear spherical collapse model in a matter dominated universe can be computed as follow:

The Friedmann equation in a matter dominated and closed universe is given by

$$\frac{H^2}{H_0^2} = \Omega_M a^{-3} + (1 - \Omega_M) a^{-2}. \tag{1.28}$$

The solution of equation (1.28) for the time dependent radial extent r of a closed universe can be stated in terms of the parameter $\theta = H_0 \eta (\Omega_M - 1)^{1/2}$ as

$$r(\theta) = A(1 - \cos\theta) \tag{1.29}$$

$$t(\theta) = B(\theta - \sin\theta), \tag{1.30}$$

where $A = r_0 \Omega_M / 2(\Omega_M - 1)$ and $B = H_0^{-1} \Omega_M / 2(\Omega_M - 1)^{3/2}$ are the initial conditions. The turn around happens at $\theta = \pi$ with $r = 2A$ and $t = \pi B$. Collapse $(r \to 0)$ occurs at $\theta = 2\pi$ with $t = 2\pi B$.

The correspondence to a local perturbation in a background cosmological model with $\Omega_M = 1$ can be achieved by

$$A = \frac{1}{2H_0 \Omega_M^{1/2}} \left(\frac{3}{5} \frac{a_i}{\delta_i} \right)^{3/2} \tag{1.31}$$

$$B = \frac{3}{10} \frac{r_i}{\delta_i}. \tag{1.32}$$

The initial conditions are given as a homogeneous over-density δ_i with radius r_i at scale factor a_i. At collapse $(\theta = 2\pi)$, the scale factor a_{col} is

$$a_{\text{col}} = \left(\frac{3}{4} \right)^{2/3} \left(\frac{3}{5} \frac{a_i}{\delta_i} \right) (2\pi)^{2/3} \approx 1.686 \frac{a_i}{\delta_i}. \tag{1.33}$$

This means that the perturbation collapsed when the linear theory predicts a density contrast of $\delta_c = 1.686$. The collapse does not proceed to a point $(r \to 0)$ but relaxes to virial equilibrium.

The *virial theorem* describes the partition of potential energy E_{pot} and kinetic energy E_{kin} as

$$E_{pot} = -2E_{kin}. \tag{1.34}$$

The total energy of the system E_{tot} is given by the potential energy at the turn-around radius r_{max}. Applying the virial theorem

$$E_{tot} = E_{pot} + E_{kin} = E_{pot}(r_{max}) = \frac{1}{2}E_{pot}(r_{vir}) \tag{1.35}$$

leads to $r_{vir} = \frac{1}{2}r_{max}$ since $E_{pot} \propto r^{-1}$. The virial radius is reached by $\theta = \frac{3}{2}\pi$ and the over-density at collapse in such a situation relative to an ever expanding critical universe is

$$\frac{\rho(\theta = 3\pi/2)}{\bar{\rho}(\theta = 2\pi)} = 18\pi^2 \approx 178. \tag{1.36}$$

The specific value of the density contrast of a collapsed halo is only valid for a matter dominated universe. The density contrast Δ_ν is often used to define a collapsed object. Equivalently, the virial mass of an object M_ν is defined as

$$M_\nu = \frac{4\pi}{3}r_\nu^3 \rho_M \Delta_\nu. \tag{1.37}$$

The model provides an approximation on how and when over-densities collapse. The underlying assumptions are however simplified. In reality, the collapse is not spherical and shell crossing can occur when δ_i is not monotonically decreasing with radius. Furthermore, in a universe with dark energy driving the expansion but not participating in the collapse, one can not consider spherical collapse to be a separate universe. Nevertheless, the description above allows a qualitative assessment of the non-linear structure formation and more detailed models are built by extensions of the spherical collapse model.

1.2.3 Press-Schechter formalism

One way to quantify the structure in the universe is by counting the number of collapsed objects as a function of their mass, the so called *halo mass function* (HMF). A theoretical

framework to predict the HMF from the statistics of the linear density field was introduced by Press and Schechter (40). This framework relies on smoothed density fields and the spherical collapse model described in section 1.2.2.

Given a density field $\delta(\vec{x})$, one can apply a filter $W(\vec{x}; R)$ with scale R to get a smoothed density field

$$\delta(\vec{x}; R) = \int \delta(\vec{x}')W(\vec{x} + \vec{x}'; R)d\vec{x}'. \tag{1.38}$$

For each filter, one can define a corresponding mass $M = \gamma_f \bar{\rho} R^3$, where γ_f is a constant depending on the shape of the filter. Smoothing can also be applied in Fourier space accordingly.

One can then define the mass variance of a smooth density field

$$\sigma^2(R) = \langle \delta^2(\vec{x}; R) \rangle = \frac{1}{(2\pi)^3} \int P(k)\widetilde{W}^2(kR)d^3k, \tag{1.39}$$

where $\widetilde{W}^2(kR)$ is the Fourier-transform of W. The same quantity can also be expressed in real space. One can equally label the filter by its radius R or by its mass M, $\sigma^2(R) = \sigma^2(M)$. The latter is called the *mass variance*.

If $\delta(\vec{x})$ is a Gaussian random field, so is $\delta(\vec{x}; R)$ and the probability distribution function $\delta_M = \delta(\vec{x}; M)$ is

$$P(\delta_M)d\delta_M = \frac{1}{\sqrt{2\pi}\sigma_M} \exp\left[-\frac{\delta_M^2}{2\sigma_M^2} \right] d\delta_M \tag{1.40}$$

where $\sigma_M = \sigma(M)$. In hierarchical models, such as the CDM model, the variance is monotonically decreasing with increasing scale R. In top-down cosmologies however, the lack of small scale structure introduces a characteristic scale with maximum variance.

Press and Schechter postulated: "The probability of $\delta_M > \delta_c(t)$ is the same as the mass fraction that at time t is contained in halos with mass greater than M". For a Gaussian random field

this statement results in

$$F(>M, t) \equiv P(\delta_M > \delta_c(t)) = \frac{1}{\sqrt{2\pi}\sigma_M} \int_{\delta_c}^{\infty} \exp\left[-\frac{\delta_M^2}{2\sigma_M^2}\right] d\delta_M = \frac{1}{2}\left(1 - \mathrm{erf}\left[\frac{\delta_c}{2\sigma_M}\right]\right). \quad (1.41)$$

$\delta_c(t)$ is the time dependent critical density for collapse when computing δ_M at fixed time, i.e. when computing δ_M on todays power spectrum and normalizing the growth factor such that $D(a=1)=1$, then $\delta_c(t) = \delta_c/D(t)$.

One can see from equation 1.41 that for $t \to \infty$, i.e $\delta_c \to 0$ will never result in more than half the mass being collapsed in halos. The Press-Schechter formalism does not account for under-dense regions ($\delta < 0$) being able to collapse as part of larger over-dense regions. This is a shortcoming. (40) "re-solved" this problem by correcting equation (1.41) with an additional factor of 2

$$F(>M, t) = 2 \cdot P(\delta_M > \delta_c(t)). \quad (1.42)$$

The expression $\frac{\partial F(>M)}{\partial M} dM$ describes the mass fraction bound in objects of mass $[M, M + dM]$. Multiplying by $\bar{\rho}$ leads to the total mass bound in those object per unit volume.

With the formalism providing $F(>M, t)$, one can define the halo mass function $n(M, t)dM$, which is the number of objects in the mass range $[M, M + dM]$ in a fixed (co-moving) volume, hence $n(M, t) = \frac{dn}{dMdV}$, as

$$n(M, t)dM = \frac{\bar{\rho}}{M}\frac{\partial F(>M)}{\partial M}dM = \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\frac{\delta_c}{\sigma_M}\exp\left(-\frac{\delta_c^2}{2\sigma_m^2}\right)\left|\frac{d\ln\sigma_M}{d\ln M}\right|dM. \quad (1.43)$$

Equation (1.41) and (1.42) and $\partial F/\partial M = \partial F/\partial\sigma_M|\partial\sigma_M/\partial M|$ were used in the last step. The expression above can be simplified by introducing the variable $\nu = \delta_c(t)/\sigma(M)$. In this form, equation (1.43) writes

$$n(M, t)dM = \frac{\bar{\rho}}{M^2}f_{PS}(\nu)\left|\frac{d\ln\nu}{d\ln M}\right|dM, \quad (1.44)$$

where

$$f_{PS}(\nu) = \sqrt{\frac{2}{\pi}} \nu e^{-\nu^2/2}. \tag{1.45}$$

The characteristic mass M_{char} is defined by $\sigma(M_{\text{char}}) = \delta_c(t)$, i.e $\nu(M_{\text{char}}) = 1$. The Press-Schechter HMF obeys the following properties:

- When $M \ll M_{\text{char}}$, $n(M, t) \propto M^{-(2+\alpha)}$, where $\alpha = d\ln\sigma_M/d\ln M$. For a CDM cosmology, it holds $\alpha \sim 0$ and therefore $n(M, t) \propto M^{-2}$ at low masses.

- For $M \gg M_{\text{char}}$, the abundance of halos is exponentially suppressed.

- Since $\delta_c(t)$ is a decreasing function with time, the characteristic halo mass ν increases with time and more and more high mass structure is formed.

From the Press-Schechter formalism, a hierarchical structure formation picture emerge. Small structures are formed first and larger and larger structure are formed with time.

1.2.4 Extended Press-Schechter formalism (EPS)

The extended Press-Schechter formalism (EPS), was introduced by (41) and is an extension of the Press-Schechter approache which consistently treat under-dense regions $\delta < 0$ and has no need of a 'fudge factor" of two (equation 1.42). Furthermore, EPS allows the description of progenitor halos and can trace back the past of the halos, which is not possible in the Press-Schechter formalism.

For a hierarchical model, $S \equiv \sigma^2(M)$ is a monotonic declining function of halo mass M. There is a one-to-one relation between S and M. Let's consider a position $\vec{x}$ with present day over-density is $\delta_0(\vec{x})$. For each value of the filter mass M (or S), the smoothed over-density $\delta_S = \delta_M(\vec{x})$ can have a different value. On very large scales $M \to \infty$ ($S \to 0$) the universe is homogeneous and therefore $\delta_S \to 0$, independent of position $\vec{x}$. Each point in space $\vec{x}$ can thus be described by a trajectory (S, δ_S) with starting point $(0, 0)$. This means that the filter S considers smaller and smaller scales and the density fluctuation δ_S within the filter changes accordingly.

If the filter S is a sharp in k-space, independent modes $\Delta\delta_S$ are added to the trajectory (S, δ_S) in a Gaussian random field. Statistically, a trajectory (S, δ_S) follows a Markovian random walk. As all random walks will at one point (with very large s) cross $\delta_c(t)$, EPS predicts that all the mass resides in halos. The EPS formalism uses explicitly the sharp k-space filter.

The EPS ansatz is: The fraction of trajectories (S, δ_S) with first up-crossing (FU) above $\delta_S > \delta_c(t)$ at $S > S_1 = \sigma^2(M_1)$ is equal to the mass fraction at time t residing in halos with $M < M_1$.

Based on the EPS ansatz, one can write the EPS mass function

$$n(M, t)dM = \frac{\bar{\rho}}{M}\frac{\partial F(>M)}{\partial M}dM = -\frac{\bar{\rho}}{M}\frac{\partial F(<M)}{\partial M}dM = -\frac{\bar{\rho}}{M}\frac{\partial F_{FU}(>S)}{\partial S}\frac{dS}{dM}dM$$
$$= \frac{\bar{\rho}}{M}f_{FU}(S, \delta_c)\left|\frac{dS}{dM}\right|dM \quad (1.46)$$

where $f_{FU}(S, \delta_c)$ is the fraction of trajectories that have their first up-crossing of barrier $\delta_c(t)$ in the range $[S, S + dS]$, given by (without proof, see (41))

$$f_{FU}(\nu) = \frac{1}{\sqrt{2\pi}}\frac{\delta_c}{S^{3/2}}\exp\left[-\frac{\delta_c^2}{2S}\right] = \frac{1}{2S}f_{PS}(\nu). \quad (1.47)$$

This yields to the same mass function as the Press-Schechter formalism (equation 1.43), but without the need of a fudge-factor of 2.

The EPS formalism has specific assumptions and shortcomings. First, as for the Press-Schechter formalism, it is anchored on the spherical collapse model. Second, individual paths of halos are not predicted by EPS neither. For example, let us consider two positions $\vec{x}_1$ and $\vec{x}_2$ which reside in the same halo. In this case, EPS predicts two different halo masses as the masks S_1 and S_2 are centered on the two different positions. EPS only predicts how much mass resides in halos of different mass only in a statistical sense. Third, the sharp k-space filter has a complicated form in real-space. The real-space filter is not spatially localized.

Nevertheless, EPS finds many applications in cosmology and was further developed. (42) introduced the elliptical EPS formalism and the original HMF were tuned to match N-body simulations (see e.g. 43).

1.2.5 Halo merger trees

Of importance in cosmology and galaxy evolution is not only the statistics of the halo popula-
tion but also the individual paths of different halos, their growth and the progenitors which
assembled the current halo. An advantage of EPS over PS is that EPS enables to calculate the
properties of the progenitors of a given class of objects, the so called *progenitor mass function.*
In particular, the progenitor mass function $n(M_1, t_1|M_2, t_2)dM_1$ gives the average number of
progenitors of mass $[M_1, M_1 + dM]$ at time t_1 that at time $t_2 > t_1$ have merged into a halo of
mass M_2. The progenitor mass functions allows to construct halo merger trees, an assembly
history of all the progenitors merging in the apparent halo.

The derivation of $n(M_1, t_1|M_2, t_2)dM_1$ is as follow: Consider a spherical region of mass M_2
with linear over-density $\delta_2 = \delta_c(t_2) = \delta_c/D(t_2)$ such that it forms a collapsed object at t_2. Of
interest is the fraction of M_2 which at an earlier time $t_1 < t_2$ resides in a collapsed halo of mass
M_1. Within EPS, this is the conditional probability that a trajectory that crossed the barrier δ_2
at S_2 has its first up-crossing of $\delta_1 = \delta_c(t_1)$ at $S_1 > S_2$. This is the same quantity as equation
(1.47), except for a translation of the origin in the (S, δ_s)-plane:

$$f_{FU}(S_1, \delta_1|S_2, \delta_2) = \frac{1}{\sqrt{2\pi}} \frac{\delta_1 - \delta_2}{(S_1 - S_2)^{3/2}} \exp\left[-\frac{(\delta_1 - \delta_2)^2}{2(S_1 - S_2)} \right] \tag{1.48}$$

and converting from mass to number-weighted leads to

$$n(M_1, t_1|M_2, t_2)dM_1 = \frac{M_2}{M_1} f_{FU}(S_1, \delta_1|S_2, \delta_2) \left| \frac{dS_1}{dM_1} \right| dM_1 \tag{1.49}$$

Several numerical methods exist in the literature to sample the progenitors of a halo M_0 at a
look-back time-step Δt, that obey the progenitor mass function and mass conservation, e.g.
(44; 45; 46; 47).

1.3 Gravitational lensing

Gravitational lensing is the effect caused by density inhomogeneities along the path of a
photon. The effect is a prediction of GR and was proposed by Albert Einstein as a test of GR
in 1916 (48). In 1919 the effect was quantitatively measured and was in agreement with the

prediction of GR (49). Gravitational lensing is caused by the *total* mass, by stars, gas and dark matter and can be used as a probe of the total mass distribution in the universe.

In this section, we first discuss the physical effects caused by a spherical mass. In a second paragraph, we discuss the geometrical component of gravitational lensing and time-delays in a third paragraph.

1.3.1 Physics of gravitational lensing

The geometric solution outside a spherical mass M of Einstein's Equation (1.1) is the Schwarzschlid metric, which is given by

$$ds^2 = \left(1 - \frac{r_s}{r}\right) c^2 dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2 \left(d\theta^2 + \sin^2\theta\, d\phi^2\right), \tag{1.50}$$

where t is the time coordinate (measured by a stationary clock located infinitely far from the massive body), r is the radial coordinate (measured as the circumference, divided by 2π, of a sphere centered around the massive body), and r_s is the *Schwarzschild radius* of the massive body given by

$$r_s = \frac{2GM}{c^2}. \tag{1.51}$$

The metric predicts a deflection and a time dilatation of light rays passing close to the object at an impact radius b relative to a flat Minkowsky metric ($b = \infty$). To analytically compute these quantities, certain approximations have to be made. In this thesis, we consider impact parameters b in the weak gravitational regime ($b \gg r_s$) and that the effective deflection of the light path is computed on the unperturbed straight path (the Born approximation). These approximation are reasonable and only break down very near a black hole and other very dense astrophysical sources, such as a neutron star. The resulting physical deflection angle $\hat{\alpha}$ is

$$\hat{\alpha} = \frac{4GM}{c^2 b} \tag{1.52}$$

and the gravitational time-delay t^{grav} caused by the change in the speed of light c' (relative to an observer at $r = \infty$) can be written as

$$t(b)^{\mathrm{grav}} = \int \frac{dl}{c} - \int \frac{dl}{c'} \approx \frac{2GM}{c^3} \int \frac{dl}{r} \tag{1.53}$$

to first order. This integral diverges for infinite path lengths, independent of the distance of closest encounter b. Of interest is the relative change in time delay (modulo a constant diverging shift). We define the relative time delay $\Delta t_\infty^{\mathrm{grav}}$ as the time difference relative to a path with $b = \infty$. In this form, the relative time delay for a Schwarzschild lens becomes

$$\Delta t_\infty^{\mathrm{grav}} = t(\infty)^{\mathrm{grav}} - t(b)^{\mathrm{grav}} = \frac{4GM}{c^3} \log(b). \tag{1.54}$$

Applying Huygens principle on a wave front traveling through a varying diffraction index with path differences of $\Delta l = c \Delta t_\infty^{\mathrm{grav}}$ leads to a deflection in the form

$$\hat{\alpha} = \frac{\partial \Delta l}{\partial b} = \frac{4GM}{c^2 b}. \tag{1.55}$$

One recovers the result of equation (1.52). Gravitational lensing follows the physics of classical and quantum optics of light propagation. Huygens principle holds and Fermat's principle, that can be derived from Huygens principle, as well.

More realistic mass distribution can be described in the weak field regime as a sum of Schwarzschild lenses and their effects of deflection and time-delay add up. Attention has to be payed to the geometrical aspects of the mass distributions and the Born approximation may no longer be valid.

1.3.2 Lens equation

Deflection of light paths act as a distortion of intrinsic luminous features and can be described phenomenological as a mapping from the image plane to the source plane. Considering the angular position on the sky from an observer, the difference between the un-lensed angular

position $\vec{\beta}$ (source plane) and the observed position $\vec{\theta}$ (image plane) is the deflection angle $\vec{\alpha}$ given as

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}). \tag{1.56}$$

Corresponding to the vector field $\vec{\alpha}$, there exists a scalar potential ψ, called the *lensing potential* such that

$$\vec{\alpha}(\vec{\theta}) = \vec{\nabla}\psi(\vec{\theta}). \tag{1.57}$$

The *lensing convergence* $\kappa(\vec{\theta})$ of the lensing potential is defined as

$$\kappa(\vec{\theta}) \equiv \frac{1}{2}\nabla^2\psi(\vec{\theta}). \tag{1.58}$$

The lensing potential itself can be expressed in terms of the convergence as

$$\psi(\vec{\theta}) = \frac{1}{\pi}\int d^2\vec{\theta}'\kappa(\vec{\theta}')\ln|\vec{\theta}-\vec{\theta}'|. \tag{1.59}$$

To first order, the lens effect described in the form of equation (1.56) can be expressed as the *Jacobian* A_{ij}

$$A_{ij} \equiv \frac{\partial\beta_i}{\partial\theta_j} = \delta_{ij} - \frac{\partial\alpha_i}{\partial\theta_j} = \delta_{ij} - \frac{\partial^2\psi}{\partial\theta_i\partial\theta_j} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{1.60}$$

where γ_1 and γ_2 correspond to the *shear* terms. The *magnification* μ is the inverse of the determinant of the Jacobian A

$$\mu = \frac{1}{\det(A)}. \tag{1.61}$$

In the thin lens approximation, the deflection $\vec{\alpha}(\vec{\theta})$ comes from one or several mass distributions which are thin compared to the light paths from the source to the observer. When only considering one thin lens, this approximation leads to a description between the visible

deflection angle on the sky $\vec{\alpha}(\vec{\theta})$ relative to the effective physical bending angle $\vec{\hat{\alpha}}$ as

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}) = \vec{\theta} - \frac{D_{\mathrm{ds}}}{D_{\mathrm{s}}} \vec{\hat{\alpha}}(D_{\mathrm{d}}\vec{\theta}) \tag{1.62}$$

where D_{d}, D_{s} and D_{ds} are the angular diameter distances from the observer to the lens, to the source and from the lens to the source [1], respectively.

In the case where there are n thin lenses along the line-of-sight (LOS), the total mapping is the sum of all the deflections, weighted by their distance relations and evaluated at the light path as

$$\vec{\beta}_s = \vec{\theta} - \frac{1}{D_s} \sum_{k=1}^{n} D_{ks}\vec{\hat{\alpha}}_k(D_k\vec{\beta}_k) \tag{1.63}$$

where $\vec{\beta}_k$ is the angle under which the k'th lens deflects the light ray and $\vec{\beta}_1 = \vec{\theta}$. The ordering in D_k follows the backwards light path such that the light passes the lens at k before $k-1$ (see also (50) for a different notation of the same expression).

1.3.3   Time-delays

So far, the geometry was folded in the mapping of the light paths from the image to the source plane. Apart of the gravitational time-delay, the light travel time is also affected by the geometric configuration.

Considering a light ray propagating at an angle $\vec{\theta}$ between two deflectors $(k, k+1)$ whose origin emerged from the direction of angle $\vec{\beta}$. The difference in the co-moving path length of this light ray compared with a light ray traveling straight at the angle $\vec{\beta}$, assuming $(\theta - \beta) \ll 1$ to second order, is given by

$$\Delta l_{k,k+1} = T_{k,k+1} \frac{(\theta - \beta)^2}{2}. \tag{1.64}$$

Time delays are determined by measuring photons emitted at the same time but received at a different time. This means that the path difference is made at recent cosmological time ($z = 0$).

---

[1] D_{ds} is *not* the subtraction $D_{\mathrm{d}} - D_{\mathrm{s}}$. In a flat universe: $D_{\mathrm{ds}} = \frac{1}{1+z_s}(T_{\mathrm{d}} - T_{\mathrm{s}})$, where T is the transverse co-moving distance.

The excess distance of equation (1.64) result in a time delay of

$$\Delta t_{k,k+1}^{\text{geo}} = \frac{\Delta l_{k,k+1}}{c} = \frac{T_{k,k+1}}{c} \frac{(\theta - \beta)^2}{2}. \tag{1.65}$$

In the general case of multiple lens planes, the total time delay relative to a straight path is given by

$$\Delta t^{\text{geo}} = \sum_{k=1}^{N} \Delta t_{k-1,k}^{\text{geo}} = \frac{1}{c} \sum_{k=1}^{N} T_{k-1,k} \frac{(\hat{\alpha}_{tot,k-1} - \beta)^2}{2}, \tag{1.66}$$

where $\hat{\alpha}_{tot,k}$ is the physical angle of the light ray after (backwards view) the k'th deflection.

The integrated gravitational time delay is the sum of the individual gravitational time delays $\Delta t_{\infty,k}^{\text{grav}}$ of equation (1.54) rescaled to the observer time at $z = 0$ given by

$$\Delta t^{\text{grav}} = -\sum_{k=1}^{N-1} (1 + z_k) \Delta t_{\infty,k}^{\text{grav}}. \tag{1.67}$$

The combined time delay from the geometric path difference (equation 1.66) and the gravitational time difference (equation 1.67) results in

$$\Delta t^{\text{tot}} = \Delta t^{\text{geo}} + \Delta t^{\text{grav}}. \tag{1.68}$$

Example: Single thin lens

As an example, we calculate the single thin lens time delay of an image visible at $\vec{\theta}$ lensed from the source at angular position $\vec{\beta}$. The relative geometric time delay is then given by

$$\Delta t^{\text{geo}} = \frac{T_{\text{d}}}{c} \frac{(\vec{\theta} - \vec{\beta})^2}{2} + \frac{T_{\text{ds}}}{c} \frac{(\vec{\theta} - \vec{\hat{\alpha}} - \vec{\beta})^2}{2} \tag{1.69}$$

which can be simplified by applying the lens equation (1.62) to

$$\Delta t^{\text{geo}} = \frac{1}{c} \frac{(\vec{\theta} - \vec{\beta})^2}{2} \frac{T_{\text{d}} T_{\text{s}}}{T_{\text{ds}}}. \tag{1.70}$$

An equivalent expression can be stated in angular diameter distances

$$\Delta t^{\text{geo}} = \frac{1 + z_d}{c} \frac{(\vec{\theta} - \vec{\beta})^2}{2} \frac{D_{\text{d}} D_{\text{s}}}{D_{\text{ds}}}. \tag{1.71}$$

The gravitational delay is given by

$$\Delta t^{\text{grav}} = -(1 + z_d) \Delta t_{\infty}^{\text{grav}}(D_{\text{d}} \vec{\theta}). \tag{1.72}$$

The convergence (equation 1.58) becomes the scaled projected surface mass density

$$\kappa(\vec{\theta}) = \frac{\Sigma(D_{\text{d}} \vec{\theta})}{\Sigma_{\text{crit}}} \tag{1.73}$$

with

$$\Sigma_{\text{crit}} = \frac{c^2 D_{\text{s}}}{4\pi G D_{\text{d}} D_{\text{ds}}} \tag{1.74}$$

being the *critical density* . The

The excess time delay of an image at $\vec{\theta}$ with corresponding source position $\vec{\beta}$ can be written as (combining equation 1.71 and 1.72)

$$t(\vec{\theta}, \vec{\beta}) = \frac{D_{\Delta t}}{c} \phi(\vec{\theta}, \vec{\beta}), \tag{1.75}$$

where

$$\phi(\vec{\theta}, \vec{\beta}) \equiv \left[\frac{(\vec{\theta} - \vec{\beta})^2}{2} - \psi(\vec{\theta}) \right] \tag{1.76}$$

is the Fermat potential with

$$D_{\Delta t} \equiv (1 + z_{\text{d}}) \frac{D_{\text{d}} D_{\text{s}}}{D_{\text{ds}}} \tag{1.77}$$

is the *Time-delay distance* for a single lens. The relative time delay difference Δt_{ij} between

two images positioned at $\vec{\theta}_i$ and $\vec{\theta}_j$, the actual observable, is then given by

$$\Delta t_{ij} = t_i(\vec{\theta}_i, \vec{\beta}) - t_j(\vec{\theta}_j, \vec{\beta}). \tag{1.78}$$

LOS structure external to the lens also affect the observed time delay distance through additional focusing or de-focusing of the light rays. A simplified way to parameterize additional deflections along the LOS on the time-delays is by a single constant mass sheet parameter κ_{ext}, the *external convergence* . The actual time delay distance $D_{\Delta t}$ relates to the one inferred by ignoring the external LOS structure by

$$D_{\Delta t} = \frac{D_{\Delta t}^{\text{model}}}{1 - \kappa_{\text{ext}}}. \tag{1.79}$$

This description is used in chapter 5. A more detailed description of the LOS structure is given in chapter 4.

2 A Simple model linking Galaxy and Dark Matter Evolution

This chapter appeared in a similar form in Birrer, Lilly, Amara, Paranjape & Refregier 2014 (28).

Galaxy evolution is a field where cosmological structure formation needs to be enriched with astrophysical processes, i.e. astrophysics has to be embedded into a cosmological model. It is the largest scale where astrophysical models have to succeed and the smallest scales where the cosmological structure formation model has to prove its validity. Galaxies, and the galaxy population, therefore offer tests for both astrophysics and cosmology.

Several approaches have been taken to understand the link between galaxies and dark matter haloes. Usually, the dark matter component is assumed to be well understood on the basis of both analytic and numerical models that are based on input parameters derived from cosmological observations, e.g. the cosmic microwave background. Small collapsed objects, i.e. 'haloes", form earlier and subsequently merge together to form more massive objects. Numerical N-body simulations provide an accurate description of the evolution of the population of dark matter haloes in the cosmological context (e.g, 51; 52). Much of the difficulty in galaxy formation and evolution arises then in understanding the actions of baryonic physics within these haloes.

A major theoretical effort has been made using so-called "semi-analytic" techniques to follow the evolution of baryons in the haloes. In semi-analytic models (or SAMs) simple parametric descriptions of the most important baryonic physics are combined with a dark matter merger

tree that is usually obtained from a large volume N-body simulation. The treatment of the relevant baryonic processes is necessarily simplified (e.g, 53; 54; 10; 55; 56; 57; 58; 59; 60). Some or all of the parameters describing these processes can be adjusted to match particular observational properties of galaxies or of the galaxy population, either at a single epoch or at many. Although much progress has been made and the range of output quantities can be large, the total number of parameters in such models is often quite large and as a result, the uniqueness and predictive power of SAMs is limited. In addition, the apparent complexity of the SAMs can often hide underlying links between different aspects of galaxy evolution.

Much progress has also been made using the alternative approach of ab initio simulations in which the baryonic physics is directly incorporated into hydrodynamic codes. However, due to the very large dynamical range that must be covered, such simulations are currently not able to resolve star formation and associated feedback processes and so cannot describe these processes from first principles. Simulation codes therefore include these as "sub-grid" physics, which leads to the emergence of a number of alternative approaches (e.g, 60; 61).

Partly in response to these difficulties, other, more phenomenological, approaches have been developed. One has been to study the statistical connection between galaxies and dark matter haloes in terms of the conditional luminosity function (CLF; 62) or the halo occupation distribution (HOD; e.g, 63; 64). These methods are anchored on our good understanding of the statistical properties of dark matter haloes in the current ΛCDM model plus the hypothesis that galaxy properties should be closely linked to the properties (and especially the masses) of dark matter haloes. A variety of statistical tools can then be used to constrain the galaxy-dark matter connection: galaxy clustering (e.g, 65), galaxy-galaxy lensing (e.g, 66; 67; 68), galaxy group catalogs (e.g, 69; 70), abundance matching (recently e.g, 71; 72; 73; 74), and satellite kinematics (e.g, 75).

In recent years, large scale surveys of the distant Universe have yielded sufficient data to apply similar approaches at significant look-back times. The differential effects with redshift then allow a phenomenological description of the evolving galaxy population using simple parametric descriptions. The parameters of these are matched to the evolving statistical descriptions of the stellar to halo mass relation (e.g, 76; 77; 78; 79). Such models can provide

consistency checks within several data sets and observables. As an example, when compiling different data sets, (78) finds a disagreement between galaxy abundances for high redshift surveys and high systematic errors in the stellar mass and star formation rate estimates.

The increasingly good observational data on the evolving galaxy population has also opened up other phenomenological approaches which instead focus on the baryonic processes. A successful approach has been to broadly classify galaxies as either actively forming stars or quiescent. Most star-forming galaxies exhibit a rather tight relation between their star formation rates (SFR) and stellar masses producing the so called Main Sequence (80; 81; 82; 83; 84). The quiescent galaxies have sSFR that are 1-2 orders of magnitude lower, and these galaxies are not forming stars at a cosmologically significant rate. We will henceforth refer to these passive galaxies as "quenched". A few underlying simplicities in the galaxy population can then be identified (such as the observed constancy of the Schechter M^* of star-forming galaxies or the separability of the fraction of galaxies that are quenched - colloquially the "red fraction"). The analytic consequences of these can then be explored using the most basic continuity equations (83; 85, hereafter P10 and P12). This has proved very successful in describing the evolution of the galaxy population and, in particular, in deriving the simple empirical "laws" for the quenching of star-formation in galaxies as a function of stellar mass (even if other parameters are involved or are even the main causal drivers). This approach has also yielded new insights into the relationships between the mass functions of active and passive galaxies, and the relative importance of mass and environment in the quenching of star-formation in galaxies.

There have also been several papers developing simple toy analytic models for the star-formation rate in galaxies (e.g, 86; 87; 88; 89; 90; 91, L13 from here on). These have been motivated by the small dispersion in the *specific star-formation rate* (sSFR = star formation rate/stellar mass) of actively star-forming galaxies, and by the strong evolution of this characteristic sSFR with time. In terms of the CLF a phenomenological approach has been chosen by (92). (93) developed a toy analytic model when comparing to hydrodynamical simulations. All these models have tried to boil down the complexity arising in numerical simulations and detailed semi-analytic models into simple analytic models that are motivated by either simulation results or observational constraints. The aim has been to provide a simple picture

of how galaxies evolve in the cosmological context and to highlight connections between different aspects of galaxy evolution. In particular, L13 developed a toy analytic model in which the star-formation rate is regulated via the variable mass of gas in the gas reservoir feeding the star-formation. Such a model links the specific star-formation rate (sSFR) to the specific accretion rate onto the regulator system. The self-regulation by the gas reservoir naturally introduces the SFR as a second parameter in the *mass-metallicity relation* $Z(m, \mathrm{SFR})$ and also naturally explains why the $Z(m, \mathrm{SFR})$ relation should be more or less independent of time. This model also links in a straightforward way the different slopes of the mass functions of galaxies and haloes.

By construction, the phenomenological analytic models in P10, P12 and L13 have only been tangentially linked to the dark matter haloes and not at all to the overall evolving *population* of haloes that is produced by hierarchical assembly in the cosmological context. The whole approach, and in particular the derivation of the numerical values of the few parameters in the models, was based on comparison with baryonic systems. This has been both a strength and a weakness of these analyses.

The aim of this chapter is therefore to explore how far we can get by taking these simple baryonic prescriptions and combine them with a dark matter structure formation formalism. Specifically, we will take the self-regulation model of L13 plus the quenching "laws" of P10 and P12. We couple them with a Monte Carlo realisation of dark matter halo merger trees. Our goal is to present a phenomenological model whose few parameters are taken from the earlier papers, and are not adjusted in the combined model. The parameters are well constrained and therefore considered as non-adjustable in this work. We can then explore how well these predictions match the observed Universe, and identify where and how it needs further improvement. In a second step, we propose two changes in the model and show their impact on the predictions.

Our approach is thus rather different to the one in (79) or (78) as we do not explore a parameter space but rather develop a physical picture without further tuning within the combined model. We stress that the current model is not intended to replace more complex SAMs whose greater sophistication will no doubt be required to account for a more multi-dimensional view of

galaxies.

The current chapter is structured as follow: In Section 2.1 we review the key concepts that were introduced in the earlier papers P10, P12 and L13 which we use to establish the characteristics of the baryonic processes. We define our notation and parameterization of these independent models and describe the dark matter structure formation formalism we apply. In Section 2.2 we describe how these are then combined into the dark matter merger tree, and what further assumptions have to be added, and how the model is then run. In Section 2.3, we present our results in terms of the most basic observables of the galaxy population such as the overall *star-formation rate density* (SFRD), the sSFR-mass relation of star-forming galaxies, the mass function of active and passive galaxies, and the form of the stellar mass vs. halo mass relation for star-forming and passive galaxies, and compare them with other work. In Section 2.4 we discuss the implications of the model and explore how one could modify it and where we are more restricted by the linkages between different parts of the model. Finally, in Section 2.5 we summarize this chapter.

Throughout this chapter, we assume a flat cosmology with $h = 0.7$ (i.e. $H_0 = 70 \text{kms}^{-1}\text{Mpc}^{-1}$), $\Omega_b = 0.045$, $\Omega_m = 0.3$, $\Omega_\lambda = 0.7$, $\sigma_8 = 0.8$ and $n_s = 1.0$ consistent with (94) WMAP7 results. We use the BBKS (95) transfer function to calculate the matter power spectrum. We define a halo as having a mean over-density $\Delta \equiv 3M_h/4\pi\Omega_m\rho_{\text{crit}}R_h^3 = 170$ to be consistent with the merger tree we use in this chapter. We use "dex" to refer to the anti-logarithm, so that 0.3 dex represents a factor of 2.

2.1 Model ingredients

In this section we review the concepts and descriptions used in our model. We start with the differential equations that control the regulator system from L13 (Section 2.1.1). We then quote the mass- and satellite- quenching expressions from P10 and P12 (Section 2.1.2). In Section 2.1.3 we describe the dark matter structure formation formalism we apply to our model. These ingredients are completely independent of each other and do not rely on mechanisms described in other subsections.

2.1.1 Galaxies as gas-regulated systems

We adopt the model proposed in L13. Several similar models have been proposed in the literature (e.g, 86; 87; 88; 89; 90) although there are significant differences in both concept and detail. We identify a galaxy as a gas-regulated system sitting in a dark matter halo. The SFR in the galaxy is set simply by the gas mass M_{gas} within a reservoir in the galaxy via a star-formation efficiency, ϵ. There is also mass-loss from the reservoir in the form of a wind that is parameterized by a mass-loading factor, λ, such that the outflow is $\lambda \cdot$SFR. Both of the ϵ and λ parameters are allowed to vary with the stellar mass M_s of the galaxy (and possibly the epoch, or redshift). In L13, the baryonic infall rate into the regulator Φ_b, which replenishes the reservoir, was assumed to be some fixed fraction f_{gal} of the baryonic infall onto the surrounding halo. Two obvious simplifications of the L13 model were that gas expelled from the galaxy in the wind was assumed to be lost forever, i.e. it does not mix with any surrounding gas in the halo, and that substructure within a halo was neglected, i.e. there was only one regulator in each halo. These issues will be discussed later in this chapter.

As in L13, the stellar mass is defined as the long lived stellar population assuming that a fraction R of newly formed stellar mass is promptly returned to the gas reservoir. The remaining stars will have a lifetime that is longer than the Universe. As in L13, we will set the mass-return factor $R = 0.4$, motivated by stellar population models (e.g, 96). The "stellar masses" used throughout this chapter will be these "long-lived" stellar masses. These are of order 0.2 dex smaller than the stellar masses that are obtained by integrating the SFR, which are sometimes quoted in the literature.

The build up in stellar mass $\dot{M}_s$ is then given by

$$\dot{M}_s = SFR \cdot (1 - R). \tag{2.1}$$

Following L13, the differential equations of the regulator in differential form can then be written as:

$$SFR = \epsilon \cdot M_{\text{gas}} \tag{2.2}$$

$$\dot{M}_{\text{gas,outflow}} = \lambda \cdot SFR \tag{2.3}$$

$$\dot{M}_{\text{gas}} = \Phi_b - \dot{M}_s - \dot{M}_{\text{gas,outflow}} = \Phi_b - \epsilon \left(1 - R + \lambda\right) M_{\text{gas}} \tag{2.4}$$

We will not go in detail into the analytic solution of these differential equations as L13 explored these in some detail.

The efficiency ϵ and the outflow load λ are intended to cover, albeit simplistically, all the baryonic processes within the galaxy system. L13 considered a power law parametrization for both these quantities as a function of the stellar mass M_s in order to match the observed Z(M_s,SFR) relation in (97). The parameterization as a function of stellar mass is a convenience and is still valid even if other quantities (e.g. halo mass) are responsible for the physical effect. The parameterization is:

$$\epsilon(M_s, z) = \epsilon_{10} \cdot \left(\frac{M_s}{10^{10} M_\odot}\right)^b \cdot \left(\frac{H(z)}{H_0}\right) \tag{2.5}$$

$$\lambda(M_s) = \lambda_{10} \cdot \left(\frac{M_s}{10^{10} M_\odot}\right)^a, \tag{2.6}$$

$H(z)$ is the Hubble rate at redshift z and H_0 the present-day Hubble constant. L13 assumed, following (98), that the star-formation efficiency would scale as the inverse dynamical time of the galaxies and haloes, which should scale as the Hubble rate, and we will do the same until revisiting this issue towards the end of the chapter[1]. For example (99) looked at the the role of the normalization and slope of the Kennicutt-Schmidt relation (our Equation 2.2) by varying this parameter and keeping all other parameters fixed. They find that a linear Kennicutt-Schmidt relation is a much better fit to observations than a strongly super-linear

---

[1]The actual redshift scaling of the efficiency in L13 is $\epsilon \propto (1 + z)$ which is a good approximation for the scaling of the Hubble rate at low redshifts.

relation, in agreement with L13.

The gas infall rate Φ_b is assumed to be closely related to the dark matter halo growth rate. We will describe this term in greater detail when discussing our model in Section 2.2 but in essence we set the f_{gal} parameter of L13 to unity, i.e. all of the gas flowing in a halo will be assigned (at least temporary) with a regulator system.

One of the most interesting features of this very simple regulator system is that the resulting sSFR is closely linked to the specific infall rate of the baryons, which L13 termed the $SMIR_B$.

Indeed, the model is motivated by the overall similarities between the observed sSFR(z) of the population of star-forming galaxies and the specific growth rate of dark matter haloes (e.g, 100, or L13). The sSFR will be exactly the specific baryonic infall rate if a constant fraction f_{star} of the incoming gas is converted into stars. If, however, this fraction increases as a given regulator evolves, e.g. if star-formation becomes more efficient as the stellar mass of the regulator increases, then the sSFR will be boosted relative to the specific baryon infall rate, as in Equation 36 of L13. Because this boosting of the sSFR is likely to be larger at low masses, this also has the effect of reversing the weak dependence of the sSFR on stellar mass relative to the dependence of the dark matter specific accretion rate on halo mass (see L13).

Another attractive feature of this regulator system is that it introduces the SFR as a second parameter in the mass-metallicity relation, producing a $Z(M_s, \mathrm{SFR})$ relation that will only change with epoch to the extent that the ϵ and λ parameters (at fixed M_s) evolve. In other words a so-called "fundamental metallicity relation" is a more-or-less natural outcome of the regulator. By comparing the expected $Z(M_s, \mathrm{SFR})$ with data from SDSS given by (97), L13 derived nominal values for the parameters ϵ_{10}, b, λ_{10} and a in Equation (2.5) and (2.6) above. Given the extreme simplicity of the model, the resulting values for $\epsilon(M_s)$ and $\lambda(M_s)$, which are quoted in Table 1 in L13 and included in Table 2.1 of this chapter, are surprisingly reasonable, giving gas depletion timescales (ϵ^{-1}) at $M_s \sim 10^{10} \mathrm{M}_\odot$ of about 2 Gyr and mass-loading factors of order unity. The gas depletion timescale and the outflow mass loading both decrease with increasing stellar mass resulting in more and more efficient conversion of inflowing baryons into stars as the stellar mass of the system increases. The fraction of incoming baryons that are converted to stars is denoted as f_{star} in L13. In the context of the simple analysis of L13,

this "saturation" of f_{star} can be traced to the pronounced flattening of the $Z(M_s)$ relation at high masses. We will return to this later in the chapter.

The processes associated with star-formation in galaxies are thus represented in our model by the four parameters (Equation 2.5 and 2.6) describing $\epsilon(M_s)$ and $\lambda(M_s)$, and taken straight from L13. As noted above, we will initially assume ϵ increases as $H(z)/H_0$, although we will revisit this assumption later.

Work by (e.g, 101; 102; 103; 104; 105; 106; 107) have emphasized the importance of supernova feedback. In L13, outflows of material represent an "inefficiency" in the production of stars, but do not "regulate" the level of star-formation, which is instead defined by the gas mass.

2.1.2 Quenching of star-formation in galaxies

In this chapter, we apply the phenomenological quenching prescriptions derived by P10 and P12. This is distinct from introducing a turnover in the efficiency parameter as done by (78) and (79) or cutting off the supply of gas, as done by e.g. (86), although the outcomes may be similar. There are many physical mechanisms that have been proposed for quenching. One popular approach is AGN feedback (see e.g, 108; 61; 109; 110). The AGN feedback also presents a viable solution to the cooling flow problem (see e.g, 111; 112; 113), hence its popularity. The P10 approach comes from the continuity in the two distinct galaxy populations and is not based on a particular physical mechanism but rather seeks to define the characteristics that any viable mechanism must satisfy.

We will assume that star formation within a galaxy stops instantaneously when it is quenched and that no significant star formation occurs afterwards. As a shorthand (and on plots) we will denote the actively star-forming galaxies as blue and those that are quenched as red although we will not consider the colors of galaxies per se. The "red fraction" will then be the fraction of galaxies of a given mass etc. that have been quenched.

P10 showed that the red fraction of galaxies as a function of mass and local projected over-density is separable in the two variables, suggesting that there are two dominant processes: one which depends on mass but not density (so-called "mass-quenching") and a second

environment-related process which should be independent of stellar mass. The mass-quenching process is then the only one that depends on mass, and therefore is the one that determines the shape of the mass-function of the surviving star-forming galaxies and, via the continuity equation, the shape of the mass function of the resulting (mass-quenched) population of passive galaxies. The observed constancy of the shape of the mass function of star-forming galaxies over a wide redshift range up to $z \sim 2$ (or even higher) imposes a strong requirement on the form of mass-quenching (see P10 and below).

Subsequently, P12 showed that the environment-quenching in the overall population can be fully accounted by a satellite quenching process that applies only to satellite galaxies. The probability that a previously star-forming central galaxy is quenched when it becomes the satellite of another galaxy is about 50%, independent of it's stellar mass. There are many possible suggestions for an environment-dependent quenching mechanism (see e.g, 114; 115; 116).

The P10 prescription for mass-quenching can be written either as a quenching rate, i.e. the probability that a given star-forming galaxy will be quenched per unit time, or as a survival probability to reach a certain mass without being quenched. The probability p_{quench} for a galaxy becoming quenched when increasing its stellar content by dM_s is given by

$$dp_{\text{quench}} = \mu dM_s, \tag{2.7}$$

for an infinitesimal dM_s. For a finite increase ΔM_s, one gets

$$p_{\text{quench}} = 1 - \exp\left[-\mu \Delta M_s\right], \tag{2.8}$$

The constant μ is required (see P10) to be M^{*-1}, where $M^* = 10^{10.68} M_\odot$ is the value of the characteristic stellar mass of the (single) Schechter stellar mass function of the blue star-forming population. Following P10, we take μ to be constant with time because M^* is observed to be more-or-less constant.

We will assume that the mass-quenching process acts on all galaxies, i.e. both centrals and satellites. This is motivated by the observational fact that M^* is the same for central and

satellite star-forming galaxies (P12). Because of the close coupling of stellar mass (and even BH mass) and halo mass for central galaxies, the action of a mass-quenching that is driven by stellar mass (as in the equation above) is hard to distinguish from one driven by halo mass for centrals, but again our point is that the outcome must be well represented by the empirical P10 quenching "laws".

For satellites, we apply an additional stochastic quenching process. When a central galaxy becomes the satellite of another galaxy because it's own halo merges with another more massive halo, the chance of it being (instantaneously) quenched is set to $p_{sat} = 0.5$. This additional quenching probability is only applied once to any particular galaxy when it first becomes a satellite. Because we do not, in the current work, consider the radial distribution of galaxies within haloes (e.g, 117), or try to compute the local over-density as in P12 or (118), we do not consider the density-dependence of p_{sat}, instead adopting a mean value. This mean value of $p_{sat} = 0.5$ is assumed to be constant with epoch, as shown in the zCOSMOS group catalogue to $z \sim 0.7$ (119; 118).

To summarize, the quenching of galaxies in this model is accounted by just two constants, $\mu = 10^{-10.6} M_{\odot}^{-1}$ for mass-quenching and $p_{sat} = 0.5$ for satellite quenching.

2.1.3 Dark matter structure formation

To describe the hierarchical structure formation process we take a simple model, far below the complexity of N-body simulation but aiming to account for most of the features of those simulations. The descriptions we apply have been incorporated in one or another way by many authors (recently e.g. by 79). We use the dark matter merger tree generator from (46), which is based on the excursion set theory (e.g, 40; 120; 41; 121) tuned to match the Millennium simulation (60). (46) showed that the tuned merger tree generator matches the overall halo mass function and the progenitor mass function for different halo masses very well back to redshift $z = 4$. The merger tree generates its trees with a Monte Carlo method. Given a halo mass M_h at redshift z it generates the progenitors at $z + \Delta z$ for small time steps Δz (backward process). In addition to a smoothed component growth there is a probability of

having a binary split in the merger tree with a host and a satellite halo:

$$M_h \xrightarrow{\Delta z} M_{\text{host}} + M_{\text{sat}} + \Delta M_{\text{smoothed}}, \tag{2.9}$$

where M_{host} is the most massive progenitor of M_h. The tree naturally divides the progenitors into a smooth component (all progenitors below a mass threshold M_{thresh}) and a merger component (growth due to accretion of mergers above M_{thresh}). We express the growth of a halo as

$$\dot{M}_h = \dot{M}_{h,\text{smoothed}} + \dot{M}_{h,\text{merger}}. \tag{2.10}$$

For the subhalo evolution we apply the formalism from (122). They used high resolution dark matter simulations with one host and one satellite halo to invert the dynamical friction time scale t_{df} and provide a fitting formula for t_{df} (Equations 5,6 in their paper with the further assumption that the last factor in their Equation 5 is equal to unity):

$$\frac{t_{\text{df}}}{\tau_{\text{dyn}}} = 0.216 \frac{(M_{\text{host}}/M_{\text{sat}})^{1.3}}{\ln(1 + M_{\text{host}}/M_{\text{sat}})} e^{1.9\eta}. \tag{2.11}$$

This formula depends on the host-to-satellite mass ratio $M_{\text{host}}/M_{\text{sat}}$ and orbital circularity η. (122) noted in their analysis that including the effect of baryonic bulges one gets an approximately 10% shorter t_{df}. This fitting formula has been tested for $0.025 \leq M_{\text{sat}}/M_{\text{host}} \leq 0.3$ and is applicable for $\eta \geq 0.2$. Note that the dynamical time $\tau_{\text{dyn}} \approx 0.1 H^{-1}$ with H being the Hubble parameter. The inverted dynamical friction time scale can be several times larger than the dynamical timescale τ_{dyn}. From numerical simulations, (123) have shown that the probability distribution of the orbital circularity η of dark matter subhaloes can be approximated by

$$P(\eta) \propto \eta^{1.2} \left(1 - \eta\right)^{1.2}. \tag{2.12}$$

For every merger event in our merger tree, we therefore draw η from this distribution and thereby introduce some scatter in the dark matter structure formation process.

So far, we have an expression for the survival time t_{df} of a subhalo. For the subhalo mass

evolution $M_{\text{subhalo}}(t)$ we implement a step function following (77)

$$
M_{\text{subhalo}}(t) = \begin{cases} M_{\text{sat}}\,(t = t_a) & t - t_a < t_{df} \\ 0 & t - t_a > t_{df}, \end{cases} \tag{2.13}
$$

where t_a is the time of accretion.

In this chapter, M_h refers to the total halo mass. The halo mass associated with the central galaxy is then given by

$$
M_{\text{central}} = M_h - \sum_i M_{\text{subhalo,i}}, \tag{2.14}
$$

where the sum is over all surviving subhaloes above a certain mass threshold M_{thresh}. We thereby identify all substructure above M_{thresh} and trace its evolution.

Our dark matter formalism clearly consists of some simplifications. The merger tree is tuned to a dark matter only simulation whereas our model contains baryonic matter too. One implicit simplification is that the baryonic matter component will not deviate from the behavior of dark matter. In other words, the gravitational forces from the dark matter are the dominant driver of baryonic structure formation and pressure terms are ignored. Likewise, there is no reverse effect from the baryons on the dark matter (see for example (124) for a more detailed description).

Also, it should be noted that the merger tree is tuned to a slightly different cosmology. However, the tuned parameters are dimensionless and as the excursion set approach is formulated for arbitrary power spectra, (46) argued that their merger tree can also be applied to different cosmologies. See also (125) for discussion on the accuracy. For the substructure evolution, we have applied a very simple description, especially for the time evolution of the substructure. Despite these simplifications, our chosen description provides us with a good picture of what is going on in the dark matter structure formation process. It does not however contain the detailed and accurate descriptions that would be needed for doing precision cosmology.

To summarize, we introduced one arbitrary parameter M_{thresh} in our structure formation model and take t_{df} from (122). The remaining parameters for the dark matter are taken from

the standard cosmology.

2.2 The model

In this section we describe how we combine all the ingredients given in Section 2.1. In particular we describe in Section 2.2.1 how we link the baryonic infall rate onto the regulator system to the dark matter structure formation process. In Section 2.2.2 we describe and discuss what happens in a galaxy-galaxy merging event in our model framework. In Section 2.2.3 we describe how the regulator at very low stellar masses can be described. The procedure to predict the cosmic abundances of galaxies and their properties is described in Section 2.2.4. Finally we emphasize in Section 2.2.5 how our model differs from others parametric approaches.

2.2.1 Link between baryonic and dark matter infall rate

To consistently integrate our regulator and quenching models into the dark matter framework, some further assumptions have to be made. First, only dark matter haloes and subhaloes above M_{thresh} in Section 2.1.3 will contain a regulator system. In other words, we ignore star-formation in haloes that are so small that we considered their infall as part of the smooth dark matter inflow. This is because they will be mostly gaseous. We set $M_{\text{thresh}} = 1.4 \cdot 10^9 M_\odot$. This is somewhat arbitrary, but is consistent with photo-ionisation heating suppressing cooling and star formation below a certain halo mass M_γ. $M_\gamma \sim 10^8 M_\odot$ during reionisation to $M_\gamma \sim$ few $\cdot 10^9 M_\odot$ (126; 127). For a more realistic model aiming to make predictions of low mass galaxies back to the epoch of reionisation, one would need to account for a change in the mass threshold. We explore the effect of changing M_{thresh} in the Appendix and show that it is small for the galaxy mass scales of interest.

In order to trace the gaseous baryons through the build-up of haloes, the following simple scheme was used. We will later refer to this model as Model A.

- First, all gaseous baryons in a given halo are associated at all times with one of the regulator systems (i.e. "galaxies") within that halo *except* for those baryons which have

been processed through a regulator and ejected from the galaxy through the wind described by λ in Section 2.1.1. These ejected baryons are assumed to be "lost" (we will revisit this assumption later in the chapter) and are no longer tracked. But, apart from this, all gaseous baryons are found within the reservoirs of the regulator systems.

- Second, when two haloes merge, the baryons that are at that time within each of the regulator systems in the haloes stay within those regulators, unless the (sub)halo subsequently decays and is disrupted (see below).

- Finally, smooth accretion of gas onto haloes, i.e. the baryonic inflow associated with the merging of haloes below M_{thresh}, is split between the sub-haloes as follows:

$$\Phi_{b,i} = f_b \dot{M}_{h,\mathrm{smoothed}} \cdot \frac{M_{\mathrm{subhalo,i}}}{M_h}. \tag{2.15}$$

This scheme ensures that every baryon which has not flown into some regulator in the past, will be assigned to a regulator when coming into a halo above M_{thresh}. It also ensures that when a regulator becomes a satellite, its infall rate and thus its SFR will not dramatically change, as observed (see P12). We note that when a galaxy is quenched, the gas inflow associated with this quenched galaxy will not be redirected to other active regulators. In our discussion later in the chapter, we will introduce a different assignment of the in-flowing gas and see how this will change our predictions, in what we will call Model B. Realizing that the gas inflow description is crucial to many observables we will then introduce a further Model C, which provides far more freedom in assigning gas to regulator systems.

As noted above, L13 considered only a single regulator in a given halo and introduced f_{gal} as the fraction of inflowing baryons that penetrate down and enter the regulator system at the center of the halo. L13 concluded that $f_{\mathrm{gal}} \sim 0.5$ was required to reproduce the stellar to dark mass ratio of typical galaxies. By associating all gaseous baryons to regulator systems, we are effectively setting f_{gal} to unity (i.e. eliminating this parameter) in Model A and B in the present work. However, because we now include multiple regulators (associated with the subhaloes) in a given halo and a two component growth (mergers and smoothed accretion), the net effect for the central regulator will be similar because only a fraction ($\dot{M}_{h,\mathrm{smooth}} / \dot{M}_h$) of the halo growth is associated with gas accretion and it will only receive a fraction ($M_{\mathrm{central}} / M_h$)

of the incoming gas. In other words, we would now understand that the adoption of the lower $f_{\mathrm{gal}} \sim 0.5$ in L13 was simply accounted for the two component growth of a halo, which was neglected in their treatment of regulator systems.

2.2.2   Subhalo disruption / galaxy-galaxy merging

We now turn to what happens when a subhalo decays according to Equation (2.13) and specifically what happens to the gas and stars within the regulator associated with that sub-halo. The two extreme cases would be adding all the stars and gas to the central galaxy or distributing them into the inter-cluster medium, which for the gas would involve re-distributing the gas amongst the surviving regulators according to Equation 2.15. Reality is likely in between these extremes. For concreteness and convenience, we set the fraction of stars and gas which are given to the central galaxy $f_{\mathrm{merge}} = 0.5$ but show in Appendix 2.A.1 that the output of the model is insensitive to this parameter. When the gas and stellar component from two different regulator systems is merged in this way, the new state of the regulator will likely not be in equilibrium with the gas infall rate. Galaxy-galaxy merging can thus lead to some scatter in the regulator properties. As discussed in L13 and illustrated in their Figure 3, the regulators adopt quickly to the new conditions and rapidly settle to the new equilibrium state.

2.2.3   Break down of the regulator description at low M_s

In L13, the parameters of the regulator (Equation 2.5 and 2.6 in this chapter) were tuned to match the metallicities of galaxies with stellar masses above $10^8 M_\odot$ and this parameterization must break down at lower stellar masses: not least, the mass-loading cannot increase without limit, simply on energetic grounds. We however need to include such low mass galaxies in our model so as to have larger galaxies later on. We therefore introduce a maximal outflow load. We set $\lambda_{\mathrm{max}} = 50$. This value is far off the regime where L13 tuned their parameters and therefore will not affect the validity of the tuning in L13. It also does not significantly affect the output of the model for galaxies above $M_s = 10^8 M_\odot$, the mass range of primary interest. Further discussion of this parameter can be found in Appendix 2.A.1.

2.2.4 Implementation

It will have been clear that the input galaxy data going into the model was derived independently of the number of galaxies, i.e. specifically it was the mean mass-SFR-metallicity relation (L13), the shape of the star-forming mass function parameterized by M^* (P10), and the red fractions of satellites (P12). A primary output of the model will be the expected number density of galaxies.

We therefore need to create a representative sample of the Universe. Merger trees derived from N-body simulations are sampled according to the halo mass function and therefore produce far more low mass halo trees than for high mass haloes. As we want to achieve the same statistical power over a wide range in halo mass, we want to equally sample the halo masses and weight their abundances in a second step. The merger tree generator provides such a possibility. The procedure is as follows: We sample 10'000 haloes at redshift $z = 0$, chosen randomly from a flat distribution in logarithmic halo mass, from $7.1 \cdot 10^9 M_\odot$ up to $1.4 \cdot 10^{15} M_\odot$. We then weight their abundance according to the halo mass function of (43) at $z = 0$. By construction, the weighted abundance of our haloes is then in perfect agreement with the input halo mass function at $z = 0$. We then let these haloes run backwards in cosmic time by applying the merger tree description. We stop when our resolution limit M_{thresh} is reached or at $z = 15$. At that point we identify our regulator systems, put in some initial stellar and gas mass and solve the differential equations for every single tree component. In parallel we apply the subhalo evolution model in the forward process. We thereby keep track of every satellite halo with its own regulator system. The model is not sensitive to the initial state of the regulators, as described in Appendix 2.A.1).

Clearly, this description has no spatial resolution, either within galaxies, within haloes or to follow the large scale distribution of haloes. The last of these would be relatively easy to implement and this will be the subject of future work. The other two would take us deeper into details, which we wish to avoid.

2.2.5   A model without re-adjusting the parameters

In Table 2.1 all of the parameters of our model are listed with a short description and reference to the input data on which they are based. These are mainly taken from the three papers P10, P12 and L13, and from cosmology and computational simplifications in the dark matter sector. The effects of the three additional parameters that we have introduced in this chapter, i.e. M_{thresh}, f_{merge} and λ_{max}, are investigated in the Appendices 2.A.1, 2.A.1 and 2.A.1. We conclude there that any reasonable variation within these parameters do not invalidate our conclusions. In essence, these parameters are introduced for practical reasons to make the model operable and the output does not depend very much on their precise values.

Within our chosen gas inflow description we therefore have virtually no freedom in changing our predictions: The model either matches observations or produces a discrepancy from which we may hope to learn. The goal is therefore not at first to produce a model that fits all available data, nor to observationally determine parameters. Rather, and in the spirit of the previous papers (P10,P12 and L13) we aim instead to provide insights into how well the ideas presented in those papers perform in the global context of a dark matter hierarchy, and to see where we encounter limitations.

2.3   Results

In Section 2.1, we reviewed the different and independent inputs that were then combined in Section 2.2 to produce a single model of star-formation and quenching in galaxies within a dark matter hierarchical framework. In this section, we compare the output of the default model A with both observations directly and with the outputs of other phenomenological approaches to galaxy evolution, most notably that of (78).

As discussed above, we will not vary any pre-adjusted parameter in our model beyond the three parameters introduced to allow the model to be computed (the values of which do not much affect the outcome) and so we can examine these comparisons one at a time. Throughout this section, we refer always to the same output sample generated with the parameters given in Table 2.1 with the inflow description of Equation 2.15, referred as our fiducial Model A.

Symbol	Description	Fixed to:	Units	Value
		Regulator Parameters (externally derived)		
ϵ_{10}	Efficiency normalization to $10^{10}M_\odot$ stellar mass	Metallicity data [a]	Gy^{-1}	0.33
b	Power law of efficiency as a function of stellar mass	Metallicity data [a]	-	0.3
λ_{10}	Outflow load normalization to $10^{10}M_\odot$ stellar mass	Metallicity data [a]	-	0.3
a	Power law of outflow load as a function of stellar mass	Metallicity data [a]	-	-0.8
		Quenching parameters (externally derived)		
M^*	Mass-quenching parameter μ^{-1}	Exponential cutoff of main sequence [b]	$M_\odot$	$10^{10.68}$
p_{sat}	satellite quenching probability	Elevated ref fraction of satellites [c]	-	0.5
f_{merge}	merging fraction of gas and stars of disrupted subhaloes	Parameter with no significant effect on our conclusions [e]	-	0.5
λ_{max}	Maximum outflow load of regulator	Upper bound provided by regulator action in tuning range [f]	-	50
M_{thresh}	Threshold in halo mass for having a regulator	Photo-ionisation model [g]	$M_\odot$	$1.4 \cdot 10^9$
		Cosmological Parameters (externally derived)		
h	dimensionless Hubble parameter	CMB [d]	-	0.7
Ω_b	Baryonic density	CMB [d]	-	0.45
Ω_m	Matter density	CMB [d]	-	0.3
Ω_λ	Dark Energy density	CMB [d]	-	0.7
σ_8	Power spectrum normalization	CMB [d]	-	0.8
n_s	spectral index	CMB [d]	-	1.0
		Additional simplification descriptions of the Dark Matter sector		
t_{df}	dynamical friction time scale	Dark Matter N-body simulation [h]	-	Eq 2.11
η	orbital circularity	Dark Matter N-body simulation [i]	-	Eq 2.12

[a] Data from (97) fitted by L13   [b] Data and model fit by (83)   [c] From (85), (118) and (119)
[d] From WMAP seven-year data (94)   [e] Further discussion in Appendix 2.A.1   [f] Further discussion in Appendix 2.A.1
[g] Model by (126) and (127), further discussion in Appendix 2.A.1   [h] Relation from (122)   [i] Relation from (123)

Table 2.1: This table lists all our model parameters (for Model A and B), its values and to what set of data they are tuned to.

It should be noted that the observational data used to determine these parameters were (a) gas metallicity data (as in L13) from (97) SDSS, specifically the $Z(M_s,\text{SFR})$-relation, (b) the red fraction of satellites (as in P12 from (128) SDSS DR7) and (c) the value of M* of star-forming galaxies (as in P10 also from SDSS). Any predictions of these particular quantities must therefore match observations, by construction, but predictions of all other quantities are bona fide and can be meaningfully compared with other data.

Comparison of these predictions with other data will enable us to draw several interesting conclusions. Some of the successes of these "predictions" will mirror conclusions that were already drawn in the original papers on which our new model is based, e.g. the discussions of mass functions and red fractions in P10 and P12, and the link between sSFR and specific accretion rate in L13. For these, it is reassuring to see them holding up in the context of a more realistic treatment of the haloes, including substructure and merging etc. None of the predictions based on the *population* of dark matter haloes could be made before, since they were not treated in the earlier works. These include the normalization of the mass functions and the computation of the star-formation rate density. We can also predict the scatter in various relations coming from different halo assembly histories.

Finally, we will make explicit comparisons with the output from the orthogonal phenomenological approach of (78). The (78) approach is anchored in the dark matter hierarchy and derives a very general description of the effect of baryonic processes within these haloes. In that work, a general M_s/M_h relation is assumed. The epoch dependent form of this is then derived by simultaneously applying statistical tools such as abundance matching of the mass functions at different redshifts, coupled with comparison of the consequent information on star-formation with a variety of observational data, including the $\text{sSFR}(M_s, t)$ and the global star-formation rate density SFRD. Our own approach is in a sense orthogonal to this as it is based on a prior determination of the purely baryonic phenomenology which is then imported into the dark matter structure. Despite the quite different approaches, and the obvious limitations of each of them, we will find that a very similar picture emerges.

2.3.1 Stellar Mass dependence of the Main Sequence sSFR at the present-day

We first plot in Figure 2.1 the specific star formation rate (sSFR) of all blue (i.e. star-forming) central galaxies of the output sample at $z = 0$ as a function of their stellar masses. The model successfully recovers the tight correlation between sSFR and mass which is known as the Main Sequence (e.g, 80; 81) and an almost constant sSFR with a scatter about this relation of about 0.2 dex.

For comparison with data we over-plot an sSFR(M_s, z) relation of the form

$$\text{sSFR} \propto M_s^\beta. \tag{2.16}$$

Observational estimates of β range between $-0.4 < \beta < 0.0$ at stellar masses above $10^9 M_\odot$, with most estimates $\beta \sim -0.1$ (e.g, 80; 81; 129; 82; 130; 131; 83). In Figure 2.1 the red line illustrates the data compilation in the form

$$\text{sSFR}(M_s, z) = 0.12 \left(\frac{M_s}{10^{10.5} M_\odot} \right)^\beta (1 + z)^3 \text{ (at } z < 2) \tag{2.17}$$

with $\beta = -0.1$ evaluated at $z = 0$ (see L13 and references therein). The observed scatter amongst real galaxies is about 0.3 dex once outliers with much higher sSFR are excluded (see e.g, 84; 132). These latter are associated with star-bursts, probably induced by mergers.

The mean sSFR(M_s) at $z = 0$ is clearly well reproduced by the model. As noted in L13 and discussed earlier in this chapter, a key feature of the kind of gas regulation considered in this chapter is that it sets the sSFR close to the specific mass accretion rate of the system, independent of the values of the parameters ϵ and λ controlling the regulator. There is a modest "boost" to the sSFR if an individual regulator system is increasingly efficient at producing stars as time passes (as would be expected if the efficiency increases with mass). This boost at $z = 0$ is expected to be of order 0.3 dex for typical galaxies. It increases to lower masses, potentially reversing the slope of the sSFR(M_s) relation relative to that of the specific accretion rate, defined as sMIR$= \dot{M}_h / M_h$. L13 took the approximation for the sMIR provided by (133). Despite our model using a more complex description for the baryonic infall rate Φ_b, we would expect to have the same underlying link between the sMIR and sSFR. The good
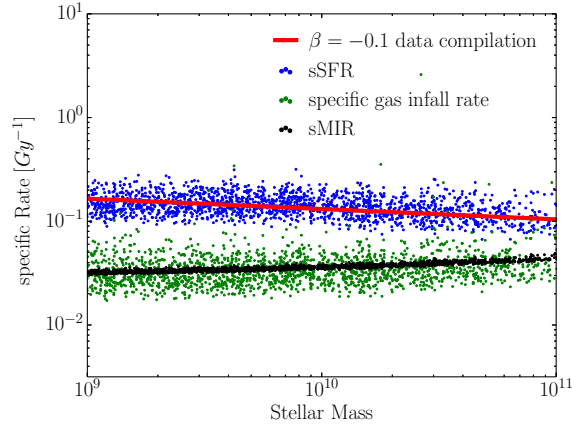
Figure 2.1: Prediction of the sSFR-mass relation at redshift $z = 0$ for blue central galaxies in Model A (blue). Dots correspond to individual galaxies in the model. The red line corresponds to the relation given in Equation 2.17 from a data compilation in slope and normalization (for citations see text). The green dots correspond to the specific gas infall rate of the same galaxies. Black dots denote the specific mass growth rate of the entire halo (sMIR) on timescales shorter than the last major merger event. The mean in the specific gas infall rate is the same as the sMIR. The sSFR is elevated by more than a factor of two. The scatter in the sSFR follows the scatter in the gas infall rate.

agreement with the mean $z = 0$ sSFR(M_s) relation in the current model which contains a wide variety of individual haloes is therefore reassuring but not unexpected given the discussion in L13 (see their Fig 9).

The scatter in sSFR(M_s) in our model is caused by the different halo formation histories, i.e. by the variation in the gas inflow rate caused by variations in the merger tree (green dots in Figure 2.1), and by the effects of galaxy-galaxy merging (see Section 2.2.2). Our model does not include any further stochastic time-variation in the gas infall Φ_b such as might be caused by other baryonic processes, and also neglects any stochastic scatter in the baryonic processes controlling star-formation within the galaxy regulator systems. Both of these could further increase the scatter (in our model there is almost no scatter occurring in the SFR-Φ_b relation). Our predicted scatter can therefore be interpreted as a lower bound in the expected sSFR(M_s) scatter. The fact that it is already 2/3 of the observed scatter suggests that these two further contributors to the scatter (stochastic infall variability and variation in the regulator) can contribute only of order 0.2 dex in normal Main Sequence galaxies.

2.3.2 Epoch dependence of the Main Sequence sSFR and the star-formation rate density

In Figure 2.2, we show the evolution in the sSFR for galaxies in the mass range $10^{10} M_\odot$-$10^{10.5} M_\odot$ back to $z = 5$, compared with data from (131) and a highly parameterized model of (78) adopted to our definition of sSFR=SFR/M_s. We also show for comparison the mean sMIR and the specific gas infall rate. As expected the sSFR tracks the increase in sMIR with redshift. While this broadly matches the data, the rise with redshift is not steep enough. As a result, the observed sSFR at $z \sim 2$ is about a factor of two higher than predicted from the model.

This is a common problem encountered in galaxy evolution models (e.g, 134; 135, and others) and is also present in the simple analysis of L13 that used an average halo growth rate from (133). Adjustment of the prediction would require a substantial modification of the accretion rate of baryons onto the regulator systems, i.e. breaking the link between the baryonic accretion rate onto the galaxy and the specific growth rate of the dark matter halo) or a rather dramatic adjustment of the efficiency with which inflowing gas is converted to stars (i.e. the f_{star} parameter of L13) so as to increase the boost factor associated with temporal changes in this quantity (see L13). We will return to this discrepancy in models B+C but note here that it is not inconceivable that some of the offset of 0.3 dex could reflect observational difficulties in determining stellar mass and star formation rates at high redshifts.

Our model naturally produces a deviation of the baryonic increase rate to the dark matter growth rate at very high redshifts as the dynamical friction time scale cannot catch up the halo growth rate resulting in far more substructure surrounding the central at high redshifts. More substructure means within our model that less baryonic infall will be assigned to the central as described in Equation (2.15).

In Figure 2.3, the overall star formation rate density (SFRD) is plotted over the whole range of cosmic time compared with data from the compilation by (136) and the phenomenological model by (78). The gray region is the 1-σ inter-publication scatter noted by (78).

The broad features of the evolving SFRD of the Universe are reproduced and our predicted value at $z = 0$ matches well the observational data of the nearby Universe. We again see a
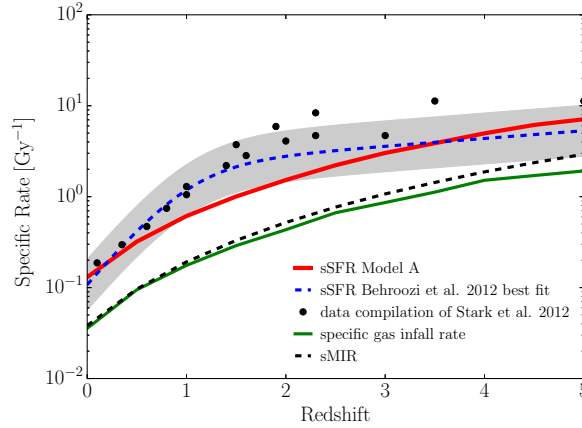
Figure 2.2: Prediction of the mean sSFR for blue galaxies within a stellar mass range of $10^{10} M_\odot$-$10^{10.5} M_\odot$ of Model A as a function of redshift (red curve). These are compared with data points (in black without errorbars) from (131) and a model based on a data compilation of (78) adjusted to our definition of SFR. The gray region reflects the 1-σ scatter between different measurements in the literature given by (78). The specific gas infall rate of the same galaxy sample of our model is over-plotted in green. The sSFR follows this quantity with an offset (boost) as discussed in L13. Furthermore the specific mass increase rate of the halo (sMIR) is over-plotted.

tension in the model that the SFRD is too low at $z = 2$. The size of the discrepancy is roughly the same as for the sSFR(z) evolution. We return to this below.

2.3.3 The evolution of the gas fraction in galaxies

In Figure 2.4, we plot the gas-to-star ratio $\mu = M_{\mathrm{gas}}/M_s$ as a function of stellar mass for different redshifts. We get about a factor of six higher gas-to-star ratio at $z \sim 4$ compared to $z = 0$. From the definition of the regulator quantities in L13, the gas ratio is simply given by the ratio of the sSFR and the star-formation efficiency ϵ

$$\frac{M_{\mathrm{gas}}}{M_s} = \frac{sSFR}{\epsilon}. \tag{2.18}$$

So the increase in the gas ratio is a direct result of the fact that the halo growth rate and thus the sSFR increases faster with redshift than the dynamical time of the galaxy which was assumed to set the redshift evolution of ϵ. Lowering the gas fraction in high redshift galaxies can be
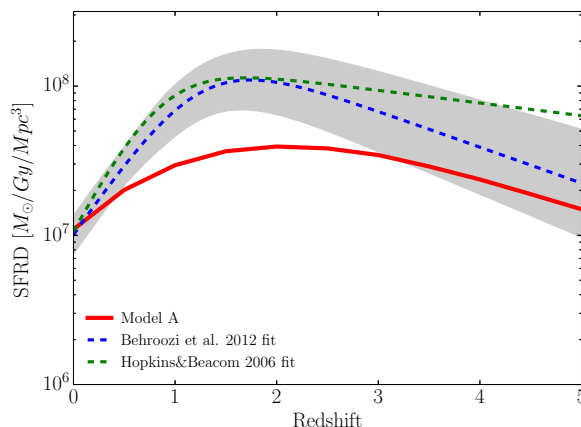
Figure 2.3: The Star Formation Rate Density (SFRD) from our model sample as a function of redshift (red line). The SFRD is the integrated SFR over all galaxies at a certain cosmic time, normalized to unit volume. The blue dashed line is the best fit model of (78) and the gray region is the 1-σ inter-publication scatter noted by them. The green dashed line is the best fit of the data compilation of (136). Our model predicts the right normalization at $z = 0$ and the drop in the SFRD at late times. Our model does not reproduce the boost in SFRD at $z = 2$ in its full strength.

done in two different ways: One either lets the efficiency ϵ increase faster with redshift or as a higher power of the gas mass within the regulator. These have similar effect because of the higher gas fractions at high redshift.

In our model A, the gas infall rate Φ_b drops faster with cosmic time than the star formation efficiency ϵ and therefore galaxies become less gas-rich at later cosmic times (a similar argument was drawn in (87)). This behavior is in qualitative agreement with observations (e.g, 137; 138).

2.3.4 Stellar Mass Function (SMF)

The galaxy stellar mass function (SMF) is a well measured quantity at low redshifts (e.g, 139; 140; 83; 141). Our model provides predictions for the overall SMF and also for the population split into blue and red galaxies (i.e. star-forming and quiescent) and into centrals and satellites. As noted above, the model is constructed to reproduce the characteristic Schechter cutoff of the blue population at $M^* \sim 10^{10.68} \mathrm{M}_\odot$ and for this to be constant with time, but we have not introduced any other parameter that is based on e.g. the faint end slope of the blue and red population, or the red fraction at M^*). The mass quenching law of P10 can directly predict the
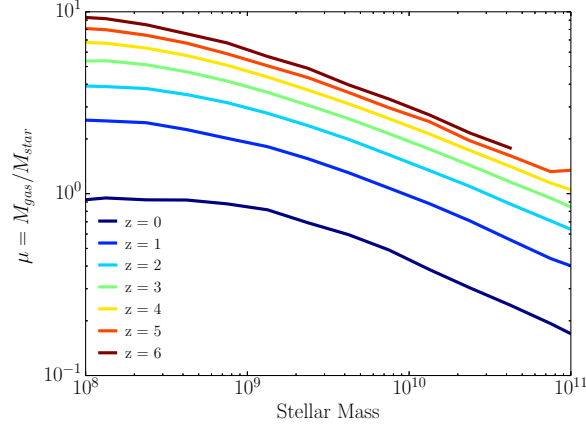
49

Figure 2.4: The gas-to-star ratio is plotted as a function of stellar mass at different redshifts for Model A. With our default regulator parameters, this ratio is increased at $z = 4$ by about a factor of six relative to locally.

relative faint end slopes of the blue and red population, but the absolute slope $\alpha_{s,blue}$ of the blue population had to be assumed. The red fraction at M* also follows from the input $\alpha_{s,blue}$.

In Figure 2.5 the model prediction for the blue, red and total population at $z = 0$ is plotted, while in Figure 2.6, we present our results for the evolution of the SMF's for different galaxy types (split into red and blue and into central and satellite) over cosmic time. The red satellite population can be better described by a double Schechter function. As shown in P12, this is due to superposition of mass- and satellite-quenching.

The model successfully reproduces the correct faint end slope of the mass function. This is a reflection of the link between the slope of the mass-metallicity relation and the faint-end slope α of the mass-function (see L13 for discussion). The relations between the Schechter parameters (M^* and α) of the different populations in Figure 2.6 are also as observed. The universality of M^* (all populations have very similar M^*) and the change in faint end slope $\Delta\alpha \sim 1.0$ between blue and red centrals, are also successfully reproduced. These follow from the forms of the quenching laws derived in P10 and P12.

Less trivial is the overall normalization of the SMF of the different populations. The ϕ^* describes the normalization at M^* in the Schechter function fits. The SMF is the convolution of the stellar-to-halo mass relation (SHMR), including its scatter, with the underlining halo mass function. We note that the underlying halo mass function is Press-Schechter like and
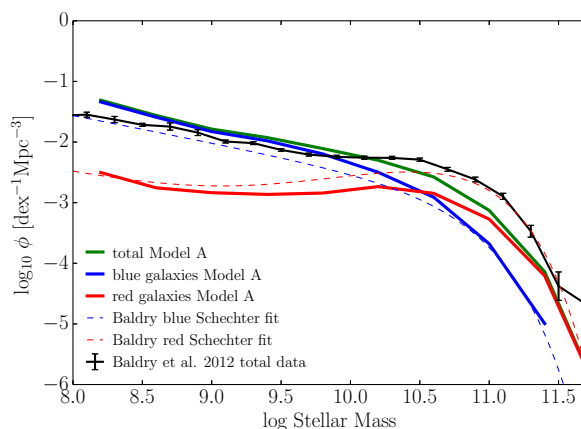
Figure 2.5: The Stellar Mass Function at $z = 0$ is plotted for Model A. The green line is the overall SMF from our Model A. The blue curves are for the blue population, the red curves for the red population (including centrals and satellites). The output is compared to the data of (141). Dashed lines corresponds to Schechter fits to the blue and red population in their paper.

not Schechter like. If we do not apply the mass quenching description, the SMF prediction would look Press-Schechter like and would have a rapidly evolving characteristic mass. At very high redshift, where the galaxy population could not build up a significant fraction of galaxies with stellar masses above M^*, we predict a Press-Schechter like SMF. In our model we see that the transition from a Press-Schechter to a "vertically evolving" Schechter-like SMF happens between $z = 6$ and $z = 4$ (from Figure 2.6). It is the moment when the stellar mass functin breaks away the halo mass function. (142) referred to this as the Phase 1 to Phase 2 transition. We can also clearly see that the satellite population grows more rapidly with cosmic time than the one of the centrals in Fig 2.6. This means that the special role of the quenching of satellite galaxies becomes more and more important with cosmic time. The satellite-quenching leads to the double-Schechter component in the SMF of the red population. The differential rate of quenching of the two populations and the fact that the quenched satellites dominate at lower masses leads to the appearance of "down-sizing" , i.e. a more gradual buildup of the stellar mass-function at lower masses.

The biggest problem with the mass functions is a surprising one. Although the shape of the mass function of passive galaxies is right, their overall number density is too low. This also produces a weaker bump in the "double" Schechter function that is caused by the superpo-
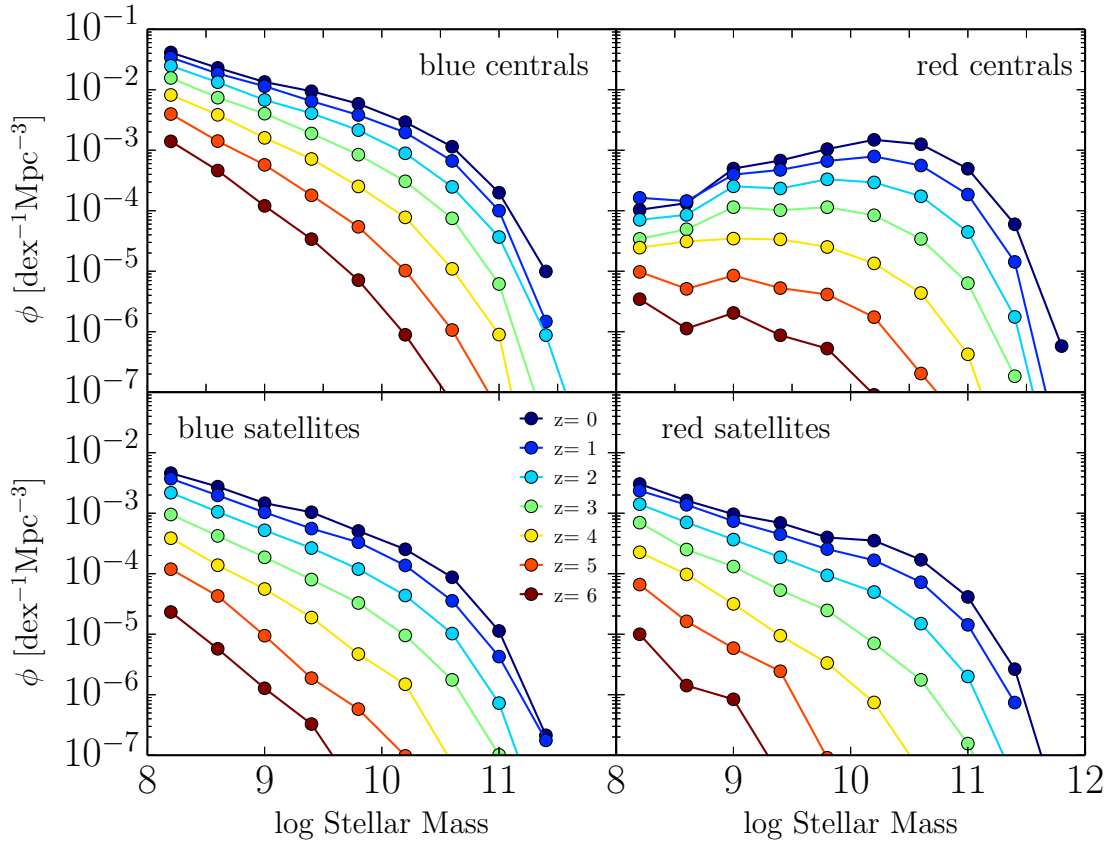
Figure 2.6: The SMF for the blue and red population, centrals and satellites are plotted for redshifts $6 \leq z \leq 0$ for Model A. Our model predicts a nearly constant fraction of blue and red population, centrals and satellites back to at least $z = 4$.

sition of the red and blue SMF (which have different faint end slopes α). This is surprising because one of the great successes of the P10/P12 quenching formalism was to explain, via the continuity equation, the ratio of these two components, which is given simply as $(1 + \alpha)^{-1}$ where α is the faint end slope of the star-forming mass function. For $\alpha \sim -1.4$ this would predict a ratio of about 2.5, close to what is observed, whereas our model predicts more like 1.5 . But we clearly note that with $\alpha \sim -1.5$ (our Schechter fit) the ratio goes already down to about 2.0 . We will return to discuss this interesting question further in Section 2.4.

2.3.5 Star formation rate history in different mass haloes and the evolution of the star-formation rate density

We now turn to comparisons with the phenomenological model of (78). In Figure 2.7, we show our prediction for the SFR in haloes (including centrals and satellites) as a function of cosmic time and halo mass. This may be compared with the similar Figure 4 from (143) which was derived from their completely different but similarly phenomenological approach.

(78) concluded that most stars were formed around $z = 2$ in haloes of about $10^{12} M_\odot$. This is a natural output of our model as the regulator is highly inefficient in producing stars at low stellar masses and (mass-)quenching is most effective above $M_s = M^*$, which corresponds to about $10^{12} M_\odot$ in halo mass.

The fact that these two orthogonal approaches produce broadly the same phenomenological picture is very reassuring. It furthermore emphasizes the operational difficulty of distinguishing, for central galaxies, whether the dark matter mass or the (baryonic) stellar mass is driving the variable efficiency with which haloes convert baryons into stars, simply because these two quantities are tightly linked.

2.3.6 Stellar-to-halo mass relation (SHMR)

One of the central properties of galaxies is the stellar-to-halo mass relation (SHMR), both for centrals and for satellite galaxies. The SHMR represents the overall efficiency with which haloes convert baryons into stars. This quantity has been extensively studied using abundance matching and other statistical techniques such as halo occupation distributions, which are
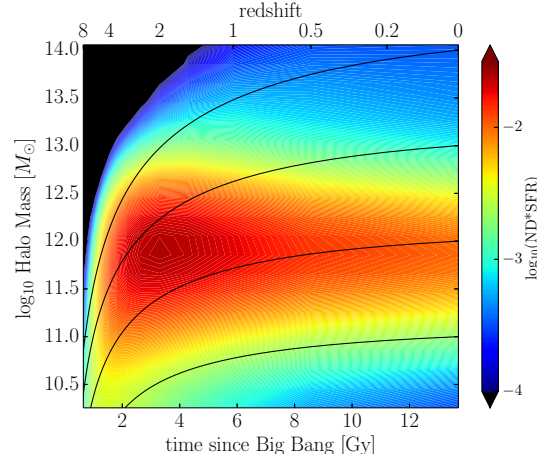
Figure 2.7: The Star Formation Rate history as a function of cosmic time and halo mass for Model A. The plot includes SFR from centrals and satellites. Black lines indicate an average growth history of different haloes. Units of the color scale are chosen to be ND*SFR $\mathrm{dlog}_{10}(M_h) = M_\odot yr^{-1}Mpc^{-3}\mathrm{dlog}_{10}(M_h)$.

based on the conviction that the SHMR should be well-behaved. Observations using weak-lensing can be used to directly test these, generally with success (e.g, 144).

The SHMR for our output sample at the present epoch is plotted in Figure 2.8 and compared with the zero-redshift relation from (78). As would be expected, the increase in the M_s/M_h ratio at low masses simply reflects the increasing efficiency of converting baryons to stars (i.e. f_{star} in L13) in more massive regulators, while the turn-over and subsequent decline is due to the mass-quenching of galaxies which becomes progressively more important at masses around and above M^*, corresponding to about $10^{12}M_\odot$ in halo mass.

The 1-σ scatter in the SHMR of the blue population in the model is about 0.21 dex. This comes mostly from the different halo assembly histories (e.g, the time when the last major merger happened). The scatter in the red population is larger and is about 0.36 dex. This ultimately reflects the quite broad range in stellar (or halo) mass over which central galaxies have been mass-quenched and the continued growth of haloes after the star-formation has been quenched.

Red galaxies have systematically lower M_s/M_h than blue ones at a given M_h because their stellar masses are frozen at quenching (apart from mass growth due to merging) while their dark matter haloes continue to grow. They may scatter down the $M_s \sim M^*$ locus. This scatter

explain the observation (e.g, 145) that at a given stellar mass, red galaxies are found in higher mass haloes (e.g. with more satellites). As the overall population of central galaxies changes from predominantly blue at low halo masses to predominantly red at higher halo masses the mean SHMR shifts from that of the blue galaxies to that of the red. The overall scatter is expected to be 0.32 dex at the peak but deviates from being a log-normal distribution in stellar mass.

Overall, the agreement between the output of our model and the reconstruction from (78) is very good. Our curves for the overall population are slightly lower around the peak, by up to about 0.2 dex at halo masses above $10^{11.5} M_\odot$ and this can be traced to the saturation of f_{star} in L13, which itself was driven by the saturation in the adopted $Z(M_s)$ mass metallicity relation. We will return to this point below and show that it is closely linked to the issue of the deficit of quenched galaxies noted in Section 2.3.4.

Our model has a slight redshift evolution in the SHMR (see Figure 2.9). Within our model, this is due to the fact, that regulators (i.e. galaxies) at higher redshifts contain proportionally more gas and thus less stellar mass as discussed in Section 2.3.3. But the general behavior remains at all redshifts the same. At very low halo masses, the stellar content remains dominated by the maximum outflow load $\lambda_{\max}$ and the saturation feature occurs at every redshift at roughly the same halo mass. The nominal drop in the SHMR at $z = 4$ is about a factor of two.

2.4 Discussion

In Section 2.3 we recovered a number of encouraging agreements of various predictions compared to the literature, both in terms of observational data and in terms of the independent and orthogonal phenomenological model of (78). In particular there is no reason for the total number density of galaxies to come out right. The models and the parameters taken from the previous papers (P10, P12, L13) did not have any information about the abundance of dark matter haloes nor were designed to match the number density of galaxies in the universe. This is a remarkable success of our model. The model is simple but still reproduces a wide range of non-trivial results. In this section, we will have a closer look at those areas where our model produces discrepancies that may give clues as to where additional features could be added, or
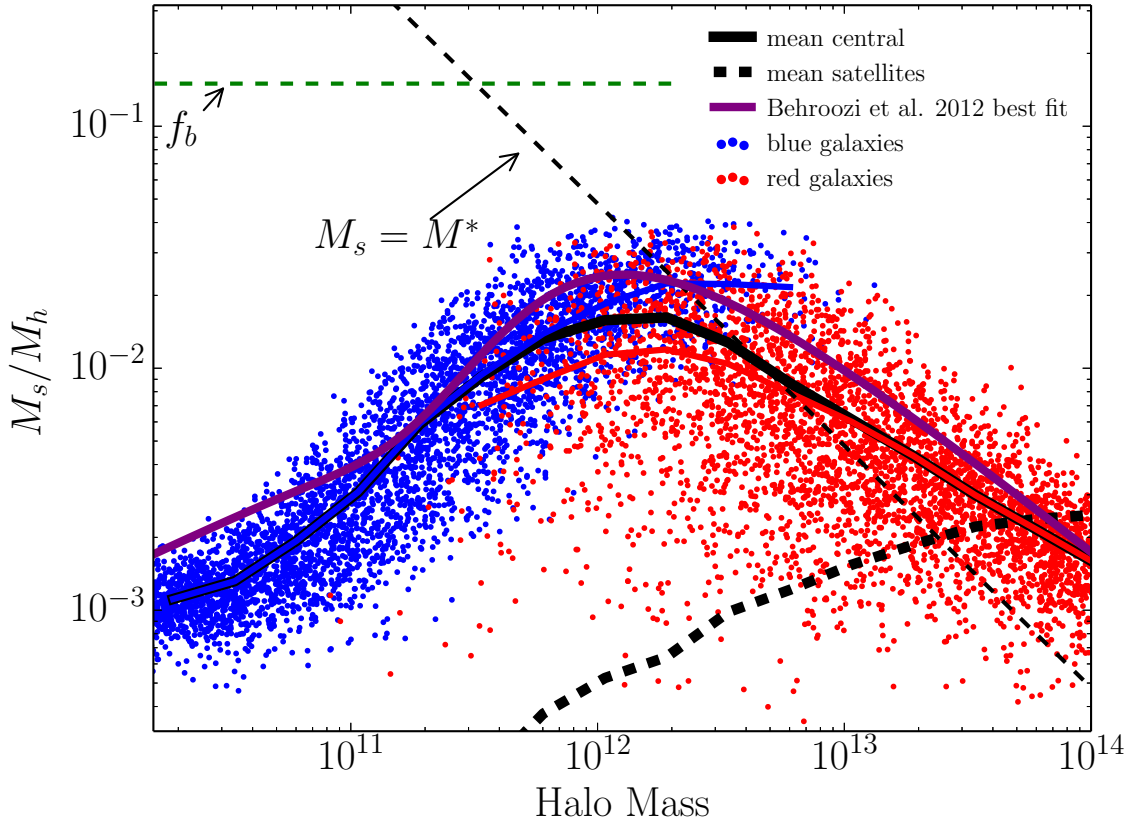
Figure 2.8: The SHMR at $z = 0$ of Model A is plotted as a function of the total halo mass M_h for the set of central galaxies, separated into red and blue. The blue (red) continuous line is the mean value of the blue (red) population in our model and the black line is the mean SHMR of the overall sample for centrals (i.e. a suitably weighted average of the red and blue lines). The thick dotted black line is the contribution of satellites to the SHMR while the green thin dotted line indicates the cosmic baryonic fraction. The global turn-over of the star formation efficiency can fully be accounted by quenching galaxies around M^*, corresponding to about $10^{12} M_\odot$ in halo mass. The agreement with the abundance matching reconstruction of (78) is quite impressive, although there is a systematic reduction in M_s/M_h above $M_h \sim 10^{11.5} M_\odot$ which may be traced to the saturation of the efficiency with which the regulator in L13 converts baryons to stars that is in turn linked to the flattening of the $Z(M_s)$ mass-metallicity relation.
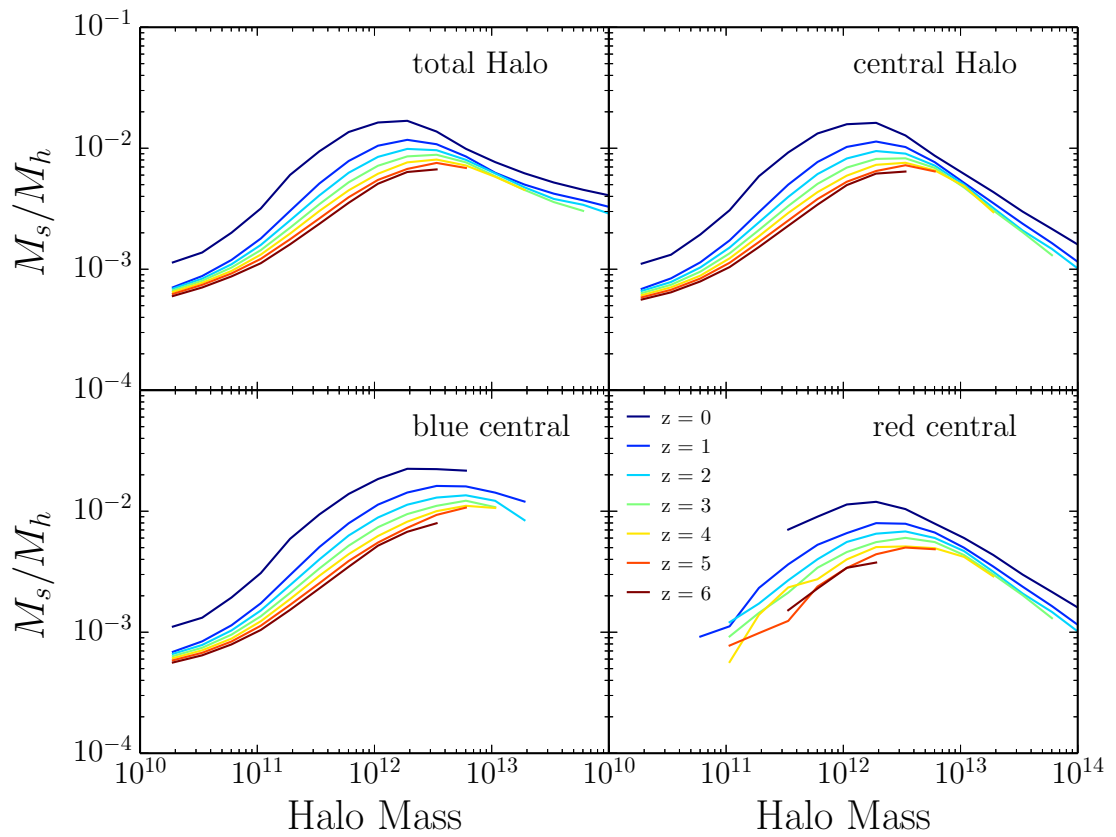
Figure 2.9: The average SHMR of Model A for different galaxy types as a function of total halo mass. Upper left: Considering all galaxies within a halo and all types of galaxies. Upper right: Considering only the central galaxy (red and blue). Lower left: Considering only blue central galaxies. Lower right: Considering only red central galaxies. The scatter in this relation can be deduced from Figure 2.8. The SHMR is predicted to be an increasing function in cosmic time (i.e. decreasing with increasing redshift) because regulators at high redshift are more gas rich.

which may highlight more fundamental tensions.

First we will have a look at the specific star formation rate evolution and note how we can, in principle, achieve a better agreement with the data compilation of (131) and (78) in Section 2.4.1. But then, relating the sSFR evolution to the SFRD evolution, we argue that we can not easily bring these two observations in agreement with each other, independent of our model assumptions (Section 2.4.2). We then turn our attention to the missing red galaxies. We discuss how this is linked to the form of the SHMR in Section 2.4.3 and we discuss its relation to the saturation feature of the L13 regulator model. In Section 2.4.4 we propose two other ways of assigning the gas in-flow to the galaxies within the halo and see that we get a further improvement in matching the SMF, sSFR and SFRD history, our Model B and C. In Section 2.4.5 we discuss a very specific feature of our models and finally in Section 2.4.6 we relate our results to abundance matching methods.

2.4.1 Modification to match the sSFR at z=2

In Section 2.3.2 and in Figure 2.2 we noted a deviation of the sSFR evolution at $z = 2$ between our predictions and the data compilation of (131) and (78) . It might be thought that one possible way of modifying our model to try to get a better match is to change the star-formation efficiency, ϵ, at high redshift. Detailed discussion about the link between star formation and gas reservoir has been made by several authors (recently e.g, 99). However, because the link between sSFR and the sMIR (specific mass accretion rate of the system) is independent of ϵ and λ (see L13, and thus also of f_{star}), modification of $\epsilon(z)$ changes the sSFR only through the "boost" effect on sMIR that is associated with a change in f_{star} with time and so the effect of this change should be quite weak. It turns out that a higher ϵ at high redshift leads to a drop in the offset of sSFR compared with the sMIR. To explore this, we modify the parameterization of ϵ to:

$$\epsilon(z) \propto (1+z)^c \qquad\qquad (2.19)$$
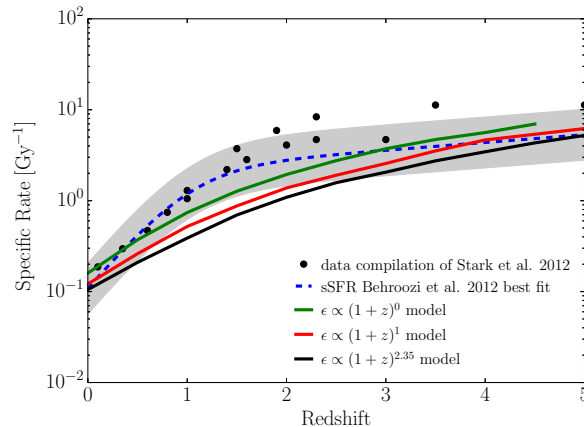
Figure 2.10: The sSFR history for three different variants of the redshift evolution of the star-formation efficiency ϵ with $c = 0, 1, 2.35$ (as defined in Equation 2.19) are plotted. Stronger evolution (higher c-values) lead to a lower sSFR at all times.

with c being the additional model parameter. In our default Model A (also Models B and C below), the efficiency scales as the Hubble rate. In Figure 2.10 we plot three different models with $c = 0, 1, 2.35$, i.e. assuming no redshift evolution, one coming close to the fiducial model and one in which the efficiency scales as the sMIR according to (133). We note that at fixed redshift, the efficiency is parameterized as a function of M_s. This parameterization is fitted at $z = 0$ and might not provide a direct link to the physical process that actually sets the efficiency.

We clearly see that *lowering* the star-formation efficiency at higher redshifts actually *boosts* the sSFR. This is because it lowers f_{star} at high redshifts and therefore increases the boost term in Equation 36 of L13. On the other hand if the efficiency increases with redshift as fast as the specific infall rate, we reduce the sSFR. In both cases, the effect of the change in the sSFR is spread out over a wide range of redshifts (because of the smooth evolution in ϵ) and we cannot get a peak at one particular redshift, or drastically change the overall slope.

An alternative approach is to decouple the specific accretion rate onto the regulator systems from the specific growth rate of the surrounding dark matter haloes. A redshift dependent cold gas accretion efficiency (e.g, 86) could do this, or some other scheme to limit the baryonic accretion onto the regulators. In Section 2.4.4 we will explore some modifications by introducing Models B and C.
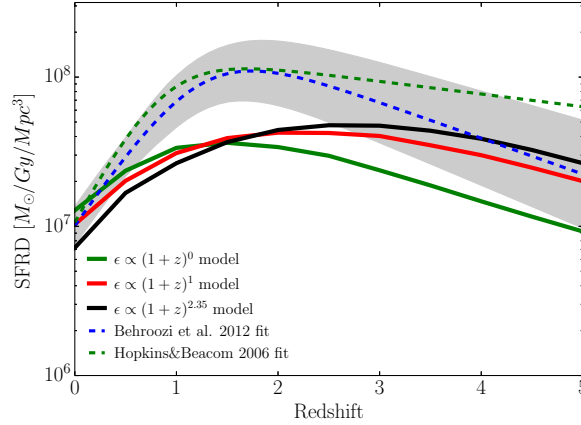
Figure 2.11: The SFRD history for the three different variants of $\epsilon(z)$ with $c = 0, 1, 2.35$ (defined in Equation 2.19). Stronger evolution (higher c-values) leads to an enhancement of the SFRD at early times and a stronger decrease at later times.

2.4.2 The link between sSFR and SFRD

Staying with the same expansion in our model as in Section 2.4.1 we turn our attention to the star formation rate density (SFRD). We plot in Figure 2.11 the SFRD history for the same three models as for Figure 2.10. The figure shows that lowering the efficiency at high redshift shifts star formation to later times. The redshift dependence of the efficiency ϵ does not have a significant influence on the outcome at $z = 0$. It has a slight effect of where the stellar mass is formed. As the model has a smoothed evolution in ϵ, significant deviations in the sSFR history from our default model can not be made.

2.4.3 Matching the red fraction at M^*

As mentioned in 2.3.4, our model under-predicts the abundance of red galaxies around M^*. In other words, the relative fraction of red to blue galaxies is too low.

The number density of red galaxies around M^* is directly related to the number of dark matter haloes between $M_h(M^*)$ and infinity. As the halo mass function is a very steeply decreasing function of halo mass, the number of red galaxies around M^* is very sensitive to the halo mass $M_h(M^*)$ that corresponds to the quenching mass M^*.

However, simply changing the parameter M^* (i.e. μ^{-1}) will have a severe impact on the blue

population that we match very well. Boosting the SHMR (e.g, by just letting more gas flow in the regulator) is also not satisfactory. By doing so, we will boost the number density of blue and red satellites by the same amount. We would be able to get the needed number density in the red population around M^* (as we lowering the halo mass corresponding to M^*) but at the same time we would end up with to many blue galaxies at the same stellar mass range. The question is: How can one change the red fraction without either changing the number density of the blue population or M^*? The fraction between blue and red galaxies around M^* is dependent on how fast galaxies are approaching M^*. We have to elevate the sSFR at M^* or in terms of the SHMR, the power law parameter for the Main sequence γ defined as

$$M_s \propto M_h^\gamma \qquad\qquad (2.20)$$

has to be steeper around M^* than our model prediction.

Our model produces a flattening of the SHMR around M^* (see Figure 2.8 for $z = 0$ and Figure 2.9 for the redshift evolution). This is an intrinsic feature of the regulator model and independent of quenching. The overall fraction of baryons in stars cannot exceed the cosmic fraction, and indeed can only asymptotically approach this. In fact, because of the "loss" of outflowing gas in this first Model A, it will saturate at an even lower value. The regulator f_{star} saturates when the gas within the halo is nearly used up.

We note that our model, even without any quenching mechanism, therefore has a saturation feature coming from the regulator because f_{star} is limited to some value. Our model predicts just at the stellar mass when quenching happens a flattening of γ due to the saturation. In contrast, we get a better match to the red population when abandoning the saturation feature or invoking an even steeper γ at M^*. This might provide a hidden link between the quenching process and the running out of gas of the galaxy. We return to this below.

2.4.4 Changing gas in-flow description

One of the weaknesses of our models is that we do not trace the out-flowing gas. The need for gas reincorporation in a cosmological context was initially analysed in (102; 103). Other recent works include (105; 146; 147). In our simple model, we don't allow the expelled gas to

get back into the same regulator or transfer it to another regulator sitting in the same dark matter halo. Letting some or all of this gas back into the regulator system will change the output of our model significantly. We note that at stellar masses around M^*, about 1/2 M^* of gas has been ejected earlier in the history of each galaxy. There is only a slight dependence of this on the adopted value of the parameter λ_{max}. From our discussion in Section 2.4.3, the saturation feature leads to a mismatch of the red population. To delay the saturation of our regulator to higher stellar masses above M^*, we might just put some of the ejected gas back into the regulator at the time when saturation occurs. This process can in principle be accomplished by setting an appropriate recycling time (of order several dynamical times). Such a behavior can consistently be applied to our model. The only worry is that this new type of metal-enriched inflow will significantly change the metallicity-fitted parameters inferred in L13 and used in our combined model. This might indicate that the metallicity modeling might be unrealistic.

Some gain in the direction can be achieved by simply modifying how gas is assigned to the regulators. In combining the different models of Section 2.1 we have a freedom in assigning the gas in-flow to the different galaxies (central or satellites). So far in our Model A we have assigned the gas according to the weights of the (sub)haloes (Equation 2.15) with the weight of the central given in Equation 2.14. The substructure fraction is increasing with halo mass and therefore the second term in Equation 2.14 assigns a smaller proportion of the infallen gas to the central galaxy as it grows in stellar mass. This can also contribute to the flattening of the SHMR.

The Model A assumed no domination of the central galaxy over its satellites at all. The other extreme would be the central galaxy dominates completely and gets all the gas in-flow and the satellites do not get any gas infall at all. Our Model B which we present here is identical to our Model A except that Equation 2.15 is changed so that all of the incoming gas is given to the central galaxy:

$$\Phi_{b,i} = \begin{cases} f_b \dot{M}_{h,\text{smoothed}} & \text{central} \\ 0 & \text{satellite.} \end{cases} \quad (2.21)$$

Figure 2.12: The same plot as Figure 2.3. The SFRD for Model A of Section 2.3 (red), Model B (black) and Model C (green) are compared with data compilations. Model B achieves some boost around $z = 2$ compared to Model A, but only Model C achieves the required amount of boost.

The result in terms of the SFRD is plotted in Figure 2.12. We clearly see an additional boost in the SFRD around z=2 or even at higher redshift. This brings the model closer to what is required by the data. The reason for the difference between the two proposed models is that at high redshift the halo merger rate is very high compared to the subhalo decay rate. This leads to more substructure within a halo at high redshift. In our Model A this leads to less gas in-flow onto the central galaxy, which is avoided in Model B. Furthermore the gas infallen onto the central galaxy is turned into stars more efficiently than in (lower mass) satellites. But despite this improvement, the Model B still under-predicts the SFRD at $z = 2$.

In terms of the sSFR history we do not get any change in the predictions form Model A to Model B, as presented in Figure 2.13. To match the sSFR history, we have to change the model further.

Looking at the SMF at $z = 0$ predicted by our Model B in Figure 2.14 we can also partially improve matching the red fraction around M^*. A discrepancy remains, however, coming from the regulator description as discussed in Section 2.4.3. The SHMR of Model B (Figure 2.15) for central galaxies is similar to Model A and also comes close to the Model of (78).

Out of this discussion, we see the importance of how one assigns the gas in-flow to the different galaxies within a halo. But we want to emphasize that no complicated description
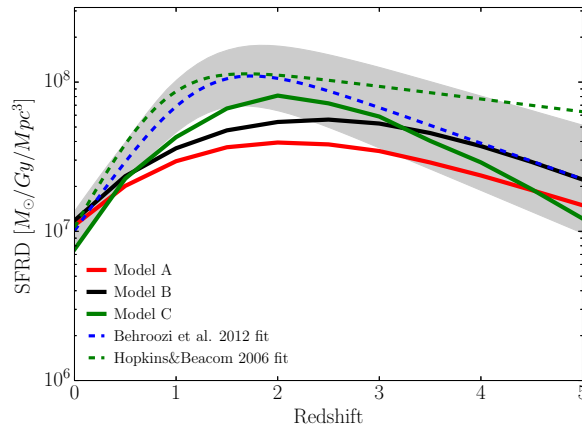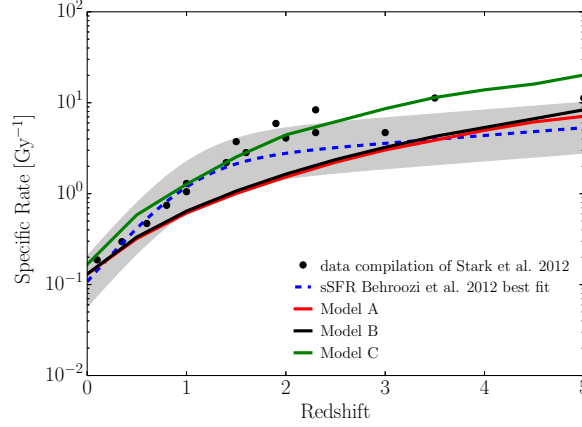
Figure 2.13: The same plot as Figure 2.2. The sSFR for Model A of Section 2.3 (red), Model B (black) and Model C (green) are compared with data compilations. Model B is very similar around $z = 2$ compared to Model A but only Model C achieves the required boost.

(e.g recycling of outflown gas, decoupling of baryonic inflow and dark matter growth, ...) is needed to achieve the level of agreement that is already presented in Model A and B. In terms of the quenching "laws", they are instead to be purely descriptive. These laws would likely be more complicated if they were formulated in terms of physical mechanisms which are still unclear.

Having said that, the red fraction problem and the sSFR and SFRD at $z = 2$ still do not match perfectly. Our next approach is the one of an 'effective SAM". From our discussion above, we concluded that the gas inflow description is crucial in perturbing our model and, doing it in the right way, matching the observables. For our Model C we introduce a redshift and halo mass dependent gas inflow. We change equation 2.15 to the form:

$$\Phi_{b,i} = \begin{cases} f_b \dot{M}_{h,\text{smoothed}} \cdot f_a(a) \cdot f_M(M_i) & \text{central} \\ 0 & \text{satellite,} \end{cases} \tag{2.22}$$

with

$$f_M(M_h) = 1 + 30 \cdot \left(\frac{M_h}{10^{12} M_\odot} \right)^{2.5} \tag{2.23}$$

Figure 2.14: The Stellar Mass Function at $z = 0$ is plotted for Model B. The green line is the overall SMF. The blue curves are for the blue population, the red curves for the red population (including centrals and satellites). The output is compared to the data of (141). Dashed lines corresponds to Schechter fits to the blue and red population in their paper.

Figure 2.15: The SHMR at $z = 0$ of Model B is plotted as a function of the total halo mass M_h for the set of central galaxies, separated into red and blue (same as Figure 2.8 for Model A). The blue (red) continuous line is the mean value of the blue (red) population in our model and the black line is the mean SHMR of the overall sample for centrals (i.e. a suitably weighted average of the red and blue lines). The thick dotted black line is the contribution of satellites to the SHMR while the green thin dotted line indicates the cosmic baryonic fraction.

and

$$f_z(z) = \begin{cases} -1.25 \cdot (1+z)^{-1} + 1.4 & z < 2 \\ 0.25 + 6.75 \cdot (1+z)^{-3} & z >= 2. \end{cases} \qquad (2.24)$$

The functions f_M and f_z are arbitrary and designed to have four desirable features:

1. f_z is a decreasing function between $z = 2$ and $z = 0$ accounting for the steep decline in the SFRD.

2. f_z is a rapidly increasing function approaching $z = 2$ accounting for the boost in the sSFR around $z = 0$.

3. f_M has an additional term such that there is significantly more gas inflow onto massive galaxies around $M_h(M^*)$ to counter-act the saturation feature of the regulator.

4. $f_z \cdot f_M$ is normalized such that the baryonic mass within the regulator never exceeds the cosmic baryonic fraction of the universe.

The functional form of f_z and f_M are completely arbitrary. The functions and values are chosen to match the four criteria mentioned above. We want to emphasize that a priori no physical argument was chosen to justify our approach except their result on the observables mentioned above. Recently (105; 146; 147) provided physical pictures or reincorporation of gas and (e.g, 100) discussed extensively the impact of different physical processes on the evolution of the SFRD. The SFRD of Model C is plotted in Figure 2.12 in red. We get about a factor of ten difference in the SFRD at $z = 2$ and at $z = 0$. The sSFR gets an additional boost at $z = 2$ (red line in Figure 2.13) and the SMF at $z = 0$ does match very well all the different galaxy populations in shape and amplitude (Figure 2.16). The resulting SHMR plotted in Figure 2.17 looks very different. The blue population is approaching the cosmic baryonic fraction very rapidly but gets quenched just before exceeding the limit (in stellar mass).

This extension can not be considered as a "best fit" model. The aim is just to indicate the power of this specific extension for future model buildings. Other predictions such as the gas-to-star ratio are only marginally affected by this extension. We will not break the degeneracy between
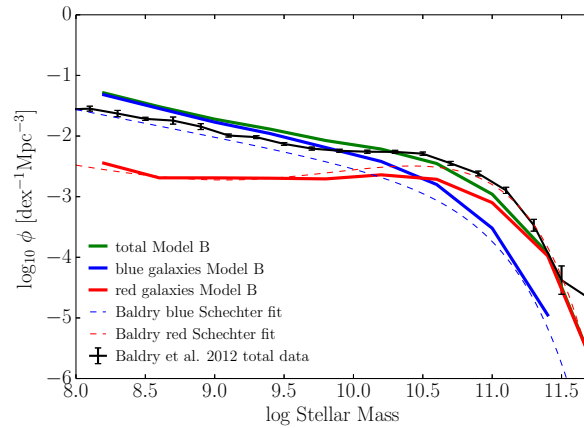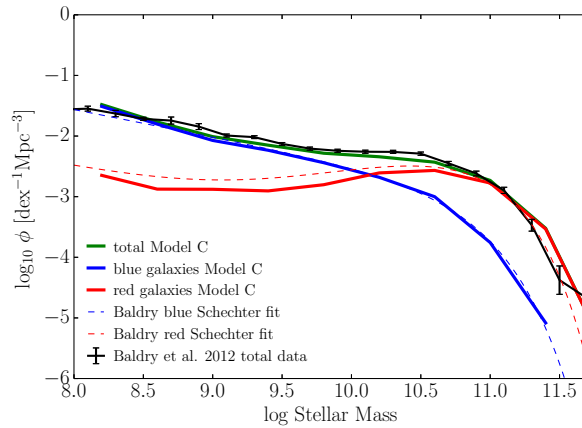
Figure 2.16: The Stellar Mass Function at $z = 0$ is plotted for Model C. The green line is the overall SMF. The blue curves are for the blue population, the red curves for the red population (including centrals and satellites). The output is compared to the data of (141). Dashed lines corresponds to Schechter fits to the blue and red population in their paper.
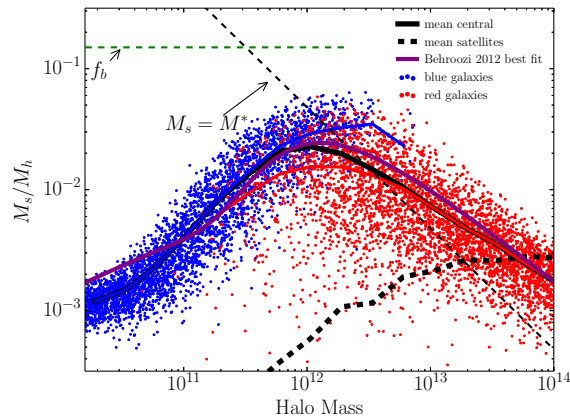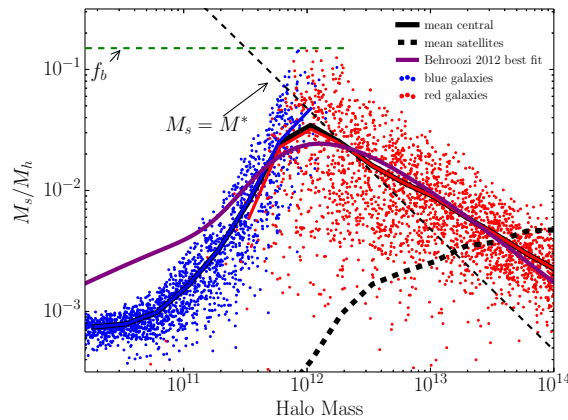

Figure 2.17: The SHMR at $z = 0$ of Model C is plotted as a function of the total halo mass M_h for the set of central galaxies, separated into red and blue (same as Figure 2.8 for Model A). The blue (red) continuous line is the mean value of the blue (red) population in our model and the black line is the mean SHMR of the overall sample for centrals (i.e. a suitably weighted average of the red and blue lines). The thick dotted black line is the contribution of satellites to the SHMR while the green thin dotted line indicates the cosmic baryonic fraction. This model best matches the abundance of red and blue galaxies. Quenching occurs just when the blue galaxies are approaching the cosmic baryonic limit.

recycled and newly infallen gas components with this extension of our model. Metallicity and HI data (see e.g, model of 148) might give further insights into this processes.

2.4.5 The coincidence of getting quenched when approaching the baryonic fraction

We notice from our analysis in Section 2.4.3 and 2.4.4, the SHMR is far below the cosmic baryonic fraction f_b at low M_s and is coming closer to f_b when approaching M^*. By "coincidence", quenching occurs in our model just when the stellar baryonic fraction approaches the cosmic fraction f_b. In our model, the regulator is not allowed to get more baryons in than the baryonic fraction (see Equation 2.15)and so will automatically saturate. It will no longer follow the power law description of Section 2.4.3 and will flatten. In our model this saturation feature is completely independent of the quenching formalism with its crucial parameter M^*.

However, apparently as a"coincidence", these two completely different features arise at the same point in the evolution history of a star-forming galaxy. It is ultimately this simultaneous appearance of these two features that led to the under-prediction of the red population around M^*. In our Model C, we see that to match the SMF we even have to steepen the SHMR of the blue population around M^* such that the blue population must approach the cosmic baryonic limit even faster, without apparently noticing it, but suddenly then quench just before reaching the ultimate limit.

If one has one mechanism suppressing star formation in low mass galaxies and quenching at high masses, a peak is inevitable. But the peak in M_s/M_h that is caused by quenching could have occurred at any mass, e.g. if it was driven by AGN feedback, morphological effects and so on. The fact that it appears to occur just when the overall efficiency of converting of baryons into stars is maximal is, in our view, noteworthy and probably tells us that it is not a coincidence.

2.4.6 Abundance matching

We note from Figure 2.8 (for Model A) and from Figure 2.15 and 2.17 (for Model B and C) that, at halo masses around $10^{12} M_\odot$, the mean value of the SHMR of the blue population is elevated

by about 0.2 dex compared to the mean value of the red population. The 1-σ dispersion in the blue population alone is about 0.2 dex, and the overall scatter in the combined red and blue populations at $10^{12} M_\odot$ is larger, 0.35 dex, and the distribution is not Gaussian in log M_s / M_h, i.e. log-normal in the ratio. Simple abundance matching techniques usually do not take into account this possible variation.

(78) noted that the range in star formation rates that is implicit in a star-forming and a passive population, is only a problem if it results in a distribution of stellar masses at fixed halo mass that cannot be reasonably modeled by a log-normal distribution (the main assumption in their work). In our particular model, we produce a clearly different distribution in stellar mass around the peak M_s / M_h. The SHMR of our Model C in Figure 2.17 is substantially different to the one of (78) but reproduces the SMF at the same accuracy. In other words, the SHMR from our Model C is effectively a kind of abundance matching, as it is specifically tuned to match abundance properties of the galaxy population, but with a different assumption (motivated by our quenching laws) of how blue and red galaxies will populate the dark matter haloes.

(74) uses measurements of the stellar mass function, galaxy clustering, and galaxy-galaxy lensing within the COSMOS survey to constrain the SHMR of blue and red galaxies over the redshift range $z = [0.2, 1]$. Their underlining assumption on the functional form of the blue and red galaxy SHMR is very different to our output. E.g. their blue population itself is described with a turn over in the SHMR.

2.5 Summary

We have presented a simple model of the evolving galaxy population that is based on importing pre-formulated baryonic prescriptions for the control of star-formation in galaxies into a dark matter halo merger tree. Specifically, the model is based on the gas-regulation model of star-forming galaxies from L13, and the empirical quenching formulae of P10 and P12.

The parameters for these baryonic prescriptions are taken directly from these earlier works and are not adjusted according to the output of the current model. A very limited number of additional a priori assumptions are however required to ensure the model can operate, but these do not greatly affect the outcome. The model allows us to make predictions about the

numbers and properties of galaxies that are independent of the observation inputs used to determine the model prescriptions in the previous papers and which can therefore be used to test the model. The observational inputs to the previously tuned parameters were: The exponential cutoff scale M^* of the main sequence galaxies at z=0, Z(M_s,SFR) data at z=0 and the averaged enhanced fraction of red galaxies in groups and clusters. The only input from the dark matter picture in the previous papers (namely L13) was the average halo growth rate of (133).

The output of this model is compared with independent observational data and also with other recent phenomenological models (78) for the evolving galaxy population that have been based on epoch-dependent abundance matching of haloes and galaxies. Output quantities examined include (a) the Main Sequence sSFR-mass relation; (b) the integrated star-formation rate density (SFRD); (c) the stellar mass functions of star-forming and quenched galaxies; (d) the $M_s vs. M_h$ relation and $SFR - M_h$ relations as well as the epoch dependence of these over the whole redshift interval $0 < z < 5$. The predicted gas content of galaxies is also presented.

The goal of this work has been to see how far we can get with this simple model and to explore how it may need to be adjusted so as to rectify any failings in reproducing the real Universe.

We have drawn the following conclusions out of this work:

1. Reassuringly, the attractive features of the input baryonic prescriptions that were high-lighted in the original papers, including the mass-dependence of the Main Sequence sSFR, the faint end slope of the galaxy mass function, the relative Schechter M^* and α parameters of the blue and red (star-forming and quiescent) galaxy populations are certainly all preserved when transplanted into a realistic dark matter structure. The argument of L13 in relating the faint end slope α from the regulator scaling laws does not suffer from the limitations of a single mean sMIR. The mass-function of star-forming galaxies is also well reproduced and the general form of the $SFR - M_h$ and $M_s - M_h$ relations are very similar to those constructed by (78) and arise from the competition between the increased efficiency of turning baryons into stars as the mass increases (due to lower mass loss in winds) and the quenching of star-formation in galaxies. The

overall forms of the sSFR(z) and SFRD(z) are also qualitatively produced by the model. These are major and rather striking successes from an simple model that are very largely independent of the original observational inputs that were used previously to define our baryonic prescriptions.

2. As with other models in the literature, our simplest model has quantitative difficulty in reproducing the steep increase back to $z \sim 2$ in both the sSFR(z) and SFRD(z). This cannot be solved by simple adjustments to the adopted star-formation efficiencies. We also find that the peak in the $M_s - M_h$ relation is a little softer than in the (78) representation and, surprisingly, that the ratio of quenched to star-forming galaxies around $M*$ is lower than observed (and than can be predicted from the original P10 formalism). We show that the latter two issues are closely related and are due to a saturation in the efficiency with which haloes form stars that is inherent in the adopted regulator model, especially as the cosmic baryon limit is approached.

3. All four of these quantitative deficiencies can be simultaneously solved by adjusting the specific infall rate of material onto galaxies by allowing them to re-ingest material previously expelled by winds provided that this occurs in a redshift- and mass-dependent way, being most effective at masses around M^* and at redshifts $z \sim 2$.

4. Our model allow us to predict the $M_s - M_h$ relation for star-forming and quiescent galaxies separately. Red galaxies always have a higher M_h at given M_s because of the continued growth of haloes after star-formation ceases, and there is a 1-σ scatter in stellar mass of 0.36 dex for haloes of mass $M_h = 10^{12} M_\odot$ with two clearly distinguishable populations. There is significantly less scatter in the blue population than in the red one. The SHMR around M^*, where mass quenching happens, has to be steeper than predicted from our original model to match the blue and red galaxy abundances at the same time. Such a qualitative behavior brings a simple regulator model to its limits as one expects the SHMR to flatten when approaching the baryonic limit.

5. While others have emphasized the "inefficiency" of star-formation in haloes, we stress instead the efficiency of M^* galaxies in forming stars. Further, we note the "coincidence" that quenching happens in our model just at the time when the regulators are rapidly

approaching the maximum possible efficiency in covering baryons into stars, even thought these two description are completely independent of each other in the model.

Our analysis emphasizes the continued importance of pinning down as reliably as possible the bulk characteristics of the evolving galaxy population over a wide span of cosmic time. One crucial factor in our model is the gas infall onto galaxies, and it will be of great importance to trace the gas in the universe in a more observationally comprehensive way.

2.A   Appendix of chapter 2

2.A.1   Model sensitivity on additional parameters and initial conditions

Model dependence on $f_{\mathbf{merge}}$

The parameter f_{merge} in the model describes the fraction of the stellar and gas mass of a satellite that enters the central galaxy when the satellite is disrupted. In the main text, this is set to $f_{\mathrm{merge}} = 0.5$. To understand the sensitivity to this parameter we present here three models that each have the same parameters as in Table 2.1 except that f_{merge} is set to $f_{\mathrm{merge}} = 0, 0.5, 1$. There is a small change in the blue population. Even with $f_{\mathrm{merge}} = 1$, a blue central galaxy almost never has more than 20% of it's mass growth through merging. For the red galaxies at $m_s >> M*$, however, which are generally located in massive haloes (e.g. above $10^{13} M_\odot$), merging is the primary channel for mass growth and there is therefore a significant effect of f_{merge} on the the mass function and the SHMR for these most massive galaxies. The effect on the SHMR is shown in Figure 2.18. If we want to predict this quantity of the stellar mass function at these high masses, then we would have to constrain f_{merge} (or vice versa). This regime of galaxy mass is not however a central consideration of this work.

Model dependence on $\lambda_{\mathbf{max}}$

In the model, λ_{max} gives the maximum mass-loading of the wind, which is required to limit the extrapolation of the λ which varies inversely with mass at higher masses. In Figure 2.19, three models with different λ_{max} are plotted. As would be expected, there is a significant dependence for the lowest mass galaxies, corresponding to haloes below $10^{11} M_\odot$, where the
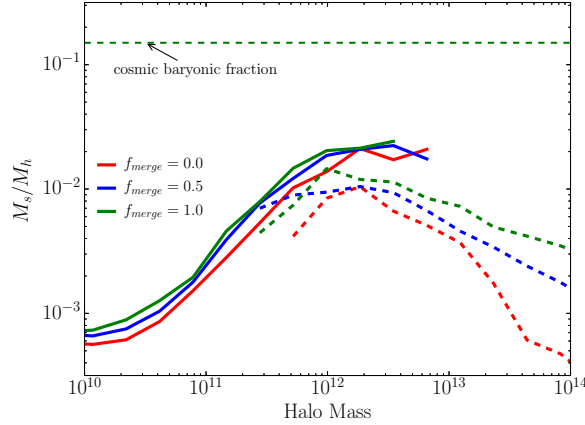
Figure 2.18: The SHMR at $z = 0$ is plotted for three different f_{merge}. Filled lines indicate the mean in the blue population and dotted lines the mean of the red populations. Increasing values for f_{merge} do only significantly affect the high halo mass end.

SHMR scales linearly with $\lambda_{\max}^{-1}$ at the very low stellar mass end. We choose $\lambda_{\max}$ in such a way that the regulator above $10^9 M_\odot$ in stellar mass is not affected by the floor value. This gives us a prior of $\lambda_{\max} \geq 20$. Values between 20-200 only affect the intermediate range marginally. We choose $\lambda_{\max} = 50$ and note that our model is not tuned to predict the SHMR below $10^{11} M_\odot$ in halo mass or $10^8 M_\odot$ in stellar mass correspondingly.

Model dependence on $M_{\mathbf{thresh}}$

The parameter M_{thresh} controls the threshold above which a (sub-) halo contains a regulator system, which in turn affects the way in which baryons are brought into the larger haloes. The dependence of our model on M_{thresh} parameter is a little more complicated. First, when lowering M_{thresh} we increase the merging component $\dot{M}_{\mathrm{merger}}$ and lower the smoothed accretion component $\dot{M}_{\mathrm{smoothed}}$. Second, the very high host-to-satellite ratio makes those low mass substructures survive very long (often longer than the age of the Universe). These results that M_{central} defined by Equation (2.14) is lowered compared to M_h. From Equation (2.15), a lowered M_{central} leads to a reduction of the infall rate in Model A. In Figure 2.20 four models with different M_{thresh} are plotted. We see a scaling difference in the SHMR. When changing M_{thresh} by four orders of magnitude, we change the SHMR by less than one order of magnitude. L13 introduced in their paper a parameter f_{gal} to account for the fact that if they let all the

Figure 2.19: The SHMR at $z = 0$ is plotted for three different $\lambda_{\max}$. In color are the mean values of the blue population. The mean values of the red fraction are plotted in red for all four models. We detect no significant deviation in the red fraction above the turn-over. The low mass population are significantly affected by $\lambda_{\max}$.

accreted baryons in their regulator they would end up with to high an SHMR. In our model, we naturally do not let all the baryons fall in the central because of the sub-structure. The parameter f_{gal} is therefore effectively absorbed into the (more physical) parameter M_{thresh}. We get an equivalent of $f_{\text{gal}} = 0.5$ with $M_{\text{thresh}} = 10^9 h^{-1} M_\odot$. This mass is consistent with photo-ionisation heating operating at low masses and suppress cooling and star formation below a certain halo mass M_γ. This halo mass scale increases from $M_\gamma \sim 10^8 M_\odot$ during reionisation to $M_\gamma \sim \text{few} \cdot 10^9 M_\odot$ (126; 127). For a more realistic model aiming to make predictions back to the epoch of reionisation, one has to account for a change in the mass threshold.

Initial conditions

When we stop expanding our merger tree (in a backward process) at either redshift $z = 15$ or at halo masses of $10^9 h^{-1} M_\odot < M_h < 2 \cdot 10^9 h^{-1} M_\odot$ we have to initialize the baryonic component of the halo. To start the forward process of the regulator system, we have to put in some initial values for M_s and M_{gas}. In principle we should start with $M_{s,\text{init}} = 0$. But with this initial condition the star-formation efficiency is zero and so the differential equation we want to solve has the solution $M_s(t) = 0$ for all times. Whether we start with $M_{s\,\text{init}} = 1 M_\odot$ or $M_{s\,\text{init}} = 10^3 M_\odot$ does not really matter when predicting the quantities in $M_s = 10^8 M_\odot$ galaxies. The time to form these first $10^3 M_\odot$ is rather short when considering a gas reservoir of order $M_{\text{gas}} \approx 10^8 M_\odot$.
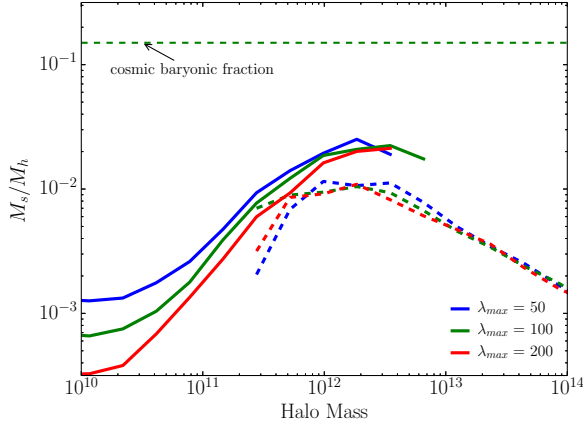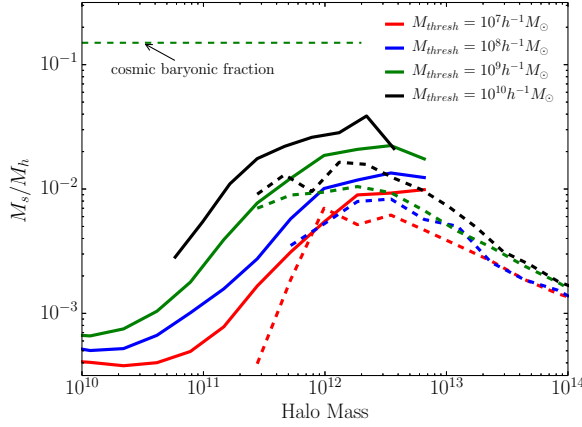
Figure 2.20: The SHMR at $z = 0$ is plotted for four different M_{thresh}. In color are the mean values of the blue population. The mean values of the red fraction are plotted in red for all four models. When changing M_{thresh} over four orders of magnitude, the SHMR changes by less than one order of magnitude.

For the initial condition of the gas content in the regulator $M_{\text{gas,init}}$ we have the freedom of $0 < M_{\text{gas,init}} < f_b M_{h,\text{init}}$. This has not more than a 1% effect on the total amount of gas that comes into a halo of mass $M_h = 10^{11} M_{\odot}$. We conclude that for merger trees which reach M_{thresh}, our freedom in the initial condition do not affect our predictions by more than 1%.

The situation for starting the forward process at $z = 15$ with a halo more massive than M_{thresh} is slightly different. This case only happens for haloes of present-day mass $M_h > 10^{14} M_{\odot}$. We cannot a priori predict the stellar or gas content of a halo of $M_h = 10^{12} M_{\odot}$ at $z = 15$. However, we study the output of the model only at $z < 8$, by which point these haloes have grown in mass by an order of magnitude. Whatever initial conditions we put in, it affects predictions at $z = 8$ by only about 10%, and even less at later epochs.

3 Gravitational lens modeling with basis sets

This chapter appeared in a similar form in Birrer, Amara & Refregier 2015 (29).

The standard cosmological model is based on the standard model of particle physics, Einsteins theory of General Relativity, a cosmological constant, cold dark matter and inflation. The physical origin of the cosmological constant, inflation and dark matter remains a mystery to date. The predictions of the expansion history of the universe has been probed with high precision and structure formation has been tested from the horizon scale down to about 1Mpc or even below (e.g., 1; 149). The smallest scale tests come from the Lyman-alpha forest (see e.g., 9) and strong and weak lensing in anomalous quadrupole lenses (e.g., 150). At even smaller scales in the non-linear regime, there are observational and theoretical challenges in bringing model and data in agreement. This problem occurs predominantly in the number, phase space densities and density profiles when comparing simulations of dark matter substructure with observations of luminous satellite galaxies in our Milky Way (see e.g., 10; 11; 12; 14). A potential non-gravitational (i.e. collisional) effect of a dark matter particle may have an effect on structure formation on small scales without having an effect on larger scales. Probing the small scale structure formation and mass distribution may thus provide information beyond the ΛCDM model.

Strong lensing is a powerful probe to test structure formation on small scales (151; 152; 153; 154; 155). Strong lensing is the effect caused by the bending of light by massive foreground

over-density (e.g. galaxy, group or cluster) such that multiple images of the same background object appears. This effect is well suited for many astrophysical and cosmological applications (see e.g. 156, a review focused on galaxy sized lenses). Strong Lensing was also proposed to detect luminous and dark substructure in the lens (157; 158). This technique has been successfully applied to data (159; 160) where sub-clumps down to about $2 \times 10^8 M_\odot$ masses have been detected. Substructure also has an effect on the flux ratios in multiple lensed quasar images (see e.g., 161; 162; 163; 164; 165). Anomalous flux ratios have thus been reported in the literature. (166) pointed out that the anomalous flux ratios measured can be accounted by line-of-sight structure and do not have to necessarily come from structure within the lens. With recent and upcoming large scale surveys new area and depth becomes available to discover strong lens systems. (167) forecasted thousands of lensed quasar systems from DES and LSST. These datasets will help to constrain the statistical features of the small scale structure imprinted in the strong lensing signal. The increasing number of strong lens systems will in the future need to be analyzed with automated modeling approaches.

The aim of this chapter is to describe a lens modeling approach that can be applied to different lens systems without adjusting parameter priors by hand and uses all the information contained in a image to constrain the projected mass density of the lens with a special emphasis on substructure. Our model approach is based on parameterized basis sets in the source surface brightness and lens model. The model framework can handle an adaptive complexity in the source and lens models. In addition to the basis sets, we show the power of modern sampling techniques and we make use of fast computational methods.

The chapter is structured as follow: In Section 3.1 we give an overview of existing lens model techniques and show how they relate to our modeling approach. In Section 3.2 the source surface brightness and lens potential basis sets on which our model relies on are introduced. Section 3.3 describes the model fitting procedure and in particular how the source surface brightness reconstruction is done and how we deal with the high number of non-linear parameters in the lens model. We test our fitting procedure on mock data and on Hubble data of the lens system RXJ1131-1231. In Section 3.5, we study how well we can detect substructure in a lens model without prior information on the mass, slope and position. This section is followed by a summary of this chapter (Section 3.6).

78

3.1   Overview of Lens Model techniques

Galaxy-size strong lenses have been modeled extensively in the literature (see references below in this section). The following aspects have to be modeled in a strong lens system when comparing data and model on the image level:

- the lens mass model

- the source surface brightness profile

- the lens surface brightness profile

- the point spread function (PSF)

Depending on the lens system and instrument, one has to also model dust extinction, external convergence, micro lensing by stars and other aspects.

Depending on the scientific aim, the main focus is typically more on the source surface brightness reconstruction or on the lens mass reconstruction. In both cases one can, broadly speaking, divide the modeling techniques in two regimes:

(1) Parametric reconstruction: Using simple and physically motivated functional forms with a controllable number of parameters ($\sim$ 10) (e.g., 168; 169; 170; 171). A controllable number of parameters implies that one can fully explore the parameter space and convergence to the best fitting configuration can often be obtained.

(2) Pixel based reconstruction: This is most often done with a grid where each pixel is treated as a free parameter. Pixelised source surface brightness inversions have been proposed by e.g. (172; 173; 174; 175; 157; 176; 177; 178; 179). These methods often rely on a regularization of the pixel grid when there is not a unique solution. Depending on the regularization procedures, priors and the pixel size, one can come to different reconstructed sources (see e.g. 177; 180). Recently (181) did an analysis of statistical and systematic uncertainties in pixel-based source reconstructions. Furthermore, these methods are computationally expensive as they rely on large matrix inversions. For the lens mass or its potential, grid based modeling has been

applied by e.g. (182; 183; 184; 157; 185; 186; 187; 158; 188; 160; 189) and even mesh-free models (190).

Computational techniques also vary for different modeling approaches. Ray-tracing has generally been used to map extended surface brightness from the source to the image plane. If significant surface brightness variations occur on very small scales, such as for quasars due to their compact size, simple ray-tracing can lead to numerical inaccuracies. One way to model such systems is to approximate quasars as point sources. One then solves the lens equation numerically for the positions in the image plane (recently e.g. 180). An alternative to avoid the point source approximation is adaptive mesh refinement (e.g., 164; 191) which changes the ray-tracing refinement scale depending on the local spacial variation of the source at different image positions.

In standard ΛCDM, the self-similarity of dark matter indicates that the same amount of complexity as seen in galaxy clusters must also be present in galaxy-sized strong lens systems. Its effect is much weaker in terms of deflection and magnification, but it must still be present. Ideally, we want a model that is flexible such that it can describe any lens mass and source surface brightness distribution. For this model we need to be able to explore its degeneracies and to converge to the 'true' solution to extract the information contained in a strong lens system.

One of the aims of our work is to fill the gap between the parametric and non-parametric models. We do so by choosing basis sets that we treat in a fully parametrized form.

3.2   Choice of basis sets

In the following sections we describe our choices for basis sets and, in addition, we present how we produce mock data given a set of parameter values.

3.2.1   Basis for the source

We make use of shapelets (introduced by 192; 193; 194) in the source surface brightness plane. We implemented the two-dimensional Cartesian shapelets (Eq. 1 and 18 in 192, or in

Appendix 3.A.1 of this work). Independent of this work (195) proposed a different method to use shapelets in the source reconstruction. Shapelets form a complete orthonormal basis for an infinite series. Restricting the shapelet basis to order n provides us with a finite basis set that is linked to the scales being modeled. If we wish to model a larger range of spatial scales in the surface brightness profile, we need to use more high order shapelets. The number of basis functions m is related to the restricted order n by $m = (n+1)(n+2)/2$. The shapelet basis functions allow us to dynamically adapt to a given problem. We can increase the complexity when we need them and reduce it when it is not appropriate. Apart from the order n one can also set the reference scale β of the basis function. Minimal and maximal scales ($l_{\min}$, $l_{\max}$) being resolved up to order n with reference scale β is $l_{\min} = \beta/\sqrt{n+1}$ and $l_{\max} = \beta\sqrt{n+1}$. The parameter β is a user specified choice. Another choice is the peak position of the shapelet center (x_0, y_0). For any finite order in n, the choice of the center is crucial for the fitting result. A natural choice for (x_0, y_0) is the center of the light profile of the source galaxy. In that sense (x_0, y_0) must be interpreted as two non-linear parameters.

3.2.2 Basis for the lens

Choosing a realistic basis set for the lens mass distribution is a challenging task as there are many different scales involved, especially when considering low mass sub-clumps. These sub-clumps are very small in scale but are also very dense. Having a basis set which allows a general description of such clumpy halos on different scales typically involves a large number of parameters. Depending on the sub-clump mass limit being considered, there are hundreds or even thousands of sub-clumps expected. A minimal description requires at least information about individual positions, masses and concentrations. Such a description leads to a degenerate and non-unique lens model (e.g. 196; 197). For cluster lenses, the typical masses of substructure are several orders of magnitude below the total lens mass, but it is possible to give strong priors on the location of the substructure, namely at the position of the luminous galaxies. For detecting invisible substructure such a prior can not be used. As often called 'non-parametric' or 'free-form' approach, meaning there are more parameters than data constraints (i.e. deliberately under-constrained) was proposed and implemented by (198) and (189). Using the catalog level image position information and time-delay measurements,

there is far less information available than parameters to be constrained. One is able to draw random realizations of lens models that meet all the constraints. Statements about the validity of a specific lens model can only be drawn statistically. Doing a comparison on the image level where about 10^3 - 10^4 pixels are involved, more information is available to constrain the model.

In our approach we start with a softened power-law elliptical potential (SPEP) (e.g., discussed by 199). The lensing potential Φ is parameterized as

$$\Phi(x_1, x_2) = \frac{2E_p^2}{\eta^2} \left(\frac{\rho_p^2 + s_p^2}{E_p^2} \right)^{\eta/2} \tag{3.1}$$

where

$$\rho_p^2 = x_1^2 + x_2^2 / \cos^2 \beta_p \tag{3.2}$$

with $\cos \beta_p$ being the axis ratio of the potential, η the power-law index, E_p the normalization of the potential, s_p the smoothing length and $x_{1,2}$ the position rotated such that x_1 is in the direction of the major axis of the potential. For an additional sub-clump, we model them either as a spherical NFW (200) profile or a spherical power-law potential (SPP). For both functions, we set the softening length $s_p = \text{const} = 0.0001"$ for computational reasons. In that sense the softening is virtually zero and is not a free parameter in this work.

Combining the two functions (SPEP and SPP) we get $6 + 4 = 10$ non-linear free and partially degenerated parameters to be fitted. With this parameterization we expect a good overall fit to many different lens systems and perhaps to catch the largest substructure within the lens, visible or invisible. Such tests are shown in section 3.4.

In addition, we include two dimensional Cartesian shapelets (same functional form as for the source in Section 3.2.1) in the potential. We choose the scale factor β to be the Einstein radius. This allows for perturbations at the global scale of the lens that can not be made with another peaked profile. The first derivatives of the potential (deflection angle) and second order derivatives (convergence and shear) can be computed analytically and can be expressed within the same shapelet basis functions (See Appendix 3.A.1), thus enabling fast

computations.

3.2.3 Basis for the lens light

For the description of the lens light, we use Sérsic profiles (201). Depending on the lens galaxy, adding multiple Sérsic profiles can lead to better fits (see e.g. 180).

3.2.4 Image making

Having a parametric description of the source surface brightness, a possible point source, the lens potential, the lens light and the PSF, an image can be generated in the following steps:

1. Starting in the image plane one evaluates the analytic expression of the deflection angle using grid based ray-tracing. The resolution has to be of order (or slightly smaller than) $\beta/\sqrt{n}$ to capture the features in the extended source model.

2. We then compute the point source image in a iterative ray-shooting procedure starting from the local minimas of the relative distance to the point source of step 1. Corrections for the next proposed ray-shooting position can be made when considering the relative displacement to the point source and the second order derivatives of the lens potential. The requirement of the precision of the point source position in the image plane of about 1/1000 of the pixel size can be reached within very few iterations.

3. For the point sources, which appear as PSF's, we normalize the externally estimated PSF to their intrinsic brightness and lens magnification. We do not lose significant computational speed when modeling the PSF further out to the diffraction spikes. For the extended surface brightness a numerical convolution needs to be made. This can be done either at the pixel or sub-pixel level. This step is the most expensive computational process in the forward image modeling. The process scales roughly linearly with the number of pixels or sub-pixels in the convolution kernel. We use Fast-Fourier-Transforms implemented in a `scipy` routine in `python`. Our default kernel size is 15×15 pixels.

4. The lens light is added with analytical Sérsic profiles convolved with the same PSF kernel

as the extended source surface brightness.

For the modeling, we do not add noise. When simulating realistic images, we add a Gaussian background noise with mean zero to all pixels and a scaled Poisson noise on the signal (pixel by pixel).

3.3   Model fitting

For the modeling, we have three questions to answer:

1. What is the best fit configuration of the model to match the data of a specific lens system? We want to find the global minima for the χ^2 value.

2. What level of complexity is needed to fit the data to a certain level? We want to compare consistency with the data by analysing the reduced χ^2 value and compare different model configurations with a Bayes factor analysis.

3. How well is the model solution determined by the data? We want to sample the parameter space and determine confidence intervals.

As a result, many choices have to be made in the lens modeling. More than 10 parameters in the lens model with non-linear behavior have to be specified. For a realistic surface brightness description the shapelet order n can be higher than 20 which corresponds to 154 basis' and their corresponding coefficients. Given this level of complexity, even the first question on its own is difficult to address. Once we have a method for addressing the first question, repeating the procedure with different choices of complexity and parameterization will provide an answer to question 2. Question 3 can then be answered with a Bayesian inference method such as a Markov chain Monte Carlo (MCMC) sampling. For this we use the software package `CosmoHammer` (202), which is based on the *emcee* method of (203) and its implementation by (204). The software package allows for massive parallelization in the sampling process. In this section, we focus on question 1. We will describe in detail the methods and procedures we apply to make the algorithm converge to the best fit lens model configuration. Question 2 and 3 are addressed with examples in Section 3.4 and 3.5.

3.3.1 Source surface brightness reconstruction

In our method we use a weighted linear least square approach to reconstruct the source surface brightness. This is a standard procedure to minimize the quadratic distance between data and model with weighted error measures. The estimation of the covariance can also be calculated (see Eqn 3.3 - 3.6 below). The minimization problem has to be linear. Let $\vec{y}$ be the data vector of dimension d. In our system, it contains all the pixel values of the image in the area of interest for a surface brightness reconstruction. Let W be the weight matrix of dimension $d \times d$. In a likelihood interpretation, W is the inverse covariance matrix of the data. Assuming the pixel errors are uncorrelated W is a diagonal matrix. Let $\vec{\xi}$ be the parameter vector of dimension m. In our case, $\vec{\xi}$ is the vector of the coefficients of the linear combination of shapelet basis functions. The number of shapelet basis functions m depends on the shapelet order n as described in section 3.2.1. Let X be the linear response matrix of the shapelet parameters on the pixel values in the image plane of dimension $d \times m$. The product $X\vec{\xi}$ describes a lensed and convolved surface brightness on the image plane. X can be computed by mapping all m shapelet basis functions from the source to the image plane, convolve and resize them separately on the pixel scale. The computational cost of this procedure is linear in the number of basis functions involved and dominates the process for low m.

Figure 3.1 illustrates how the shapelet basis functions are mapped. The problem of finding the best source configuration $\vec{\xi}_0$ given the data $\vec{y}$ and the weights W can be posed as a weighted linear least square problem:

$$\vec{\xi}_0 = \min_{\xi} \| W^{1/2} (\vec{y} - X\vec{\xi}) \| \tag{3.3}$$

This equation can be written as

$$(X^{\top} W X)\vec{\xi}_0 = X^{\top} W \vec{y} \tag{3.4}$$

whose solution is given by

$$\vec{\xi}_0 = (X^{\top} W X)^{-1} X^{\top} W \vec{y}. \tag{3.5}$$
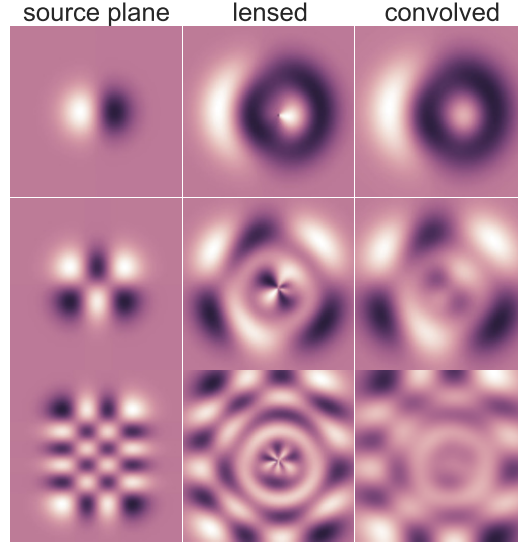
Figure 3.1: An illustration of the modeling of the source surface brightness with three different shapelet basis functions. Left panels: Shapelet basis function in the source plane. Middle panels: Mapped shapelets in the image plane with a SIS lens via ray-tracing. Right panels: PSF convolved image. From top to bottom: Shapelets with $(n_1, n_2) = (1,0), (2,1), (3,5)$.

The covariance matrix of $\vec{\xi}$, M^{ξ} is therefore given by

$$M^{\xi} = (X^{\top} W X)^{-1}. \tag{3.6}$$

M^{ξ} becomes important when marginalizing the probability distribution over $\vec{\xi}$.

The procedure involves a matrix inversion of dimension $m \times m$. The computational cost and memory allocation of this inversion becomes more significant with larger m. Moreover, the matrix $(X^{\top} W X)$ has to be invertible. If not, this method fails to find a solution and regularization is needed. A grid based regularization was introduced by (173). Conceptually and computationally, the method of (173) and the one presented in this chapter differ significantly. The matrix $(X^{\top} W X)$ is a dense matrix where as the matrix in grid based regularization can be sparse. A sparse matrix can only be maintained when having a small PSF(e.g. 5×5 pixel). We use in our method a default PSF kernel of 15×15 pixels and a further extension affects only the FFT-convolution of the lensed shapelet basis functions. Our method is well suited to

reconstruct also lensed sources in images with larger PSF's than HST images. But the main gain of our method is in terms of the number of parameters (i.e. the size of the matrix). Well chosen basis sets can allow for a smaller number of parameters compared to grid based methods significantly.

In Figure 3.2, we take a mock image produced with a chosen $\vec{\xi}_0$ (incl. point sources) with maximum shapelet order $n_{\mathrm{max}} = 40$ and added Poisson and Gaussian background noise to the image. We check the reconstruction by computing the relative residuals and their correlation function. We see from Figure 3.2 that the results are almost consistent with purely uncorrelated noise. Only for very small separations, the correlation is marginally smaller than with noise. This effect highly depends on the signal-to-noise ratio of the shapelets. Since we know the input source surface brightness, we can also check its reconstruction. The error in input vs. output in the source has features which represent the scales of the shapelet basis functions. The relative error of the surface brightness is about 10% or less. This reconstruction process with $n_{\mathrm{max}} = 40$ and 15×15 pixel convolution kernel takes about 4s on a standard personal computer. When reducing the number of shapelet coefficients to $n_{\mathrm{max}} = 20$ the reconstruction falls below 1s.

The specific reconstruction depends on the noise realization. By repeating the reconstruction 1000 times with different noise realizations, we find that the reconstruction is stable. In Figure 3.3 we plot the reduced χ^2 distribution of the different realizations. We find a mean $\chi^2_{\mathrm{red}} = 1.0015$ with a standard deviation of $\sigma = 0.009$.

3.3.2 Convergence techniques

In the previous section, we showed that we can linearize all parameters of the source model given a specific lens model and thus we can express it as a linear minimization problem. The marginalization of the linear parameters can be made analytically (see Section 3.3.2 below). Changes in the lens model however have a non-linear effect on the image. In that sense, we can marginalize over many parameters in our model and are left with about 10-30 non-linear parameters. To explore this space we use a Particle Swarm Optimization (PSO) (205) algorithm to find the minimum of the parameter space. The algorithm is described in more detail with
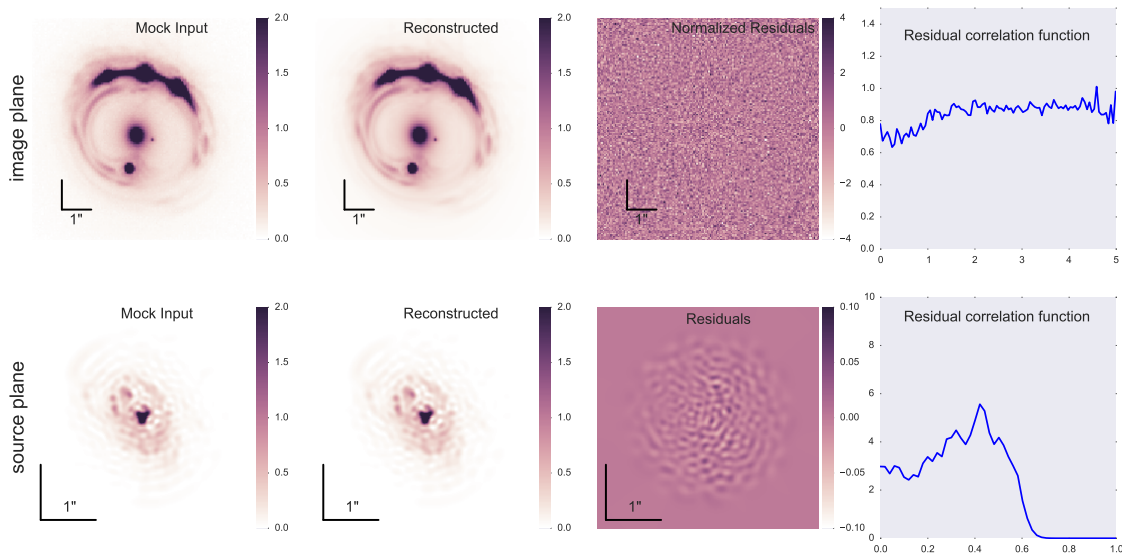
Figure 3.2: Demonstration of the source surface brightness reconstruction with upper panels showing the image plane and lower ones for the source plane. From left to right: Initial mock image (source), reconstructed image (source), relative residuals, 1D correlation function of residuals. The image is almost perfectly reproduced even without significant residual correlations. The features of the source surface brightness profile is very well reproduced. The relative intensities of input vs. output is 10% or below. The spacial correlation of the relative difference is enhanced. This feature reflects the properties of the shapelet basis functions involved and the minimal and maximal scales of those.
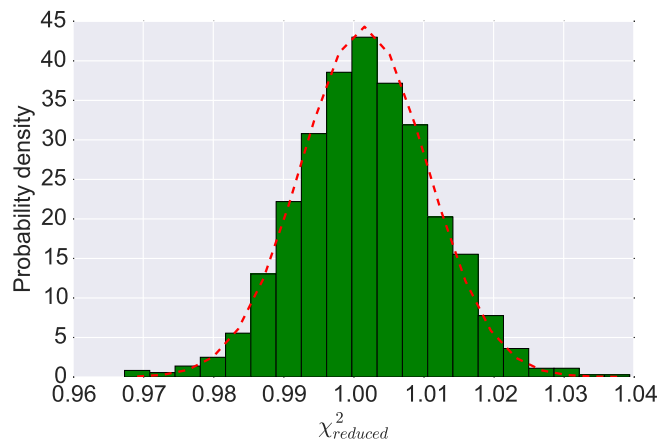
Figure 3.3: The distribution of the reduced χ^2 values of 1000 realizations of the image reconstruction process (see Figure 3.2) is plotted. Each realization differ in the Gaussian and Poisson noise realization. The mean value of the distribution is $\chi^2_{\mathrm{red}} = 1.0015$ and the spread is $\sigma = 0.009$.

an illustration in Appendix 3.A.3.

Convergence towards the global minimum in parameter space can depend on several factors. It depends on the volume of the parameter space, the number of local minima and the shape of the cost function around the absolute minimum. As we are marginalizing over all the source surface brightness parameters, one can have unexpected behavior of the cost function over the lens parameters. In the following we describe our convergence method which goes beyond simply applying the PSO algorithm and which is important for the performance of our method.

Parameterization

The sampling in parameter space can be made in any parameterization with a bijective transformation to the originally described form. The parameterization can have a significant impact on the convergence capacity and performance of a specific algorithm. If there are periodic boundaries in a specific parameterization, some algorithms can have difficulties. In our model, this is the case for the parameter of the semi-major axis angle of the elliptical lens potential θ_0 which is defined in the range $[0, \pi)$. The model can continuously rotate the axis but the parameter space has to jump from 0 to π, or vise versa. Mapping θ_0 and the axis ratio $q = \cos\beta$ into ellipticity parameters (e_1, e_2) with $f : [0, \pi) \times (0, 1] \rightarrow (-1, 1) \times (-1, 1)$ given as

$$f(\theta, q) = \left(\frac{1-q}{1+q} \cos(2\theta), \frac{1-q}{1+q} \sin(2\theta) \right) \tag{3.7}$$

provides a continuous link between the lensing potential and the parameter space. Reducing the surface area of boundary conditions in the parameter space can also reduce the number of local minima at the boundary surface. The fewer local minima there are the better one can find the global minimum. Priors on (θ, q), i.e. based on the observed light distribution, must be transformed into priors on (e_1, e_2) accordingly. In this work we assign uniform uninformative priors on (e_1, e_2).

In general, the particular choice of the parameterization can be crucial. The smoother a change in parameter space reflects a small change in the model output, the better a convergence algorithm can deal with the system. The fewer constraints and boundary surfaces there are
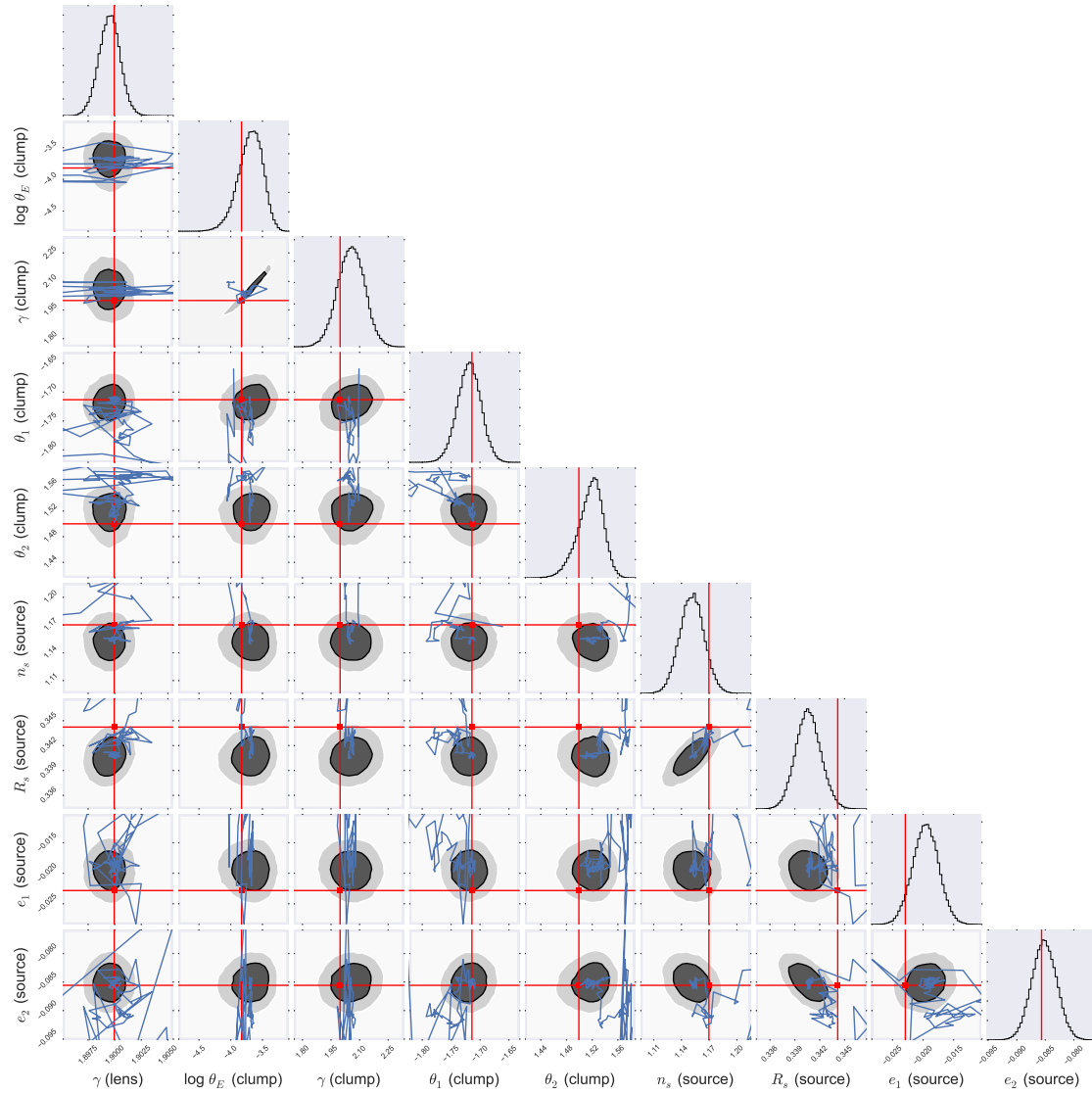
Figure 3.4: Illustration of a combined PSO and MCMC chain in a 9 dimensional non-linear parameter space. The blue lines connect the best fit particle during the PSO process. The red lines mark the true input parameter. Dark (light) gray contours mark the 68%-CL (95%-CL) interval estimated from the MCMC process.

in the parameter space, the more general convergence algorithm manage to converge to the global minimum.

Convergence with additional constraints

In cases where the source galaxy hosts a quasar that dominates the luminosity, its lensed positions in the image plane can be determined by the data without knowledge of the lens, source position or the extended surface brightness. The feature in the image is very well predicted by the PSF model and dominates the brightness over an extended area in the image. Any proposed lens model that predicts the image positions displaced from the features in the image will be excluded by the data with high significance. The quasar point sources introduce a degeneracy of acceptable solutions within the original parameter space. Knowing about this degeneracy can lead to faster convergence.

When having N bright point source images, there are $2N$ constraints to the system (their positions in the image plane). This reduces the effective dimensionality of the parameter space by $2N$. Lensing has three symmetries imprinted in the positional information: Two translations and one rotation. These transformations do not change the lens model apart from its own transformation.

In general, we can use any parameterization θ_i of an originally M-dimensional parameter space to dimension $n = M - 2N + 3$ (with $N >= 2$) if there exists a bijective transformation (an exact one-to-one mapping of the two sets) to the original parameter space with the applied constraints. In the case of four bright images of a quasar, we determine an $(M - 5)$-dimensional parameter space and solve for the source plane position of the quasar and five additional lens model parameter with a non-linear solver. This reduces the non-linear parameter space in the PSO process and leads to faster convergence without breaking any degeneracies. The choice of the five lens model parameters is arbitrary as long as the parameters can provide a solution of the point source mapping. Priors on these parameters have also to be applied in the sampling process.

Likelihood computation

The likelihood calculation on the image level is the product of the likelihoods of each pixel (see e.g. 180, for a similar approach). We estimate the variance on the intensity at pixel i as

$$\sigma_{\text{pixel,i}}^2 = \sigma_{\text{bkgd,i}}^2 + f d_{\text{model,i}} \tag{3.8}$$

where $\sigma_{\text{bkgd,i}}$ is the background noise estimated form the image, $d_{\text{model,i}}$ the model prediction at pixel i and f a scaling factor. A pure Poisson noise results in f being the product of exposure time and gain. The likelihood of the data d_{data} with N_{d} image pixels given a model d_{model} with non-linear lens model parameters $\boldsymbol{\theta}$ can then be written as a marginalization over the linear parameters $\boldsymbol{\xi}$, the source surface brightness parameters:

$$P(d_{\text{data}}|\boldsymbol{\theta}) = \int d\boldsymbol{\xi} \, P(d_{\text{data}}|\boldsymbol{\theta},\boldsymbol{\xi}) P(\boldsymbol{\xi}) \tag{3.9}$$

where

$$P(d_{\text{data}}|\boldsymbol{\theta},\boldsymbol{\xi}) = \frac{1}{Z_{\text{d}}} \exp \sum_{i=1}^{N_{\text{d}}} \left[-\frac{(d_{\text{data,i}} - d_{\text{model,i}})^2}{2\sigma_{\text{pixel,i}}^2} \right]. \tag{3.10}$$

with $\boldsymbol{d}_{\text{model}} = \boldsymbol{X}(\boldsymbol{\theta})\boldsymbol{\xi}$. Z_{d} is the normalization

$$Z_{\text{d}} = (2\pi)^{N_{\text{d}}/2} \prod_i^{N_{\text{d}}} \sigma_{\text{pixel,i}} \tag{3.11}$$

and $P(\boldsymbol{\xi})$ the prior distribution of the shapelet coefficients. We assume a uniform prior distribution which is independent of the lens model. The integral in equation (3.9) can be computed around the maximum $\boldsymbol{\xi_0}$ coming from equation (3.5) with covariance matrix M^ξ from equation (3.6). With a second order Taylor expansion around $\boldsymbol{\xi_0}$, equation (3.9) can be written as

$$P(d_{\text{data}}|\boldsymbol{\theta},\boldsymbol{\xi_0} + \Delta\boldsymbol{\xi}) \approx P(d_{\text{data}}|\boldsymbol{\theta},\boldsymbol{\xi_0}) \cdot e^{-\frac{1}{2}\Delta\boldsymbol{\xi}^T (M^\xi)^{-1}\Delta\boldsymbol{\xi}}. \tag{3.12}$$

Integrating equation (3.12) over $\Delta\boldsymbol{\xi}$ results in

$$P(d_{\text{data}}|\boldsymbol{\theta}) = P(d_{\text{data}}|\boldsymbol{\theta},\boldsymbol{\xi_0}) \left[(2\pi)^m \det(M^\beta) \right]^{\frac{1}{2}} \tag{3.13}$$

In principle, equation (3.10) is the cost function to use for image comparison. The information about the image positions is included in this cost function. The problem with this cost function is that convergence to a good model can be difficult. The use of additional or derived information, such as the explicit image positions, can facilitate convergence.

Steps towards convergence

Having presented our model parameterization in Section 3.2 and discussed certain aspects of model fitting and convergence in the previous paragraphs, we describe our steps which allows us to find a reasonable fit to the data. Figure 3.5 illustrates the framework. Prior to the convergence algorithm, the image data has to be analyzed, the model configuration has to be chosen, the prior values have to be set and the specific configuration of the PSO process have to be given as an input. The fitting should be done within an automated process where no interaction of the modeler is needed. The output of the PSO run can then be analyzed by the modeler in terms of convergence and quality of fitting. This may lead to a change in the model parameters, functions, configuration etc and the process is run again. Once convergence is achieved and the fitting result is good, the MCMC process is run to map out the valid parameter space given the model parameters chosen. Figure 3.4 illustrates the PSO and MCMC process in a 9 dimensional parameter space. The thin blue lines corresponds to the PSO process. Once this process is converged we start the MCMC process around this position (light and dark gray contours). Certain parameters are more degenerate than others. We try to map the parameter space such that remaining degeneracies are controllable.

3.4 Example - RXJ1131-1231

In the following, we test our method on the gravitational lens RXJ1131-1231. This lens was discovered by (206) and the redshift of the lens $z_l = 0.295$ and of the background quasar source
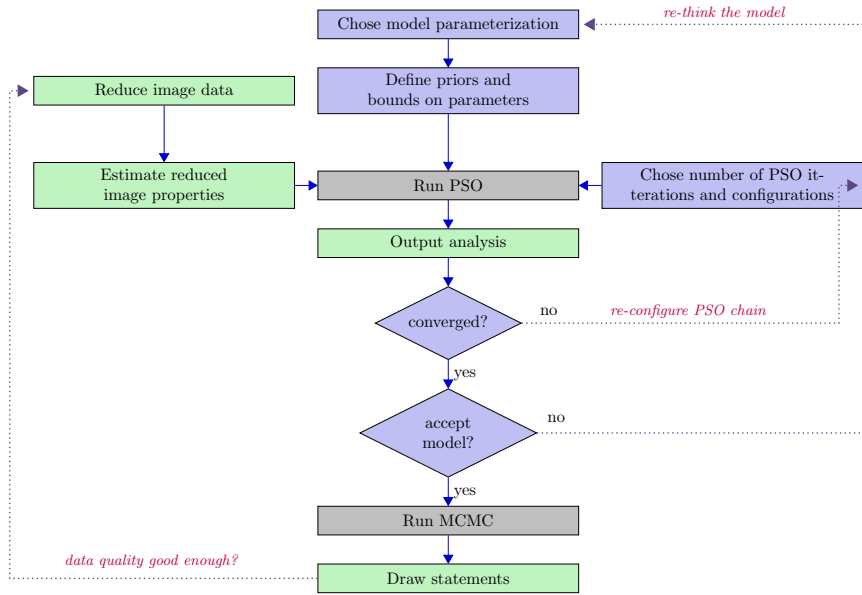
Figure 3.5: Chart of the framework highlighting user interactions. Human interactions are needed for some tasks (green) and decisions (blue). Automated tasks are shown in gray. The core of this framework is to clearly split the preprocessing from the fitting algorithm.

$z_s = 0.658$ was determined spectroscopically. The lens was extensively modeled by (207; 208; 180). We use the same archival HST ACS image in filter F814W (GO 9744; PI: Kochanek) as (180) for our lens modeling and follow a similar procedure for the reduction process and error estimation. We make use of the `MultiDrizzle` product from the HST archive. The PSF is estimated from stacking of nearby stars. We estimate a PSF model error by computing the variance in each pixel from the different stars after a sub-pixel alignment with an interpolation done using all the stars. We assume that this model error is proportional to the intensity of the point source. This method is meant to demonstrate our method in fitting the best configuration. The lens model is parameterized as a SPEP (ellipsoid) and a second SPP (round) profile (see Eqn 3.1). Furthermore we choose 15 shapelet basis sets in the potential and a constant external shear component. For the lens light we follow (180) and use two elliptical Sérsic profiles with common centroids and position angles to describe the main lens galaxy and a circular Sérsic profile with $n_{\mathrm{sersic}} = 1$ for the small companion galaxy.

Figure 3.6 shows our result of the fitting process to the HST image. In the upper left panel we show the reduced data. Upper middle shows the best fit model. On the upper right the

normalized residuals are plotted. The reduced χ^2 value of this fit is $\chi^2_{\text{red}} = 1.5$ without adjusting any Poisson factors nor the background noise level originally derived from the image data products. We clearly see that there are significant residuals around the point sources which indicates clearly that our PSF model needs further improvement and that even our error model on the PSF seems to underestimate the model error in certain regions. Furthermore, extended regions of over- or under-fitting indicate that the lens model can be improved. Source surface brightness adoptions could have acted to reduce the error in the fit in case of a perfect lens model. The lower left panel shows the reconstructed extended source surface brightness profile. We clearly see the presumably star forming clumps which lead to the features in the extended Einstein ring. In the lower middle panel of Figure 3.6 our lens model is shown in terms of the convergence map. We notice that without mass-to-light priors, the position of the two modeled clumps is strikingly close to the position of the luminous galaxy and its companion. In the lower right panel, the magnification map is shown. The reconstruction of the image depends on the number of shapelet coefficients used. In Appendix 3.A.2 we discuss the effect of n_{max} on the quality of the source and image reconstruction for this particular lens system.

Comparing different lens and source model reconstructions from different methods is difficult. Different source surface brightness reconstruction techniques use different number of parameters and thus can have different χ^2 values without changing the lens model. Error models and masking do have a significant impact on χ^2. Setting priors may also lead to different results (In case of (180) the position of the sub-clump modeled as a singular isothermal sphere was fixed at the position of the luminous companion and additional information from velocity dispersion measurements). All in all, different lens modeling techniques can only be properly compared based on mock data. And even on mock data, different input types of lens and source models might have a significant influence on the relative performance of the methods.

3.5 Detectability of substructure

One of our main focuses for the model we present is to find and quantify substructure within a lens. In this section, we want to discuss the following issues:
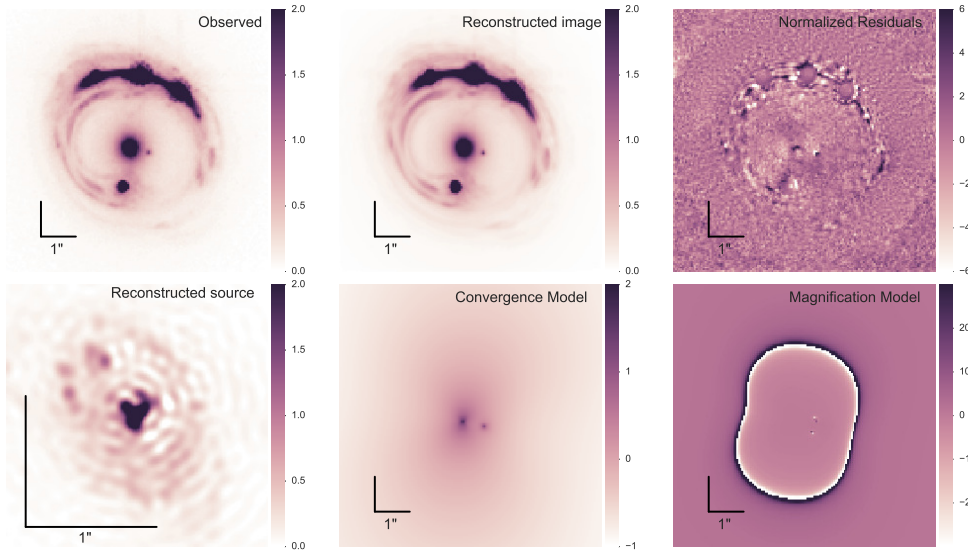
Figure 3.6: Modeling of RXJ1131-1231 HST ACS F814W image. Upper left: Observed image. Upper middle: Best fit pre-construction. Upper right: normalised residuals of the reconstruction. Lower left: Reconstructed source with 1326 shapelet coefficients (up to order 50). Lower middle: Convergence model of the lens. Lower right: Magnification model of the lens.

1. To what extend are our model basis functions and description able to reproduce the true image?

2. In case of a perfect modeling: Are we able to recover the true parameter configuration in a large parameter space?

3. In case of an imperfect modeling: How does this affect the sensitivity limit, finding and quantification of substructures?

To answer our first question, we refer to our data example of RXJ1131-1231 in Section 3.4 of Figure 3.6. Even though the observed and predicted images can hardly be distinguish by eye, the residual map indicates room for improvements in our modeling. Nevertheless the fact that our mass-to-light prior-free lens model provides us with a realistic solution might indicate that we are not far from reality. A priori, we do not know whether the solution found in Section 3.4 is the global minimum of the parameter space chosen and therefore the best reachable solution within the choices and parameters made. We will investigate whether the finder algorithm is able to recover the true input parameters when fitting mock images in the next section.

3.5.1 Substructure finding

To approach question 2 above we take a mock image that is highly inspired by RXJ1131-1231 of section 3.4. We keep the image quality fixed (i.e. noise levels, pixel size and PSF) but change the lens model such that we have one big SPEP profile and a minor sub-clump, a spherical power law potential (SPP). Ideally, we do not want to set any prior on the position, mass, shape and number of substructures. If we were interested in luminous sub-structure we could add mass-to-light priors. As we want to use our method to potentially detect dark sub-structure, we are not allowed to give any mass-to-light prior. Therefore we want to check whether our algorithm finds the preferential parameter space in the model. The main focus is on the position of the sub-clump. To explore our capability of finding sub-clumps, we generate mock data with a sub-clump in the lens model at a random position. We add Poisson and Gaussian noise on the mock image. We then run the convergence method on that image with the same weak prior information as was done for the real image in section 3.4. We repeat this procedure 10 times. Our result is:

- Success rate in position of 100%. For our setting with a random sampling of the prior parameter space, all the runs ended around the right solution (PSO).

- Detectability down to 10^{-4} level of the total lens mass in the arc of the Einstein ring (MCMC).

- Time for convergence of about 10^5 evaluations of a model configuration needed. One evaluation takes few seconds.

For one realization of the input-output process the comparison is shown in Figure 3.7 in terms of convergence and magnification and their residuals. We clearly see that the position of the sub-clump can be well recovered and the appearance of the critical line do match very well. This means that there is no other degenerate solution within the parameter space that can reproduce a similar feature like a sub-clump no matter what combination of source surface profile and lens model we chose. This test shows that with a ideal model we can find a single sub-clump in the lens mass without any prior on its existence and position. We also highlight that the relative likelihood comparison is large (more than 5 sigma compared with the best fit
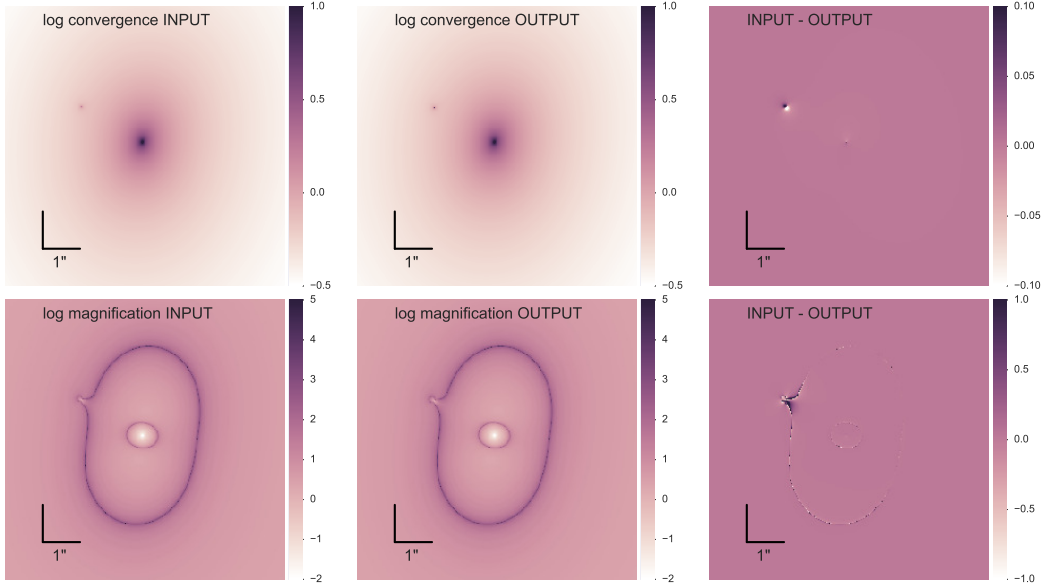
Figure 3.7: Lens model input-output comparison for convergence and magnification log(abs(magnification)). Without priors on the size or position of the sub-clump of order $10^8 M_\odot$, the PSO can find the lens configuration. The input-output comparison of the image is illustrated in Figure 3.2. The sensitivity for the sub-clump detection is analyzed in Figure 3.4. We clearly see that the we are sensitive to the position up to 0.1".

model without a sub-clump). This statement in this form can only be made if other effects (such as error in the PSF model) do not interfere. As we showed for the real image in section 3.4, errors in the PSF model alone can potentially lead to a higher increase in the minimal χ^2. This test together with the finding of a sub-clump in RXJ1131-1231, in both cases without setting priors on position, mass-to-light, concentration and mass, are encouraging hints that our model approach can extract valuable information about a lens system in a rather unbiased way.

3.5.2   Substructure sensitivity

The last section discussed the potential to recover sub-structure. We showed that this is possible for substructure with mass ratios of $M_{sub} \sim 10^{-4} M_{lens}$ consistent with a direct detection of a sub-clump in (160). This implies, that we are sensitive to this mass regime in fitting a mass-concentration relation. The concentration of the sub-clump is not exactly matched when comparing with the most likely solution of the PSO in Figure 3.7. To effectively see how

well we can constrain certain parameters in the model from the data, we do a full likelihood analysis with a MCMC. The mapping of the entire parameter space of this realization with an MCMC is illustrated in Figure 3.4 for exactly the same realization as Figure 3.7. The red lines mark the input parameters. We see, that the input parameters are always within 2 sigma of the output parameter distribution. We see that there is a partial degeneracy between mass and concentration of the sub-clump (γ (clump) and $\log \theta_E$ (clump) in Figure 3.4. Not surprising, it is difficult to constrain the profile of a very small clump which itself is close to the detection limit. Even better data than HST, such as JWST or ALMA can potentially detect clumps down to lower mass levels and also constrain the profiles of these small clumps. The sensitivity limit relies mostly on three criteria: FWHM of PSF, magnification at the position of the sub-clump and source surface brightness variation at the lensed position of the sub-clump. For any different data qualities, telescopes etc, we are able to perform such sensitivity tests.

3.6   Summary

In this chapter, we introduced a new strong lensing modeling framework which is based on versatile basis sets for the surface brightness and lens model. We identified the following aspects of our framework:

- Its modular design allows for a step-by-step increase in complexity. We are able to determine which part of the modeling needs more complexity to reproduce a lens system.

- It allows for automated or semi-automated fitting procedures. An adaptive cost function combined with a best fit algorithm, allows it to fit different strong lens systems without giving specific priors to each one of them. This allows for faster and more systematic analyses of large numbers of lens systems.

- It is suitable for a wide range of strong lensing systems and observing conditions. Our framework can be applied to various levels of image quality, different type of lens system, sizes of the lens, etc, as the convergence algorithm does not rely on strong initial priors.

- It features fast source reconstruction techniques. The evaluation of the cost function

given a position in parameter space including the simulation of the image can be achieved within seconds. Furthermore our convergence algorithm allows for massive parallelization on a distributed computer architecture.

We further proposed a way to model strong lens systems to extract information about the substructure content within the lens. Such investigations can potentially provide useful constraints on the abundances of low mass objects. To learn about the dark matter properties from strong lensing, one needs to combine well chosen descriptions for the source light and lens mass, algorithm techniques which can find solution in high dimension parameter spaces and a combination of different data sets to break degeneracies (multi-band imaging, spectroscopy, etc.). A special focus has to be made in choosing the right set of basis functions and the algorithmic design of the convergence method. Our approach is encouraging as it succeeds in recovering substructure in the lens without setting mass-to-light priors.

3.A   Appendix of chapter 3

3.A.1   Shapelets

The two dimensional Cartesian shapelets as described in (192) are the multiplication of two one-dimensional shapelets $\phi(x)$:

$$\phi_{\mathbf{n}}(\mathbf{x}) \equiv \phi_{n_1}(x_1)\phi_{n_2}(x_2). \tag{3.14}$$

The one-dimensional Cartesian shapelet is given by:

$$\phi_n(x) \equiv \left[2^n \pi^{\frac{1}{2}} n!\right]^{-\frac{1}{2}} H_n(x)e^{-\frac{x^2}{2}} \tag{3.15}$$

where n is a non-negative integer and H_n the Hermite polynomial of order n. The dimensional basis function is described as

$$\phi_n(x;\beta) \equiv \beta^{-\frac{1}{2}}\phi(\beta^{-1}x). \tag{3.16}$$

In two dimensions, we write

$$\mathbf{B_n}(\mathbf{x}; \beta) \equiv \beta^{-1}\phi_\mathbf{n}(\beta^{-1}\mathbf{x}). \tag{3.17}$$

This set of basis functions is used for the source surface brightness modeling and the lensing potential. To find the derivatives of this functions, (192) introduced raising and lowering operators which act on the basis functions as

$$\hat{a}\phi_n = \sqrt{n}\phi_{n-1}, \hat{a}^\dagger\phi_n = \sqrt{n+1}\phi_{n+1}, \tag{3.18}$$

the derivative operator can be written as

$$\frac{d}{dx} = \frac{1}{\sqrt{2}}\left(\hat{a} - \hat{a}^\dagger\right) \tag{3.19}$$

and therefore any derivative can be written as a superposition of two other shapelet basis functions (for further discussions, see (192)) In Figure 3.8, it is illustrated how shapelet basis functions in the potential space do map in the deflection angle and convergence.

3.A.2 Number of shapelet coefficients

The choice of the maximal order of the shapelet coefficients $n_{\max}$ and its corresponding number $m = (n_{\max} + 1) \cdot (n_{\max} + 2)/2$ has a significant influence in the goodness of fit to imaging data. In Figure 3.9 we illustrate this by reconstructing the source surface brightness with different $n_{\max}$. We use the same lens model as in Figure 3.6. We see that even with $n_{\max} = 10$, most of the features in the arcs of the image could be reconstructed qualitatively but significant residuals remain. By increasing $n_{\max}$, more and more details in the source appear and the residuals go down.

3.A.3 Particle Swarm Optimization

The Particle Swarm Optimization (PSO) description was introduced by (205) as a method to find the global minima in a high dimensional non-linear distribution. The algorithm is motivated by the physical picture of a swarm of particles moving in a physical potential. Every
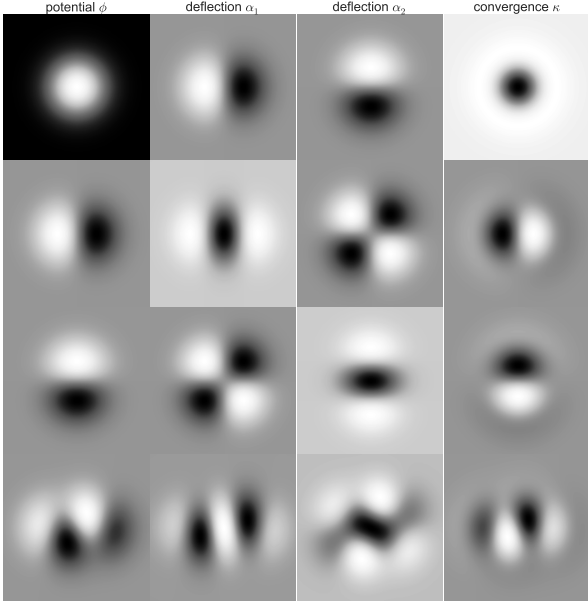
Figure 3.8: The shapelet functions in potential space are plotted in the first row. From top to bottom: (0,0), (1,0), (0,1), (0,1) + (3,0). The second and third rows show the deflection angles α_1 and α_2. The last row shows the corresponding convergence κ.
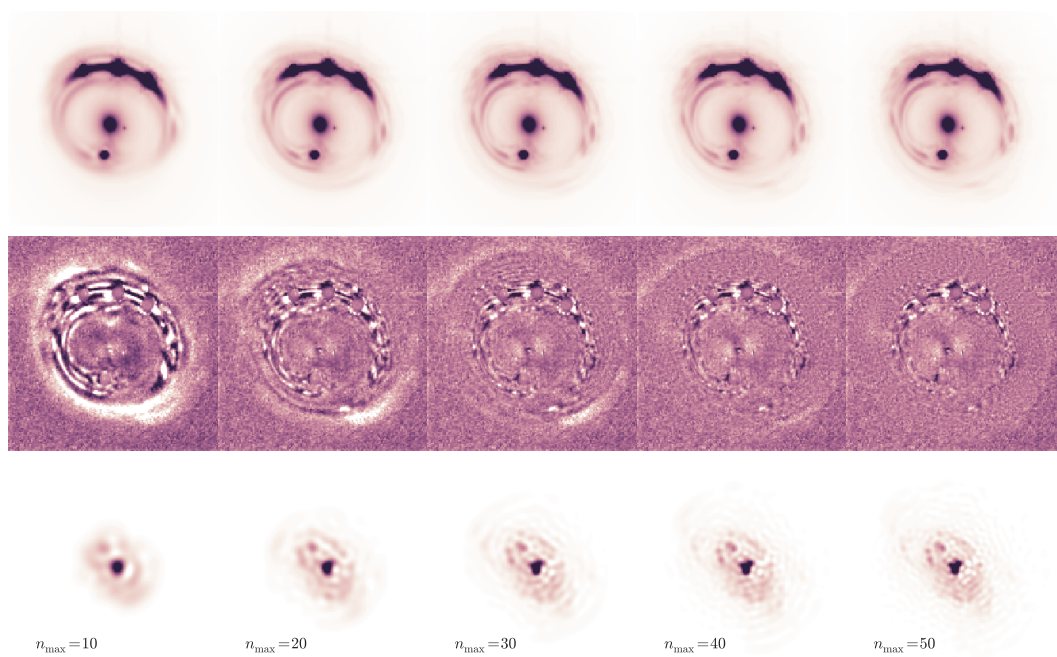
Figure 3.9: The source surface brightness reconstruction of the lens system RXJ1131-1231 is modeled with different shapelet orders n_{max}. Upper panel: The reconstructed image. Middle panel: The normalized residual maps. Lower panel: The reconstructed source. From left to right: Increasing number of shapelet order n_{max} from 10 to 50.
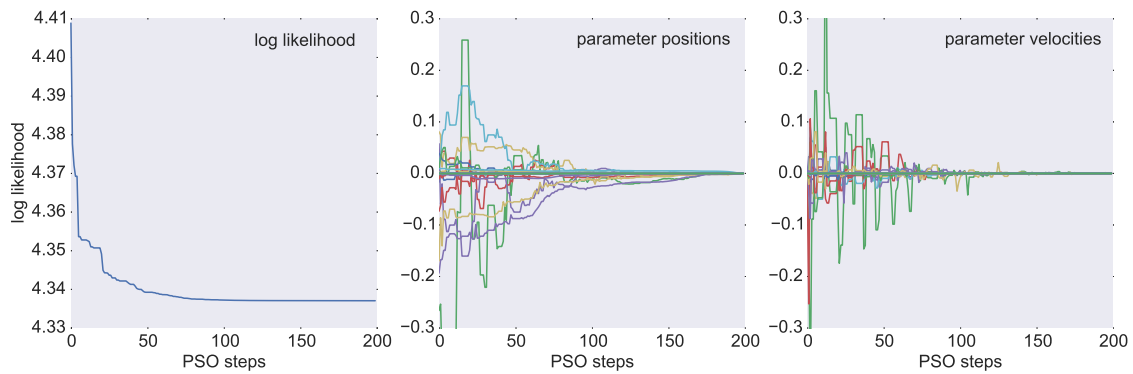
Figure 3.10: Illustration of the PSO process in 20 dimensions with 160 particles and 200 iterations. Left panel: Evolution of the log likelihood of the best fit particle. Middle panel: The difference of the parameter values from the best fit at each iteration relative to the end point of the PSO process. Right panel: Velocity of the best fit particle at each iteration. Different colors are used for each of the parameters.

particle gets assigned a position in parameter space, a function evaluation (the log likelihood value) and a velocity in parameter space. The particles is assigned a swarm "physical" behavior when moving up or downwards a potential and a "swarm" behavior when redirecting their velocity towards the particle at the deepest place of the potential. The PSO process is illustrated in Figure 3.10 in a 20 dimensional parameter space. The implementation of the PSO algorithm used in this work is publicly available as part of the `CosmoHammer` (202) software package. The inertia weight strategy comes from (209) and the stopping criteria of (210) was implemented.

4 Line-of-sight effects of strong lensing: Putting theory into practice

Gravitational lensing is a unique probe for measuring dark matter and dark energy by mapping the mass distribution of the universe on different scales. On the largest scales, weak gravitational lensing surveys measure the linear and non-linear regime of structure formation (see e.g. reviews of (211; 212) and references therein). On scales of individual galaxies, time-delay cosmography measures angular diameter distance relations ((213; 214) or (215) as a recent review), which gives us information on the background expansion of the universe. On sub-galactic scales, the abundances of dark substructure that can be studied using strong lensing is sensitive to the physical properties of dark matter (151; 152; 153; 163; 154; 155).

These different gravitational lensing regimes each developed formalisms to connect their observables with the underlying physical distribution of dark matter being studied. The distinction between these domains effectively reflected the simplifying assumptions they each used. For example, the simplest approach in modeling strongly lensed systems is to describe the process in terms of a single strong perturber and to neglect contributions of other masses along the line-of-sight (LOS). On the other hand, for weak lensing studies, the focus is on integrated tidal distortions due to structure along the LOS, but higher order non-linear stronger lensing effects are typically not included. However, with increasing volumes and

quality of data, this distinction between the different regimes is no longer sufficient and an integrated approach to lensing problems needs to be adopted.

There are several examples in strong lensing where the LOS needs to be considered carefully. One is the inference of dark matter substructure properties from Quasar flux ratios (168), where LOS structure can have a significant impact on this observable and therefore can affect the interpretation of the data (216; 150; 166; 165). Another example is strong lens cosmography (213; 214; 217; 218) since integrated LOS structure in the vicinity of galaxy scale strong lens systems can have a significant impact on relative time-delay measures. These must be taken into account to perform precision cosmographic estimates (219; 220; 221; 222; 223).

Early work studying external shear and ellipticity in gravitational lensing described the impact of LOS mass distribution as an equivalent additional mass sheet at the redshift of the main deflector with uniform surface mass density κ_{ext} (224; 225). In the literature, the LOS structural parameter κ_{ext} is typically estimated using: (i) imaging and spectroscopy of objects in the neighborhood of the lensing systems (e.g. 226; 227; 220; 228; 179; 221; 229); (ii) weak lensing (e.g. 230); and (iii) using comparison with cosmological numerical simulations (e.g. 231; 232). Recently, a general multi-plane lensing framework has been introduced (50). The authors later use this approach to test their accuracy in modelling LOS structure using mock position data (223) for quadrapole lens systems. They point on the need for properly accounting for LOS structure in precision lens modeling.

In this chapter, we present a set of simplified approximation of the multi-plane framework for accounting for LOS structures. These give reliable reconstructions for strong lens systems around the Einstein ring. The advantage of this is that accuracy is maintained while also allowing us to separate the calculations of the LOS effects from the strong lensing deflections of the main lens. This, in turn, allows us to incorporate LOS modelling into our modelling tools of chapter 2 and 3 that aim to reconstruct the full extended lens system. We apply our modeling formalism to the lens system COSMOS0038+4133 and demonstrate the power gaining insights into the LOS structure through strong lens image reconstruction.

The chapter is structured as follows: In section 4.1 we revisit the geometry of multi-plane gravitational lensing, review the approaches being taken in the literature and introduce our

notation. In section 4.2, we state our approximations, the phenomenological modeling parameterization for strong lens image reconstruction, provide the link to the physical mass distribution in the universe and present test on mock data. In section 4.3, we apply our modelling formalism to a strong lens in the COSMOS field. Independently, we perform an environmental analysis based on the galaxies in the vicinity of the lens and show the strength of the combination of strong lens and environment analysis. Finally, in section 4.4 we summarize this chapter.

Throughout this chapter, we assume a flat ΛCDM cosmological model with parameters $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $H_0 = 67 \,\mathrm{km\,s^{-1}Mpc^{-1}}$.

4.1 Multi-plane gravitational lensing

In this section, we review multi-plane gravitational lensing, the joint effect caused by multiple lens planes at different distances. We further introduce our notation and state suitable approximations to the full non-linear multi-plane ray-tracing in the regime of one main strong lens. Mathematical aspects of multi-plane strong gravitational lensing were studied in (233; 234; 235; 236; 237). Of practical use for our analysis is the multi-plane lens equation (238; 239; 240).

4.1.1 General description

In gravitational lensing, the mapping from source to image is given by the lens equation (e.g. see 241, for background material) and the introduction of this thesis 1.3

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}), \tag{4.1}$$

where $\vec{\theta}$ is the un-lensed angular position, $\vec{\beta}$ is the post lensing position and α is the deflection angle. When studying strong lens systems, the thin lens approximation is widely used. This models the light travel path as straight between lens planes with sharp deflections at the
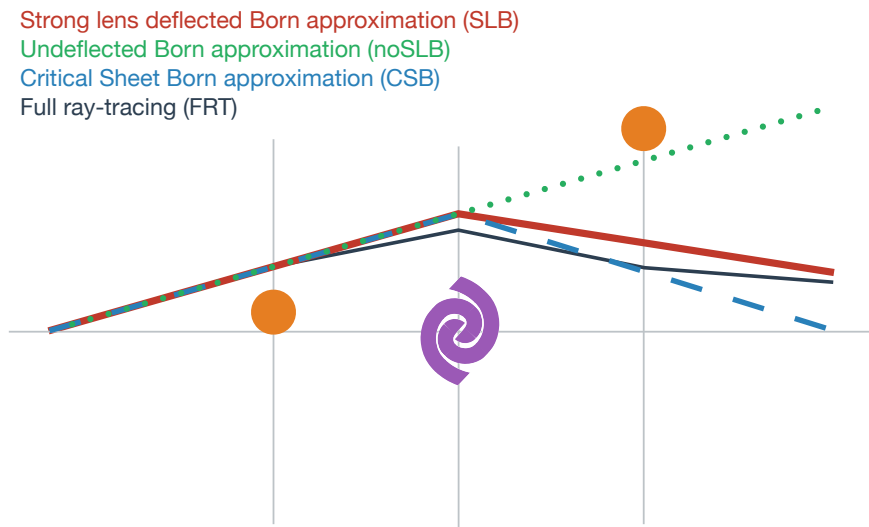
Figure 4.1: Illustration of different approximations of the light path on which to compute the LOS effects. Strong lens deflected Born approximation (SLB, red line): Computation along the strong lens deflected path, which is accurate but leads to non-linear couplings to the strong lens. Undeflected Born approximation (noSLB, green dotted line): Born approximation ignoring the strong lens. This method is inaccurate for accounting of the effects of background perturbers. Critical Sheet Born approximation (CSB, blue dashed line): Replacing the strong lens deflected path (SLB) by a critical mass-sheet deflected path (see section 4.2). Full ray-tracing (FRT, thin black line): No approximations and every (small) deflector is modeled as a single lens plane.

positions of the lenses. In the case of a single lens, the lens equation becomes,

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}) = \vec{\theta} - \frac{D_{ds}}{D_s}\vec{\hat{\alpha}}(D_d\vec{\theta}), \tag{4.2}$$

where $\vec{\hat{\alpha}}(D_d\vec{\theta})$ is the physical bending angle that is linked to $\vec{\alpha}(\vec{\theta})$ through the angular diameter distance D_s (the angular diameter distance from the observer to the source), D_d (the angular diameter distance from the observer to the lens) and D_{ds} (the angular diameter distance from the lens to the source).

In the case where there are n thin lenses along the LOS, the total mapping is the sum of all the deflections, weighted by their distance relations and evaluated at the light path as

$$\vec{\beta}_s = \vec{\theta} - \frac{1}{D_s}\sum_{k=1}^{n}D_{ks}\vec{\hat{\alpha}}_k(D_k\vec{\beta}_k) \tag{4.3}$$

where $\vec{\beta}_k$ is the angle under which the k'th lens deflects the light ray and $\vec{\beta}_1 = \vec{\theta}$. The ordering in D_k follows the backwards light path such that the light passes the lens at k before $k-1$ (see also (50) for a different notation of the same expression).

4.1.2 Case of one strong deflectors and several weak ones

Often in strong lensing, the dominant deflection is due to one single object at a specific redshift with additional deflectors causing secondary weaker effect. The black path shown in Figure 4.1 shows an illustration of the light travel path through a multiple lens system (full ray-tracing FRT). One way to simplify the calculation is to treat the smaller additional lenses as tidal perturbers (neglecting higher order terms). This tidal approximation depends on the light path. The undeflected Born approximation (noSLB, green dotted line in Figure 4.1) computes the light path ignoring any deflector. The noSLB approximation leads to a distortion that is given by the distortion matrix (Γ_{ij}):

$$\Gamma_{ij} \equiv \frac{\partial\beta_i}{\partial\theta_j} \equiv \begin{bmatrix} 1-\kappa-\gamma_1 & \gamma_2 \\ \gamma_2 & 1-\kappa+\gamma_1 \end{bmatrix} \tag{4.4}$$

where κ is the convergence and γ_1, γ_2 are the shear components of the linear distortion matrix. Another approach, we call it "strong lens deflected Born approximation" (SLB), takes the main deflector into account when evaluating the effects of the LOS perturbers. This approximation (the red path shown in Figure 4.1). The advantage of this approach is that this approximation replicates that of the full ray-tracing calculation with high fidelity relative to noSLB. Higher order effects (flexion) are small for LOS perturber (see e.g. (223)). The problem, however, is that the tidal effect can only be calculated after the light path has been found. This coupling between the ray-tracing and the impact of the secondary lenses makes such a calculation impractical when modelling strong lens systems. For this reason, there are significant advantages to finding further simplifications that are able to separate the two computations. The most commonly used method for achieving this is to perform ray-tracing for the main lens (red in Figure 4.1) with the additional effect from the secondary lenses calculated along the un-lensed path (noSLB, green dotted path of Figure 4.1).

As we will demonstrate later, ignoring the bending of the main deflector can lead to significant inaccuracies in strong lens calculations. To over come this problem we introduce a new approximation (detailed in the next sections) that better captures effects of secondary perturbers close to strong lenses without a coupling to the ray-tracing by the main lens.

4.2   Critical Sheet Born Approximation (CSB)

Since most of the information from strong lens systems typically comes from regions close to the Einstein radius, we have focused on finding an approximation that is valid in this region. At the Einstein radius the light bending is critical ($\vec{\alpha}(\vec{\theta}_E) = \vec{\theta}_E$) irrespective of the mass and detailed shape of the strong lens.

A critical mass sheet holds $\alpha = \theta$ everywhere and fulfills by construction the condition at the Einstein radius, irrespective of the Einstein radius of the strong lens. The ray-tracing through the full main deflector and evaluating the LOS effects based on light rays bent by a critical mass sheet leads to the same LOS effects around the Einstein radius compared to SLB. We call this approximation the Critical Sheet Born (CSB) approximation. In Figure 4.1, CSB is indicated with the blue dashed line. With such a description of the light path in computing

LOS effects, we can avoid non-linear coupling between main deflector and LOS perturbers.

4.2.1 Formalism

The lens equation 4.2 with additional first order distortion effects from LOS structure can be written as

$$\vec{\beta}_s = \vec{\theta} - \frac{D_{ds}}{D_s}\vec{\alpha}_d(D_d\Gamma^A_{ij}\vec{\theta}) - \left(\Gamma^B_{ij} + \Gamma^C_{ij}\right)\vec{\theta}, \tag{4.5}$$

where $\vec{\alpha}_d$ is the physical deflection angle of the main deflector, Γ^A_{ij} is the distortion matrix at the deflector plane caused by foreground perturbers, Γ^B_{ij} is the distortion caused by the same foreground perturbers at the source plane and Γ^C_{ij} is the distortion caused by background perturbers on the source plane.

For foreground perturbers, we evaluate the distortion effect on undeflected light paths $\vec{\beta}_{k<d} = \vec{\theta}$. For background perturbers, we use instead the light path caused by a critical mass-sheet at the main deflector plane. This results in a light path

$$\vec{\beta}_{k>d} = \vec{\theta}\left(1 - \frac{D_{dk}}{D_k}\frac{D_s}{D_{ds}}\right). \tag{4.6}$$

Given these approximations on the light paths $\vec{\beta}$, the distortion matrices of equation 4.5 are given by

$$\Gamma^A_{ij} = \delta_{ij} - \sum_{k<d}\frac{D_k D_{kd}}{D_d}\frac{\partial\hat{\alpha}^i_k}{\partial x_j}, \tag{4.7}$$

$$\Gamma^B_{ij} = \sum_{k<d}\frac{D_k D_{ks}}{D_s}\frac{\partial\hat{\alpha}^i_k}{\partial x_j}, \tag{4.8}$$

and

$$\Gamma^C_{ij} = \sum_{k>d}\frac{D_k D_{ks}}{D_s}\left(1 - \frac{D_{dk}}{D_k}\frac{D_s}{D_{ds}}\right)\frac{\partial\hat{\alpha}^i_k}{\partial x_j}. \tag{4.9}$$

$\hat{\alpha}^i_k$ are the physical deflections caused by the LOS perturbers and x_j are physical distances at

the perturber. The only explicit deflection in equation 4.5 is the main deflector $\hat{\alpha}_d$. As pointed out by (223), the non-linear effect of the term Γ^A_{ij} on $\hat{\alpha}_d$ is important and not taking this effect into account can lead to significant biases in the lens model inference. Furthermore, the LOS structure close to the source plane is of less importance as the light rays are bent and get closer to each other and reduced the induced tidal distortion.

4.2.2 Phenomenological parameterization

In this section, we discuss what the observables from strong lensing image reconstruction are when the underlying description is approximated by equation (4.5). The effect on the lens equation of the LOS structure can be expressed as tidal distortions (equation 4.4). Equation 4.5 becomes

$$\vec{\beta}_{\text{true}} = \vec{\alpha}_{\text{true}} \left((1 - \kappa_d) D_{\text{d}}^{\text{bkgd}} \begin{bmatrix} 1 - \gamma'_{1,d} & \gamma'_{2,d} \\ \gamma'_{2,d} & 1 + \gamma'_{1,d} \end{bmatrix} \vec{\theta} \right) + (1 - \kappa_s) \begin{bmatrix} 1 - \gamma'_{1,s} & \gamma'_{2,s} \\ \gamma'_{2,s} & 1 + \gamma'_{1,s} \end{bmatrix} \vec{\theta}, \quad (4.10)$$

where $\gamma' = \gamma/(1 - \kappa)$ is the reduced shear. $\vec{\beta}_{\text{true}}$ and $\vec{\alpha}_{\text{true}}$ indicate that in the stated form above including the true physical lens model is recovered. $D_{\text{d}}^{\text{bkgd}}$ states the cosmological background angular diameter distance. Additionally to the main deflector, 6 additional parameters, namely the shear and convergence terms to the lens ($\gamma_{1,d}, \gamma_{2,d}, \kappa_d$) and the source plane ($\gamma_{1,s}, \gamma_{2,s}, \kappa_s$) describe the LOS effect.

The convergence terms κ_d and κ_s lead to particular degeneracies with other lensing effects. A non-zero convergence κ_d changes the angular diameter distance according to $D_{\text{d}}^{\text{lens}} = (1 - \kappa_d) D_{\text{d}}^{\text{bkgd}}$. The angular diameter distance $D_{\text{d}}^{\text{lens}}$ must be considered when computing other physical quantities of the lens, such as lensing potential and kinematics. The angular diameter distance $D_{\text{d}}^{\text{lens}}$ can not be determined from strong lens image reconstruction without relying on other information. The effect of κ_s leads to a rescaling of the lens equation (Equation 4.1, 4.5 or 4.10) without changing image observables. The physical interpretation of the rescaled quantities can change significantly. This is known as the mass-sheet degeneracy (242; 243; 244; 245).

The convergence effects can be decoupled from the image reconstruction. Instead of modeling $\vec{\alpha}_{\text{true}}$, $\vec{\beta}_{\text{true}}$ κ_d and κ_s one can model a rescaled lens equation

$$\vec{\beta}_{\text{scaled}} = \vec{\alpha}_{\text{scaled}} \left(D_{\text{d}}^{\text{bkgd}} \begin{bmatrix} 1 - \gamma_{1,d}'' & \gamma_{2,d}'' \\ \gamma_{2,d}'' & 1 + \gamma_{1,d}'' \end{bmatrix} \vec{\theta} \right) + \begin{bmatrix} 1 - \gamma_{1,s}' & \gamma_{2,s}' \\ \gamma_{2,s}' & 1 + \gamma_{1,s}' \end{bmatrix} \vec{\theta}. \tag{4.11}$$

The physical interpretation of the inferred deflection angle $\vec{\alpha}_{\text{scaled}}$, source scale $\vec{\beta}_{\text{scaled}}$ and shear terms on the lens plane γ_d'' change according to the convergences. The actual physical deflection relates to the scaled one as $\vec{\alpha}_{\text{true}} = (1 - \kappa_d)(1 - \kappa_s)\vec{\alpha}_{\text{scaled}}$. The source plane coordinate scales as $\vec{\beta}_{\text{true}} = (1 - \kappa_s)\vec{\beta}_{\text{scaled}}$ and the shear induced on the main deflector γ_d'' scales as

$$\gamma_d'' = \frac{\gamma_d}{(1 - \kappa_d)^2 (1 - \kappa_s)}. \tag{4.12}$$

4.2.3 Validity test

To test the accuracy of the approximations stated in Section 4.2.1, we construct a test scenario and compare the full multi-plane ray-tracing solution with our proposed formalism. For this purpose, we position a singular isothermal sphere (SIS) lens with velocity dispersion $\sigma_v = 200$ km s^{-1} at a redshift $z_d = 0.5$ and a source at $z_s = 2$. The Einstein radius of this configuration is $\theta_{\text{E}} = 0.73"$. We place a single perturber in the form of a Navarro-Frank-White (NFW) profile (200) with an angular separation of 8" away from the center of the SIS profile. We chose the mass within a mean over-density of $200\rho_c$ as $M_{200} = 10^{13} M_\odot$. We vary the redshift of the perturber ($z = [0.1, 0.3, 0.5, 0.7, 0.9]$) to test our formalism for different LOS positions.

The perturber is a group scale halo close to the main deflector. This is a relatively strong LOS perturber. Any perturber less massive and/or further away in angular separation will have a weaker impact on the deflection angles and will be approximated as well or better in our formalism. The integrated lensing effect from multiple perturbers adds linearly on the shear and convergence terms. The accuracy of multiple perturbers should be valid as long as the integrated shear terms are of comparable strength as the single group scale halo.

Convergence maps

We first test the accuracy of the predicted convergence map. In Figure 4.2, we compare the computed convergence maps of the full ray-tracing (FRT) with the approximation of our formalism (CSB). The convergence map for FRT is computed with differential ray-tracing. The top panel of Figure 4.2 shows the deviation of the convergence map of CSB to the full solution (FRT) ($\kappa_{CSB} - \kappa_{FRT}$). The lower panel shows a same comparison with the Born approximation (noSLB) of the LOS perturber ($\kappa_{noSLB} - \kappa_{FRT}$). The Einstein radius is plotted in black dashed lines.

The main differences in accuracy occur when the perturber is placed in front of the lens ($z < z_d$). The non-linear effect on the lens model can be well captured by CSB whereas noSLB ignores those effects and leads to significant residuals in the convergence map. The higher order distortion effects of the LOS perturber results in $\Delta\kappa < 0.01$ at the Einstein radius for CSB.

When the perturber is placed between the source and the lens ($z_d > z > z_s$), the two approaches have different predictions but neither of them can predict the convergence map accurately over the entire area of the lensing system to a precision better than $\Delta\kappa \approx 0.1$. The main difference is that CSB reproduces the mean convergence within the Einstein radius while noSLB over-estimates the convergence induced by the perturber significantly. CSB reproduces the mean convergence within the Einstein radius by construction while the induced error in the mean convergence in the noSLB is $\Delta_k \approx 0.05$. This behavior of the two approximations becomes emergent when looking at extended surface brightness simulations (see section 4.2.3 below).

Extended surface brightness

To test how well extended lensed surface brightness information can be predicted and re-produced by the CSB formalism, we take the same test case of 4.2.3 and model a Gaussian source surface profile with a width $\sigma = 0.02$" in the source plane positioned in line with the center of the main deflector. In the absence of external perturbers, this configuration leads to a perfectly circular Einstein ring in the image plane. In Figure 4.3, the simulated mock images are shown for the different computations of the LOS structure. For the moment, we do not
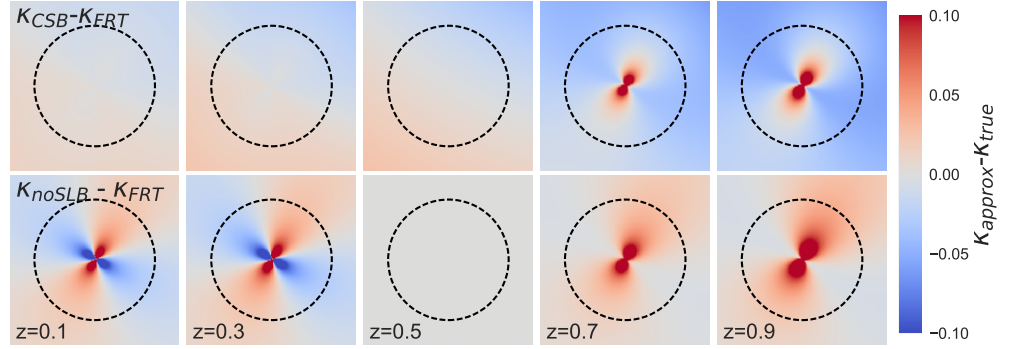
Figure 4.2: Comparison of different approximations with respect to the convergence prediction. Blue (red) corresponds to an under-(over-)estimation of the convergence by the approximation scheme. Upper row: Relative convergence of CSB compared to FRT. Lower row: Relative convergence of the Born approximation (noSLB) compared to FRT. From left to right: Comparisons with increasing LOS perturber redshift. The main deflector is placed at $z_d = 0.5$. The black dashed circle indicates the Einstein radius of the main deflector.

include any observational effects into the simulation (e.g. PSF and noise). In the top row, the full ray-tracing (FRT) simulations are shown. In the middle row, the predictions of the CSB formalism are shown. We see no distinguishable effect in the image plane. In the bottom row, the predictions with a Born approximation (noSLB) is shown. This test shows that the CSB approximation provides a good description for lens and source configurations that form an Einstein ring-like extended structure. We also see that the features in the image are not predicted accurately by a noSLB approximation. In particular, for foreground perturbers (first two columns in Figure 4.3), the real feature is a sheared Einstein ring/ellipsoid. This feature can only be reproduced when the non-linear effect of the perturber on the main deflector is taken into account. For background perturbers, the CSB approximation for the light paths is valid around the Einstein ring and can accurately predict the observational features. A noSLB approximation overestimates the induced tidal distortion and convergence (see also (223)).

4.2.4 Testing the constraining power of LOS inference

We analyze the information content of strong lens imaging data on constraining the LOS structure parameters. To do so, we set up a test case with realistic observational conditions
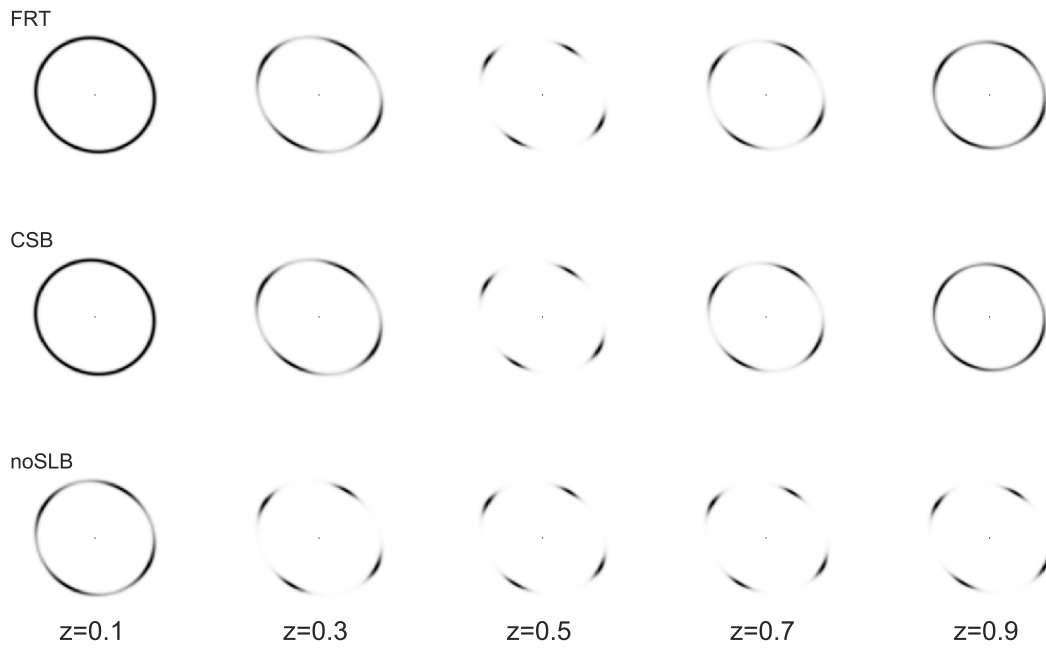
Figure 4.3: Comparison of different approximations with respect to the image prediction. In this model, a Gaussian source is placed at $z_s = 2$ and a main deflector as a SIS profile at $z_d = 0.5$. In addition, a LOS perturber is placed at different redshifts. Top column: Image computed with full ray-tracing (FRT). Middle column: Image computed with the CSB formalism presented in this work. Bottom column: Image computed with the noSLB approximation of the LOS perturber.

and source surface brightness. We then do a parameter inference and test the recovery of the LOS induced effects.

Specifically, we generate a mock image of a source at redshift $z_s = 2$, a main deflector at redshift $z_d = 0.5$ and a LOS perturber at $z_{los} = 0.1$. The main lens is modeled as a SIS with velocity dispersion $\sigma_v = 200 \text{km s}^{-1}$. The LOS perturber is positioned 8" from the main deflector with an NFW profile with mass $M_{200} = 10^{13.5} M_\odot$. We model the extended light emission from the source as a Gaussian light profile with width $\sigma_s = 0.02"$. We compute the observable light emission with sub-pixel resolution ray-tracing, convolution with a HST-like PSF and adding Poisson noise on the observed flux and a Gaussian noise realization of the background comparable to HST image quality of the COSMOS field.

We reconstruct the mock imaging data described above to infer the lens model parameter posteriors (including the LOS terms of Equation 4.11). For the lens model, we choose a smooth power-law elliptical mass profile (SPEMP), which allows for arbitrary elliptical mass distributions and power-law slopes. In the reconstruction modeling, we rescale the source size by $1/(1 - \kappa_s)$ to ensure that the same source description is applied in the reconstruction. Source size - power-law slope degeneracies are known and highly depend on the source reconstruction technique applied (see e.g. chapter 5). The more general lens model compared with the mock realization tests more rigorously the capability of recovering the LOS structural parameters. The inference is done with the formalism presented in chapter 3 with a Monte Carlo Markov Chain (MCMC), implemented in the `CosmoHammer` (202) software. In Figure 4.4, the inferred parameter posteriors are illustrated. Red vertical and horizontal lines indicate the input parameters for the lens model and the expected scaled shear parameters caused of the LOS perturber. The inference accurately recovers the expected lens and shear terms provided by the CSB formalism and shows that a separability of main deflector and LOS structure can be made with the given lens model assumptions. The posteriors in the ellipticity of the lens and the external shear terms are degenerate but the effects in the image (i.e. the ellipticity of the ring) can not be fully reproduced by an elliptic lens model configuration of the specifically used parameterization. The marginalized constraints on all the shear terms results in constraints of ± 0.003, a high precision measurement of the reduced shear field at a specific angular position on the sky.
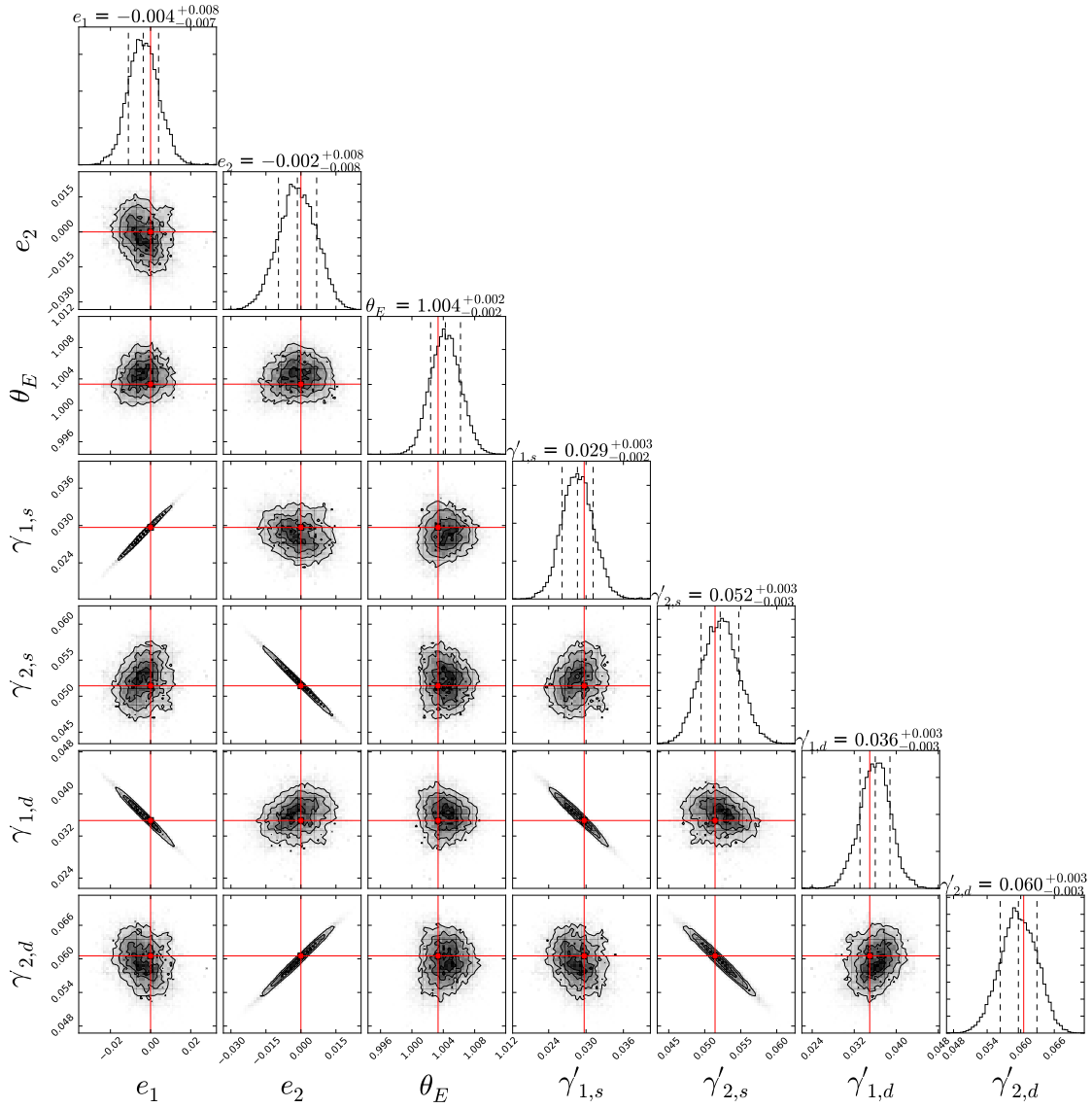
Figure 4.4: Parameter inference of the mock image described in Section 4.2.4. Red lines indicate the true input for the lens model parameters e_1, e_2 and θ_E. For the LOS shear parameters, the red lines correspond to the prediction of the CSB formalism modulo the convergence parts (Equation 4.11, 4.12). All parameter posteriors are consistent with the input model. The four external shear terms can be accurately determined to a precision of ± 0.003.

4.3 The strong lens COSMOS0038+4133 and its environment

In section 4.2, we used mock data to show that strong lensing systems can allow us to infer scaled shear terms of the LOS structure with high precision. The prediction of the same environmental quantities can be independently inferred by specifically modeling the LOS structure. Additionally, an explicit modeling of the mass structure enables us to simultaneously infer the external convergences κ_d and κ_s. The external convergences are important for many strong lens studies that involve the knowledge of the physical scales at the lens and/or source plane.

We next apply our formalism to the strong lens system COSMOS 0038+4133 and its environment to test the capabilities of our method on real data. The lens system COSMOS 0038+4133 (R.A. = $10^h00^m38.2^s$ DEC = $+02°41$arcmin33arcsec J2000) was chosen as our primary target as the configuration is close to an Einstein ring and there are massive galaxies in its close proximity that potentially add significant external shear and convergence contributions to the lens system. Detailed information about the lens system is provided in appendix 4.A.1 and the data, catalogs and derived stellar masses of the galaxies in appendix 4.A.2.

First, we perform the strong lens modeling in section 4.3.1. Second, we perform the independent LOS structure modeling based on galaxy catalogs in section 4.3.2. Third, we combine the constraints of the two approaches and show the results on the inferred external convergence and on the halo mass of specific galaxies in the vicinity in section 4.3.3.

4.3.1 Strong lens reconstruction

We model a 120^2 pixels cutout centered on the lensing galaxy. The lensing galaxy light profile is modeled with an elliptical Sérsic profile (201). The source is modeled with shapelet basis sets (192) with $n_{\max} = 10$, which corresponds to 66 basis functions. The shapelet scale is chosen to be $\beta = 0.016"$, which provides a good fit to the data (modulo mass-sheet transform, see e.g. chapter 5). For the lens model, we model a Singular Isothermal Ellipsoid (SIE) and in addition the external reduced shear components of Equation (4.11). We use the framework of chapter 3 as in section 4.2.4 to infer the parameter posteriors. In this particular inference, we further
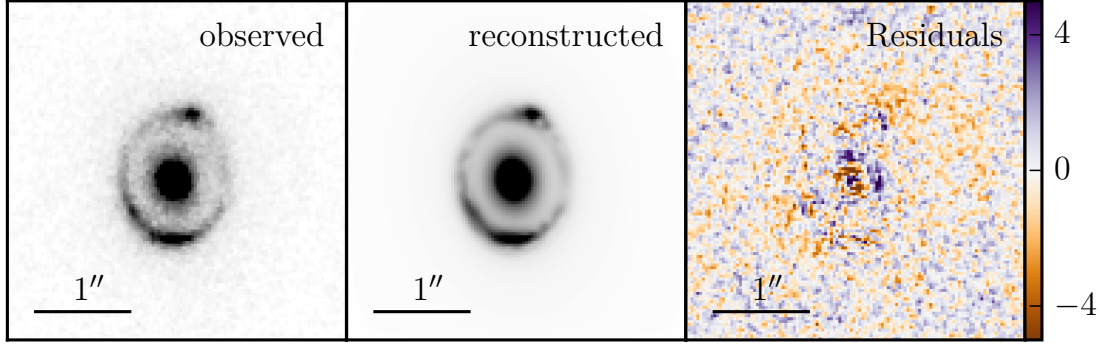
Figure 4.5: The strong lens COSMOS 0038+4133. In the left panel, the observed HST F814W image is illustrated. In the middle panel, the best fit reconstruction is shown. The right panel shows the reduced residuals. The model allows to reconstruct the arc features in detail. The largest residuals are present in the central part of the lens galaxy.

	max likelihood	1-σ posteriors
θ_E	0.663	0.663 ± 0.001
e_1	-0.105	-0.104 ± 0.006
e_2	0.018	0.017 ± 0.004
γ_1^d	-0.114	-0.115 ± 0.002
γ_2^d	-0.061	-0.061 ± 0.002
γ_1^s	-0.219	-0.218 ± 0.003
γ_2^s	-0.034	-0.035 ± 0.002

Table 4.1: Lens model parameter inference for COSMOS 0038+4133. Not included in this list are the lens light model parameters. The column labeled "max likelihood" shows the parameter position of the global maximum in the likelihood. The image reconstruction of this lens model is shown in Figure 4.5 middle panel. The second column shows the 1-σ marginal posteriors.

assume that the lens mass of the SIE is centered at the position of the luminous profile of the lensing galaxy.

Figure 4.5 shows the original HST F814W image (left), the best fit reconstructed model (middle) and the reduced residuals (right). In Table 4.1 the lens model parameter posteriors are stated for COSMOS0038+4133. We see that the precision on the shear terms is comparable to the mock example in Section 4.2.4 (Figure 4.4). With the stated model assumptions, the shear parameters associated with the LOS structure can be inferred with an uncertainty of ±0.002.

4.3.2 Halo rendering

We estimate the mass distribution in the vicinity of COSMOS 0038+4133 by linking the galaxies in the catalog (see appendix 4.A.2 for details about the galaxy catalog used) to the underling matter distribution. Similar approaches have been taken by (e.g. 228; 246) on the COSMOS field. Direct halo mass measurements come for example from galaxy-galaxy lensing (e.g 66; 68). Indirect methods use galaxy clustering (e.g 64; 247) or abundance matching (e.g 248; 249; 250). Phenomenological evolutionary models incorporating galaxy evolution in dark matter halos are presented in (e.g. 78) or alternatively in chapter 2 of this thesis.

We use the stellar-to-halo mass relation (SHMR) by (71), which is based on simultaneously modeling galaxy-galaxy lensing, galaxy clustering and abundance matching on data from the COSMOS survey. The scatter in the SHMR is described as a log-normal probability distribution function $M_* = f_{\mathrm{SHMR}}(M_h)$ and its inverse (250). We use the best fit parameters found in three redshift bins, which can be found in (71, Table 5). For $z > 1$ we use the same parameters as for $z \in [0.74, 1]$.

Uncertainties in the involved stellar mass estimates propagate non-linearly through the SHMR and affect the halo mass function, in particular it leads to a more frequent sampling of rare high mass halos. This is in contradiction to the method applied to determine the SHMR, which is based on a given fixed halo mass function. To avoid this inconsistency, we apply a conditional rendering on a fixed halo mass function.

For the spacial distribution of the mass, we assume spherical symmetric Navarro-Frenk-White (NFW) profiles (251). The masses inferred are taken to be the masses enclosed in a mean over-density of $200\rho_{\mathrm{crit}}$. The mass and redshift dependence of the NFW concentration parameter c is taken from (252; 253). The object-by-object dispersion in c at fixed halo mass and redshift is assumed to be log-normal as 0.08 dex. The lensing distortions of the NFW profiles are computed following e.g. (254; 255). Uncertainties in the measurements and modeling (i.e. stellar mass, SHMR, mass-concentration relation, redshift) can be incorporated by rendering different realizations of the uncertain quantities and propagate their uncertainties through their dependencies.

We only model over-dense regions of the universe explicitly. This leads to a manifestly over-dense universe compared to the assumed underlying cosmological model. (223) compensated this effect by ray-tracing through a homogeneous under-dense universe populated with over-dense halos. We chose a different approach. The necessary and sufficient requirement to keep the mean curvature of the universe to the one imposed by the background is that the mean convergence of all angular directions in the universe to all redshifts is zero $\langle \kappa \rangle = 0$. A homogeneous under-dense mass distribution contributes a negative convergence $\kappa_{m<0} < 0$. The model thus has to satisfy

$$0 = \langle \kappa_{\text{halo}} + \kappa_{m<0} \rangle = \langle \kappa_{\text{halo}} \rangle + \kappa_{m<0}. \tag{4.13}$$

This results in a shift of the convergence estimate of

$$\kappa_{\text{render}} = \kappa_{\text{halo}} - \langle \kappa_{\text{halo}} \rangle. \tag{4.14}$$

The term $\langle \kappa_{\text{halo}} \rangle$ is the mean convergence in a randomly sampled distribution of the galaxies in the field. This method is valid when the universe is homogeneous on the scale being rendered.

Figure 4.6 illustrates the environment of COSMOS0038+4133 and the influence on the shear of the nearby galaxies in two different zoom-out regions. The shear estimate converges with a mask of 6.5 arcmin around the strong lens system. For the final sampling, we take a mask of 13 arcmin around the strong lens system. In the selected area, more than 22'000 galaxies are found in the catalog and the contribution to the shear and convergence of each of them is rendered individually and summed up according to equation 4.7, 4.8 and 4.9.

The conservative stellar mass estimate uncertainties of the galaxies and the uncertain SHMR, especially at high stellar masses, results in weak constraints on the shear and convergence estimates. Nevertheless, a clear direction (sign) of the shear components is inferred (see Figure 4.7 and further discussions in section 4.3.3). Furthermore, the mass rendering indicates a highly over-dense LOS, which is not surprising given the way we selected the lens system.
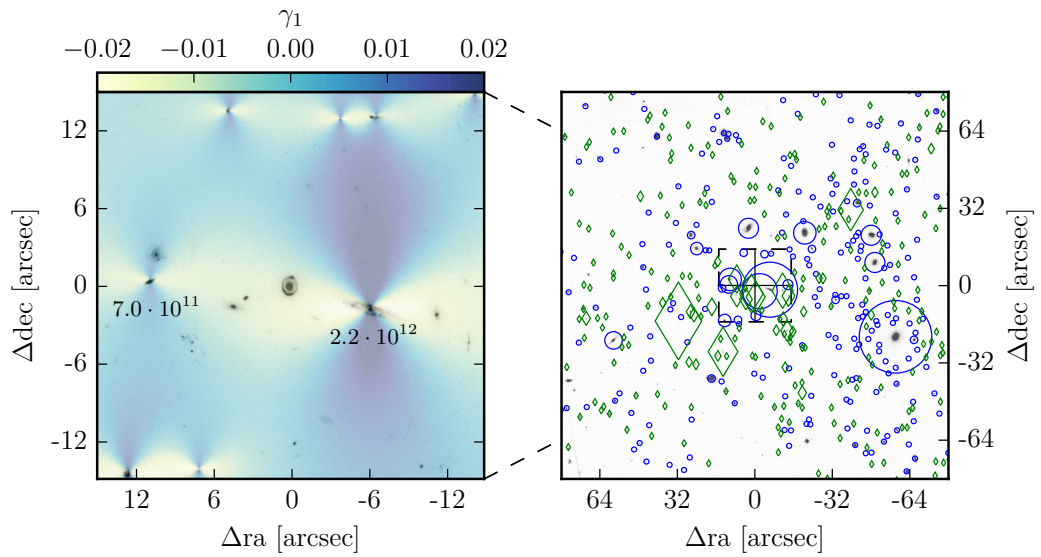
Figure 4.6: Illustration of the environment of the lens COSMOS0038+4133 and its influence on the shear at the position of the strong lens. Left panel: 30" cutout of the HST COSMOS field centered around the lens system. The shear contribution on the source plane $\gamma_{1,s}$ of the neighboring galaxies is shown. Right panel: 160" zoom-out of the left panel. Blue circles indicate galaxies between the observer and the strong lens. Green diamonds indicate galaxies between the strong lens and redshift $z = 2.7$. The size of the circle/diamond indicates the shear strength induced on the strong lens system from the galaxy.

4.3.3   Combining mass rendering and strong lens inference

Figure 4.7 shows the posterior distributions of the scaled reduced shear components. We see a consistent inference of the two independent methods in all the four shear terms. The strong lens image analysis leads to much tighter constraints on the shear terms as compared to the halo rendering approach.

From the halo rendering, we can compute the probability distribution

$$P_{\text{halo}}(\gamma_{1,d}, \gamma_{2,d}, \kappa_d, \gamma_{1,s}, \gamma_{2,s} \kappa_s, M_h^1, c^1, ..., M_h^n, c^n), \tag{4.15}$$

which involves all the galaxies (mass and profile parameters) and the lensing quantities, including the external convergence terms. The strong lens modeling provides the probability $P_{\text{SL}}(\gamma_{1,d}'', \gamma_{2,d}'', \gamma_{1,s}', \gamma_{2,s}')$. For the combined analysis, the two probability distributions can be taken as two independent unnormalized likelihoods as

$$P_{\text{Halo+SL}} \propto P_{\text{Halo}} \cdot P_{\text{SL}}. \tag{4.16}$$

The two probabilities P_{Halo} and P_{SL} are represented by a discrete sample. To numerically combine the two likelihoods, we use kernel density estimators for P_{SL} to evaluate for each sample in P_{Halo} a probability weight from the strong lens analysis. The marginalized errors on the parameters of interest come from the samples of P_{Halo} with their weights from P_{SL}.

Figure 4.8 shows the inferred external convergence at the lens plane κ_d (left) and the source plane κ_s (right). The halo rendering only constraints are drawn in green and halo rendering and strong lens reconstruction joint constraints are drawn in blue. The tight constraints on the reduced shears of the strong lens image reconstruction leads to a significant increase in precision of the inferred convergence values.

The additional constraints on the scaled shear terms from the strong lensing image reconstruction can also help constrain the halo masses of individual galaxies neighboring the strong lens system. Figure 4.9 shows the constraints on the halo mass for a selected massive and
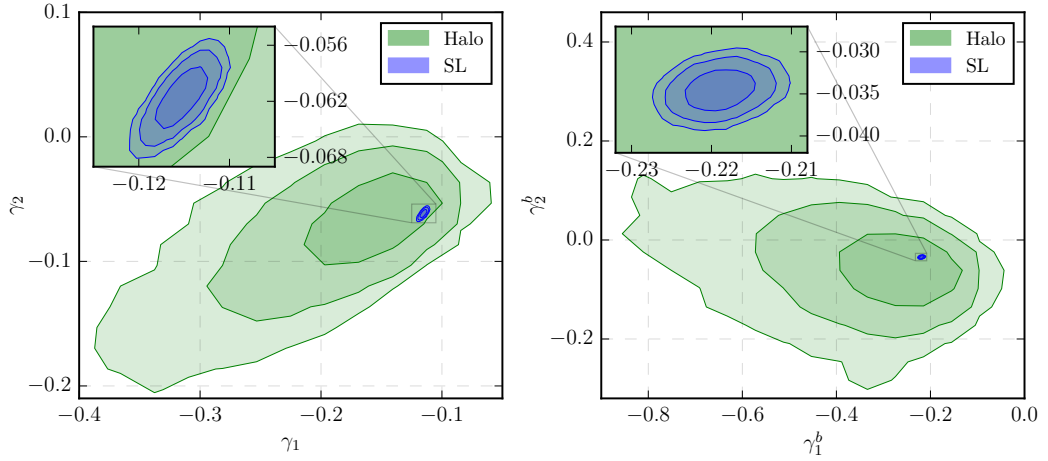
Figure 4.7: Reduced shear at the lens position $\gamma_{1,2}$ (left panel) and the source position $\gamma_{1,2}^b$ (right panel). Green contours: The 1-2-3 σ posteriors of the halo rendering approach based on the galaxy catalogue (section 4.3.2). Blue contours: The 1-2-3σ posteriors of the strong lens image reconstruction based on HST image (section 4.3.1). Both independent approaches are in agreement with each other. The strong lensing analysis provides much tighter constraints on the reduced shear components. In addition, the mass rendering approach simultaneously provides information about the external convergences and the halo masses of each individual galaxy in the catalogue.
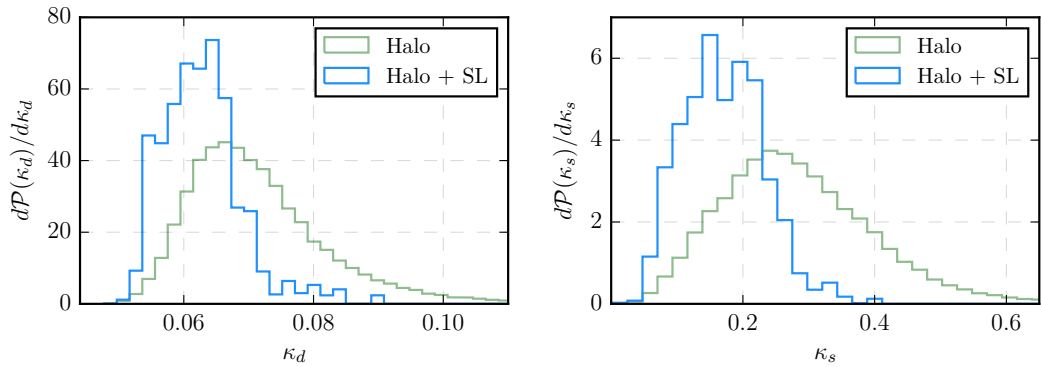
Figure 4.8: The inferred external convergence at the lens plane κ_d (left) and the source plane κ_s (right). The marginalized posteriors are shown for the halo rendering only constraints (green) and with the additional constraints on the shear from the strong lens modeling (blue). The tight constraints on the reduced shears of the strong lens analysis leads to double the precision in the inferred convergences.
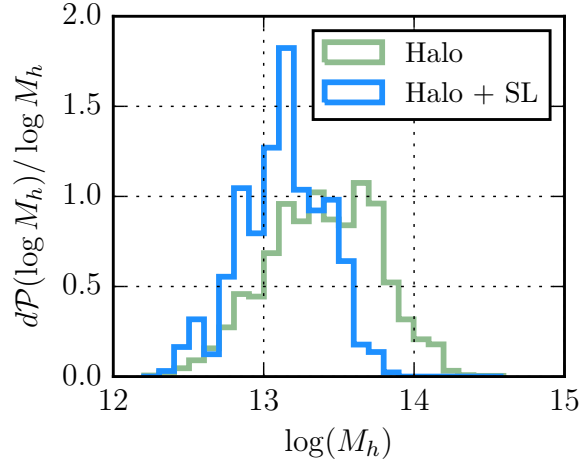
Figure 4.9: Halo mass constraints for the galaxy 62 arcsec away along the LOS of the strong lens system at $z = 0.342$. This is the galaxy with the large blue circle in the lower right half of Figure 4.6, right side.

nearby galaxy. The strong lens inference implies for this particular galaxy that a very high halo mass can be ruled out. In particular for a nearby massive galaxy, including the strong lens information, the posterior on the halo mass shifted by 0.4 dex to lower halo masses. The statistics of one single strong lens system does not allow to draw significant constraints on the SHMR. Joint constraints of multiple strong lens systems may add valuable information to the galaxy-halo connection.

4.4   Summary

We have presented a method to infer line of sight integrated lensing effects for galaxy scale strong lens systems through image reconstruction. Our approach enables us to separate weak lensing line of sight effects from the main strong lens deflector and allows a physical interpretation of both effects in parallel without relying on additional estimates in the image reconstruction. In particular, our approach reconstructs the observed shape and structure of extended arc and ring structure in strong lens systems and the approximations of the light paths and the parameterization allows us to decompose non-linear shear effects caused by intervening structure from the main deflector. We validated our method on mock data and demonstrated that strong lens systems can be accurate and precise probes of cosmic shear. In

a second step, we applied our formalism to the COSMOS field. We reconstruct the HST image, including the extended strong lens features. Independently, we modeled the LOS structure inferred from halo rendering using galaxy position, redshift and stellar mass estimates. When performing a combined analysis of our formalism with the halo rendering approach, we improve the constraints on the external convergence by a factor of two compared with a halo-rendering only analysis of the environment.

Strong lenses also allow a very precise direct shear measurement at few specific positions on the sky. This is complementary to galaxy shape weak lensing measurements. Including strong lensing constraints in large scale lensing surveys might thus help in calibrating galaxy shear measurements and constraining the mass distribution in the universe. Furthermore with increasing samples of strong lenses, one can gain insights into the galaxy-halo connection by combining strong lens image modeling and halo rendering of their environments.

4.A   Appendix of chapter 4

4.A.1   The strong lens COSMOS 0038+4133

The strong lens COSMOS0038+4133 was discovered and first quantified by (256). This system has a lensing arc including four images of a source object, see Figure 4.5 left panel. The redshift of the lens in (256) was calculated with the publicly available *Le Phare* photometric redshift estimation code using 8 bands, to be $z = 0.89^{+0.05}_{-0.03}$ at 68% confidence level. (257) released a revisited redshift estimate inferred from 30 bands at $z = 0.733^{+0.008}_{-0.012}$ at 68% confidence level. In this analysis we take the more recent redshift estimate of the lens. The Einstein radius is about $\theta_E = 0.73"$ and the effective radius of the lens galaxy $R_{\text{eff}} = 0.72"$. The magnitude of the lensing galaxy was determined $\text{mag}(I_{814w}) = 20.4$ and the maximum brightness of the ring as $\text{mag}"^{-2}(I_{814w}) = 20.5$. The (unknown) redshift of the source was placed to be at twice the co-moving distance to the lens at $z_s = 2.7$ for their lens kinematics and mass estimates. This choice maximizes the lensing efficiency and therefore provides lower bounds on the mass of the lensing galaxy. We adopt the same choice in our analysis for the source redshift.

4.A.2 COSMOS data and catalogues

The COSMOS field (see e.g. (258)) has been continuous covered by the HST Advanced Camera for Surveys (ACS) Wide Field Channel (WFC) in filter F814W. The median exposure depth is 2028 s and the limiting point-source depth is $\text{F814W}_{AB} = 27.2\,(5\sigma)$. This results in a 50% completeness for galaxies with a radius of 0.25arcsec at $I_{AB} = 26.0$ mag. The images were combined with the MultiDrizzle software (259) where the final resolution of the drizzled data is 0.03arcsec/pixel. We use the third public release v2.0 of the COSMOS ACS data (31. Oct. 2011) [1]. Details of HST ACS/WFC observations, the data calibration and processing procedures are explained in (260). The raw data of the ACS WFC were corrected for the charged transfer inefficiency by (261).

The COSMOS field provides, apart from the HST coverage, a wealth of additional data products to reconstruct the environment of the lens. Detailed information of the HST observations can be found in (262). We take the redshifts and magnitudes from the COSMOS photometric redshift release (257) for the neighboring galaxies, including apparent magnitudes provided by (263).

The photometric redshifts in (257) were calculated using fluxes in 30 different bands (broad and narrow bands covering UV, visible near-IR and mid-IR). Up to $z \sim 2$ the accuracy is $\sigma_{\Delta z/(1+z_s)} = 0.06$ at $i^+_{\text{AB}} \sim 24$, where $\Delta z = z_s - z_p$ and z_s are the spectroscopic redshifts of a comparison sample. We do not include the redshift uncertainty in our analysis as they are of order the cosmological uncertainties.

We use the NIR K band to calculate luminosities and the broad bands g^+ and i^+ as a color indicator to estimate the mass-to-light ratios. We take the color dependent mass-to-light ratio by (264; 265) in the functional form of (e.g. 266)

$$\log\left(\frac{M}{L_{\mathcal{B}}}\right) = a_{\mathcal{B}}^{\mathcal{C}} + b_{\mathcal{B}}^{\mathcal{C}} \cdot \mathcal{C} \,, \tag{4.17}$$

where in our case $\mathcal{C} = (g^+ - i^+)$ and $\mathcal{B} = K$. To calibrate the coefficients $a_{\mathcal{B}}$ and $b_{\mathcal{B}}$ we

---

[1]STScI-MAST: http://archive.stsci.edu/ or IPAC/IRSA: http://irsa.ipac.caltech.edu/data/COSMOS/

use a sample of stellar masses from the COSMOS group membership catalog (267) where the stellar masses are calculated according to the method described in (68). We split the calibration sample in six redshift bins and infer the coefficients for the different redshift samples independently.

5 The mass-sheet degeneracy and time-delay cosmography

This chapter appeared in a similar form in Birrer, Amara & Refregier 2016 (30).

Strong lensing systems and the time delays between different images of the same background source can provide information about angular diameter distance relations (see (213) and review of (268) for the early work). Cosmographic analyses rely on measurements of time delay (see e.g., 269; 270; 271; 272; 273; 274, and the COSMOGRAIL collaboration)[1] and estimates of the line-of-sight structure and lensing potential. This cosmography technique has been applied to determine the Hubble parameter H_0 using different strong lens systems (see e.g. 243; 214; 217; 275; 276; 277; 278; 279; 280; 281; 179; 180; 282) and also by applying statistics to multiple systems (see e.g. 185; 218; 283). In the past, some of the measurements have produced a wide range of results for H_0 (e.g. see section 8.2 of 179). One concern has been to evaluate the impact of potential systematic errors. In particular, the mass-sheet degeneracy (MSD) (242) and related degeneracies that cause biases due to model assumptions (e.g. 243; 244; 245; 284; 285) need special consideration. For instance, this has been illustrated by (286) where they show that assuming a power-law lens model can cause significant biasing of results.

In this chapter, we introduce a new treatment of the MSD and source reconstruction for cosmographic analyses. This approach integrates information coming from imaging, velocity dispersion, external convergence and time delay measurements. For the choice of data and

---

[1] www.cosmograil.org

the parameterization of the lens we follow the work of (180), and we infer the values of the parameters using our recent framework presented in chapter 3. In our framework we reconstruct the source using shapelet basis sets. This allows us to explicitly set an overall scale for the reconstruction. We will show that this enables us to better disentangle the effects coming from source structure and MSD. This then makes it simpler to robustly combine the information coming from the different data sets.

The chapter is organized as follow: Section 5.1 presents the data used in this chapter. Section 5.2 describes the details of the lens modeling, including kinematics, likelihood analysis and the source reconstruction technique of (29) introduced in chapter 3. In Section 5.3, we show that the use of this reconstruction technique turns out to be well designed for mapping out the MSD. Section 5.4 describes the combined likelihood analysis and posterior sampling. Section 5.5 discuss the cosmological constraints in terms of angular diameter relations and cosmological parameters. In Section 5.6, we compare our results to others. We summarize our conclusions in Section 5.7.

5.1 RXJ1131-1231 system

The quadrupole lens system RXJ1131-1231 (Figure 5.1) was discovered by (206) and the redshift of the lens $z_l = 0.295$ and of the background quasar source $z_s = 0.658$ was determined spectroscopically by (206). The lens was modeled extensively by (207; 208; 180; 287) and in chapter 3 with single band images. We use the archival HST ACS WFC1 images in filter F814W and F555W (GO 9744; PI: Kochanek). The filter F814W was also used for lens modeling in (180), (282) and (29). We make use of the `MultiDrizzle` product from the HST archive. We use a 160^2 pixel image centered at the lens position with pixel scale 0.05". This corresponds to a FOV of 8".

For the analysis in this chapter, we take the time delay measurements and uncertainties from (288), namely $\Delta t_{AB} = 0.7 \pm 1.4$ days, $\Delta t_{CB} = -0.4 \pm 2.0$ days, and $\Delta t_{DB} = 91.4 \pm 1.5$ days, where $[A, B, C, D]$ represent the quasar images in Figure 5.1. This data was used in (180), where they also measure the LOS velocity dispersion of $\sigma_\nu = 323 \pm 20$ km s^{-1}, that we use in our analysis.

For the external convergence κ_{ext}, we take the estimate of (180) based on relative galaxy counts
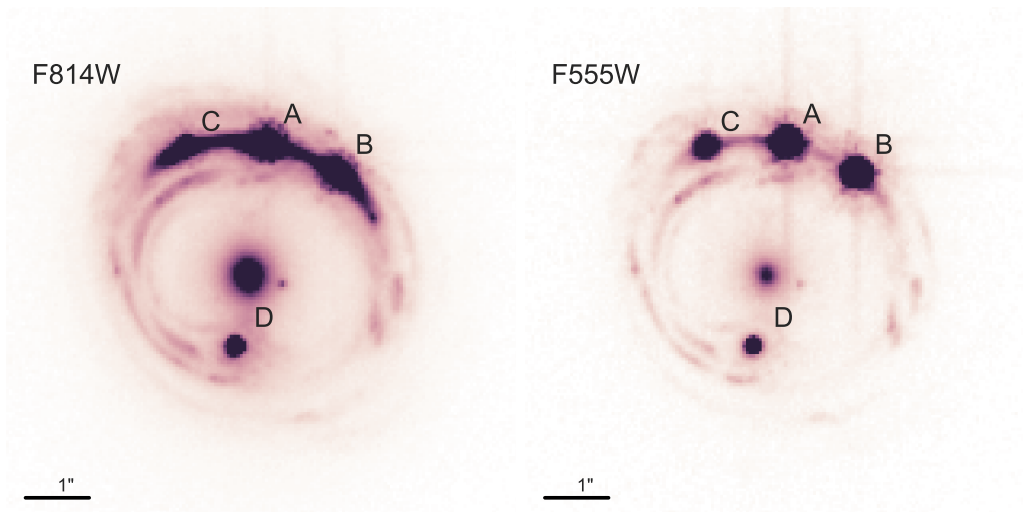
Figure 5.1: HST ACS WFC1 images in filters F814W (left) and F555W (right). The F814W filter has more high signal-to-noise pixels than the F555W filter. In the F555W filter, the substructure in the Einstein ring and the diffraction spikes of the quasar images are more prominent. The letters A,B,C,D indicate the quasar images for the time delay differences.

in the field (229) and their modeled external shear component compared with ray tracing of the Millennium Simulation (see their Figure 6). As their probability density function for κ_{ext} is not given in a parameterized form, we use an approximation of their PDF in the form of a skewed normal distribution with mean $\mu_\kappa = 0.1$, standard deviation $\sigma_\kappa = 0.042$ and skewness $\gamma_\kappa = 0.8$. This function is illustrated in Figure 5.2 and described in Appendix 5.A.5.

5.2 Lens modeling

In this section, we present the parameterization of the lens model, the lens light description, the source reconstruction technique, PSF modeling, the modeling of the lens kinematics and the likelihood analysis.

5.2.1 Lens model parameterization

For the lens model, we use:
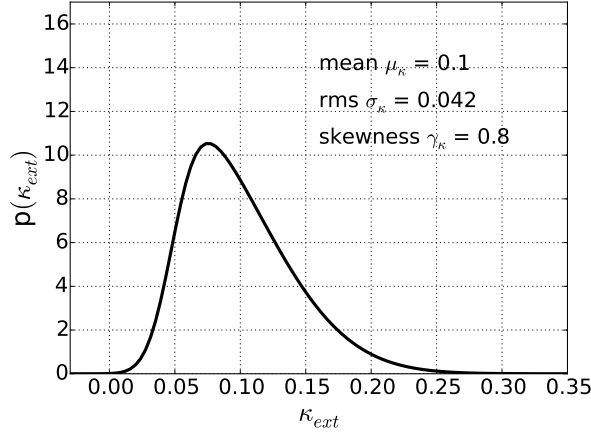
Figure 5.2: Probability density function of the external convergence in the form of a skewed normal distribution. The parameters chosen are designed to match well the probability density function quoted in (180) (their Figure 6).

1. An elliptical power-law mass distribution parameterized as

$$\kappa_{\text{lens}}(\theta_1, \theta_2) = \frac{3 - \gamma'}{2} \left(\frac{\theta_E}{\sqrt{q\theta_1^2 + \theta_2^2/q}} \right)^{\gamma' - 1} \tag{5.1}$$

where θ_E is the Einstein radius, q is the ellipticity and γ' is the radial power-law slope.

2. A second spherical isothermal profile (Equation 5.1 with fixed $\gamma' = 2$ and $q = 1$) centered at the position of the visible companion of the lens galaxy about 0.6 arc seconds away from the center.

3. A constant external shear yielding a potential parameterized in polar coordinates (θ, φ) given by

$$\psi_{\text{ext}}(\theta, \varphi) = \frac{1}{2} \gamma_{\text{ext}} \theta^2 \cos 2(\varphi - \phi_{\text{ext}}) \tag{5.2}$$

with γ_{ext} is the shear strength and ϕ_{ext} is the shear angle.

5.2.2 Lens light parameterization

The light distribution of the lens is modeled in a parameterized form. We use the same profiles as (180), namely two elliptical Sérsic profiles (289) with common centroid for the central

elliptical galaxy and an additional spherical Sérsic profile for the companion galaxy. The intensity profile is parameterized as

$$I(\theta_1, \theta_2) = A \exp\left[-k\left(\left(\frac{\sqrt{\theta_1^2 + \theta_2^2/q_L^2}}{\theta_{\text{eff}}}\right)^{1/n_{\text{sersic}}} - 1\right)\right] \tag{5.3}$$

where A is the amplitude, k is a constant such that θ_{eff} is the effective half-light radius, q_L is the axis ratio and n_{sersic} is the Sérsic index. We use the value of half-light radius θ_{eff} as the effective radius in the kinematics modeling of Section 5.2.5.

5.2.3 Source surface brightness reconstruction

We use the source reconstruction method presented in chapter 3 based on shapelet basis functions introduced by (192). To apply this method, three choices have to be made. (1) The shapelet center position, which we fixed to quasar source position. The determination of the quasar source position is explained in detail in 3.3.2. (2) The width of the shapelet basis function β (see Section 5.3 for its impact). (3) The maximal order n_{max} of the shapelet polynomials. We set $n_{\text{max}} = 30$ for modeling and parameter inference. With this, most of the features in the extended source can be modeled. Given these three choices, one can reconstruct the angular scales between $\beta/\sqrt{n_{\text{max}}+1}$ and $\beta\sqrt{n_{\text{max}}+1}$ around the center of the shapelet in the source plane.

5.2.4 PSF modeling

We use four bright stars in the same ACS image to model the PSF. After normalizing for flux, we apply a sub-pixel shift to recenter the stars and then stack. When comparing the individual star images and the stack, we see significant variations that we need to consider in our analysis. To do this by measuring the scatter for each pixel and assume that the scatter in high signal-to-noise pixels is due to a model error that we quantify as a fraction of the flux. This leads to an additional error term, beyond the Poisson and background contribution, that is important close to the center of the bright point sources (see Section 5.2.6). For the quasar point sources, we use a cutout of the PSF of 111^2 pixels to cover most of the diffraction spikes. For the

extended surface brightness we apply a PSF-convolution kernel of 21^2 pixels.

5.2.5 Stellar kinematics

We follow the analysis of (179) for the modeling of the stellar velocity dispersion. The mass profile is assumed to be a spherical symmetric power-law in the form of

$$\rho_{\text{local}}(r) = \rho_0 \left(\frac{r_0}{r}\right)^{\gamma'} \tag{5.4}$$

where ρ_0 is the density at radius r_0 and γ' is a power-law slope of the mass profile (the same γ' as for the lens model in Equation 5.1). The normalization of the mass profile can be expressed in terms of the lensing quantities as

$$\rho_0 r_0^{\gamma'} = (\kappa_{\text{ext}} - 1)\Sigma_{\text{crit}}\theta_{\text{E}}^{\gamma'-1} D_{\text{d}}^{\gamma'-1} \frac{\Gamma\left(\frac{\gamma'}{2}\right)}{\pi^{1/2}\Gamma\left(\frac{\gamma'-3}{2}\right)}. \tag{5.5}$$

where κ_{ext} is the external convergence, Σ_{crit} is the critical projected density, θ_{E} is the Einstein radius, D_{d} is the angular diameter distance from the observer to the lens and Γ is the Gamma function. The estimation of the projected velocity dispersion along the line of sight requires a description of the anisotropic velocity component split in radial and tangential component

$$\beta_{\text{ani}} \equiv 1 - \frac{\sigma_z^2}{\sigma_r^2}. \tag{5.6}$$

Massive elliptical galaxies are assumed to have isotropic stellar motions in the center of the galaxy ($\beta_{\text{ani}} = 0$) and radial motions in the outskirts ($\beta_{\text{ani}} = 1$). A simplified description of the transition can be made with an anisotropy radius parameterization r_{ani} defining β_{ani} as a function of radius r as

$$\beta_{\text{ani}}(r) = \frac{r^2}{r_{\text{ani}}^2 + r^2}. \tag{5.7}$$

Assuming a Hernquist profile (290) and an anisotropy radius r_{ani} for the stellar orbits in the lens galaxy, the three-dimensional radial velocity dispersion σ_r at radius r from Jeans modeling is given by

$$\sigma_r^2 = \frac{4\pi G a^{-\gamma'} \rho_0 r_0^{\gamma'}}{3-\gamma'} \frac{r(r+a)^3}{r^2+r_{\mathrm{ani}}^2}$$
$$\times \left(\frac{r_{\mathrm{ani}}^2}{a^2} \frac{{}_2F_1\left[2+\gamma',\gamma';3+\gamma';\frac{1}{1+r/a}\right]}{(2+\gamma')(r/a+1)^{2+\gamma'}} + \frac{{}_2F_1\left[3,\gamma';1+\gamma';-a/r\right]}{\gamma'(r/a)^{\gamma'}} \right), \quad (5.8)$$

where a is related to the effective radius of the lens light profile θ_{eff} by $a = 0.551\theta_{\mathrm{eff}}$ and ${}_2F_1$ is a hyper geometric function. The modeled luminosity-weighted projected velocity dispersion σ_s is given by

$$I_H(R)\sigma_s^2 = 2\int_R^\infty \left(1-\beta_{\mathrm{ani}}(r)\frac{R^2}{r^2}\right)\frac{\rho_* \sigma_r^2 r\,dr}{\sqrt{r^2-R^2}} \qquad (5.9)$$

where R is the projected radius, ρ_* is the stellar density and $I_H(R)$ is the projected Hernquist distribution. The luminosity weighted LOS velocity dispersion within an aperture $\mathcal{A}$ is then (see also equation 20 in (179))

$$(\sigma^{\mathrm{P}})^2 = \frac{\int_{\mathcal{A}}\left[I_H(R)\sigma_s^2 * \mathcal{P}\right]R\,dR\,d\theta}{\int_{\mathcal{A}}\left[I_H(R) * \mathcal{P}\right]R\,dR\,d\theta} \qquad (5.10)$$

where $*\mathcal{P}$ indicate the convolution with the seeing. In Appendix 5.A.1 we describe in detail how we compute a modeled σ^{P} in a numerically stable way. This calculation assumes no rotational behaviour of the lensing galaxy. Priors on the anisotropic behaviour $\beta_{\mathrm{ani}}(r)$ are discussed in section 5.4.3.

Equation (5.10) can be expressed as a function of angular scales of r_{ani} and θ_{eff} paired with a cosmological dependent angular diameter distance relation and an external convergence factor as

$$(\sigma^{\mathrm{P}})^2 = (1-\kappa_{\mathrm{ext}})\cdot\frac{D_{\mathrm{s}}}{D_{\mathrm{ds}}}\cdot H(\gamma',\theta_{\mathrm{E}},\beta_{\mathrm{ani}}(r),\theta_{\mathrm{eff}}) \qquad (5.11)$$

where H is capturing all the computation of equation (5.10) without cosmological and external convergence specifications. With this calculation, we see that any estimate of the (central)

velocity dispersion is dependent on the ratio of angular diameter distance from us to the source and from the deflector to the source. This fact is important when kinematic modeling is used to infer cosmographic information. We separate in the modeling the angular and the cosmological information. The separability allows us to consistently infer cosmographic information without the need of cosmological priors in the kinematic modeling.

5.2.6 Likelihood analysis

We estimate the pixel uncertainty in the image with a Gaussian background contribution σ_{bkgd} estimated from an empty region in the image and a Poisson contribution from the model signal d_{Pi} scaled by the exposure map t_{i}. In addition, the modeling uncertainty of the PSF of the bright point sources with amplitude A_j, PSF kernel k_{ij} and model uncertainty coming from the star-by-star scatter δ_{PSF} is given as

$$\sigma_{\mathrm{PSF},i} = \sum_{j=1}^{N_{\mathrm{AGN}}} A_j k_{ij} \delta_{\mathrm{PSF},ij}, \tag{5.12}$$

at a pixel i, where N_{AGN} is the number of quasar images. All together, the uncertainty for each pixel i sums up in quadrature as

$$\sigma_{\mathrm{pixel},i}^2 = \sigma_{\mathrm{bkgd}}^2 + t_{\mathrm{i}}^{-1} d_{\mathrm{Pi}} + \sigma_{\mathrm{PSF},i}^2. \tag{5.13}$$

For the linear source surface brightness reconstruction d_{Pi} is replaced by the image intensity $d_{\mathrm{ACS},i}$.

The likelihood of an image $\boldsymbol{d}_{\mathrm{ACS}}$ given a model $\boldsymbol{d}_{\mathrm{P}}$ is

$$P(\boldsymbol{d}_{\mathrm{ACS}}|\boldsymbol{d}_{\mathrm{P}}) = \frac{1}{Z_{\mathrm{d}}} \exp \sum_{i=1}^{N_{\mathrm{d}}} \left[-\frac{\left(d_{\mathrm{ACS},i} - d_{\mathrm{P},i}\right)^2}{2\sigma_{\mathrm{pixel},i}^2} \right] \tag{5.14}$$

with N_{d} being the number of pixels in the modeled image and Z_{d} is the normalization

$$Z_{\mathrm{d}} = (2\pi)^{N_{\mathrm{d}}/2} \prod_{i}^{N_{\mathrm{d}}} \sigma_{\mathrm{pixel},i}. \tag{5.15}$$

At this stage, it is useful to separate the model into nonlinear parameters $\boldsymbol{\eta}$ and linear parame-

ters $\boldsymbol{s}$. The likelihood of the non-linear parameters is given by

$$P(\boldsymbol{d}_{\mathrm{ACS}}|\boldsymbol{\eta}) = \int d\boldsymbol{s}\, P(\boldsymbol{d}_{\mathrm{ACS}}|\boldsymbol{\eta},\boldsymbol{s})P(\boldsymbol{s}). \tag{5.16}$$

The integral is computed in equation 3.6 assuming flat priors in $\boldsymbol{s}$, which we adopt in this chapter.

The likelihood for the time delays $\boldsymbol{\Delta t}$ is the product of the likelihoods of all relative delays of the quasar pairs (ab)

$$P(\boldsymbol{\Delta t}|D_{\Delta t}^{\mathrm{model}},\boldsymbol{\eta}) = \prod_{(ab)} \left(\frac{1}{\sqrt{2\pi}\sigma_{ab}} \exp\left[-\frac{\left(\Delta t_{ab} - \Delta t_{ab}^{\mathrm{P}}\right)^2}{2\sigma_{ab}^2} \right] \right). \tag{5.17}$$

The likelihood of the LOS central velocity dispersion is given by

$$P(\sigma_\nu|\boldsymbol{\eta}) = \frac{1}{\sqrt{2\pi}\sigma_\sigma} \exp\left[-\frac{\left(\sigma_\nu - \sigma^{\mathrm{P}}\right)^2}{2\sigma_\sigma^2} \right]. \tag{5.18}$$

5.3 The mass sheet degeneracy

There exists many different degeneracies in strong lens modeling (e.g., 244; 291). In this section we focus on the MSD (242) and in particular its impact on time delay cosmography as it was pointed out by (286). As shown by (242), a remapping of a reference mass distribution κ by

$$\kappa_\lambda(\vec{\theta}) = \lambda\kappa(\vec{\theta}) + (1-\lambda) \tag{5.19}$$

combined with an isotropic scaling of the source plane coordinates

$$\vec{\beta} \to \lambda\vec{\beta} \tag{5.20}$$

will result in the same dimensionless observables (image positions, image shapes and magnification ratios) regardless of the value of λ. This type of mapping is called mass-sheet-transform (MST), and shows that imaging data, no matter how good, can not break the MSD.

The additional mass term in MST (Equation 5.19) can be internal to the lens galaxy (affecting the lens kinematics) or due to line-of-sight structure (not affecting the lens kinematics) (see e.g., 244; 245). The external part of the MST can be approximated by an external convergence κ_{ext}, which rescales the time delays accordingly. The external contribution also rescales the source plane. Lens modeling often only explicitly models the internal structure of the lens. The inferred source scale has to be rescaled by the external mass sheet to match the physical scale.

5.3.1 Source scaling and the MSD

An important parameter in the lens model inference is the physical source scale. Neither the lens model nor the source size are direct observables, but they share the MST in each others inference. Given a lens model, certain source sizes are preferred. The opposite is also true: Given a source size, certain lens models are preferred. This is a direct consequence of the MST (Equation 5.19 and 5.20). Therefore, it is important to control the prior on the assumed source scale in the modeling. A particular source surface brightness reconstruction method, depending on the choice of regularization, basis set, pixel grid size or parameters of the source reconstruction, will potentially favor a certain size of the reconstructed source and therefore may indirectly lead to priors on the internal mass model through the MST. As one does not know a priori the physical scales in the source galaxy, this may lead to significant biases in the inference of the lens model.

We use shapelets (192) as the source surface brightness basis functions as implemented in chapter 3. These basis functions form a complete basis set when the order n goes to infinity. When restricting the shapelet basis to a finite order n_{max}, the reconstruction of an image depends on the chosen scale β of the shapelet basis function. As pointed out by (192), for a given n_{max}, there is a scale $\hat{\beta}$ that best fits the data. From Equation 5.20, we see that changes in β can be remapped into changes in the lensing potential through the linear parameter λ. Therefore, since our source reconstruction technique has an explicit scale, we have a tool to walk along the MST.

5.3.2   Varying source scale in the ACS WFC1 images

We have identified the source scale to have an impact on the inference of the lens model within the MST. To investigate the specific dependence of the shapelet scale in the source reconstruction in combination with lens model parameterization (Section 5.2.1) in our analysis of RXJ1131-1231, we model the ACS WFC1 F814W and F555W images with different choices of the shapelet scale β. For the F814W image, we use the range 0.14" - 0.19" and for the F555W image the range 0.13" - 0.18". The shapelet order was held constant at $n_{\max} = 30$. To find the best fit model, we used a particle swarm optimization as used in chapter 3 to maximize the likelihood (Equation 5.16). In this section, we only use the HST images for our modeling. Time-delay and kinematic data will be added in Section 5.4.

Figure 5.3 shows the source reconstruction of the best fit models of filter F814W for six different scales β. We see that the source reconstructions are very similar but scaled by the relative factors of the chosen shapelet scale. More explicitly, we overlay in Figure 5.4 the intensity contours of the different source reconstructions rescaled by β. We also show the reconstructions for the F555W image, which shows the same behavior. On the right of Figure 5.4 we over-plot a joint source reconstruction of the two bands in a fake color image. In Appendix 5.A.2, we present the corresponding normalized residuals for this analysis of the F814W reconstruction.

The difference in the likelihood value for different scales β from the imaging data exceeds the 10-σ level between each modeled scale β. This reflect the fact that the chosen lens model parameterization (see Section 5.2.1) does not allow for the full freedom needed for a perfect transform according to the MST (Equation 5.19). The source scale β can not be fixed to an arbitrary value and caution on any scale dependent source reconstruction description is needed. When assigning a prior on β and infer this parameter together with all the lens model parameters from the image reconstruction, we are able to very precisely determine the corresponding source scale and the parameters of the given functional form of the lens model.

5.3.3   Relaxing on the lens model assumption

As pointed out by (286), there can also be an internal component to the MST. Namely when the lens model can not reproduce the underlining internal mass distribution. The assumption
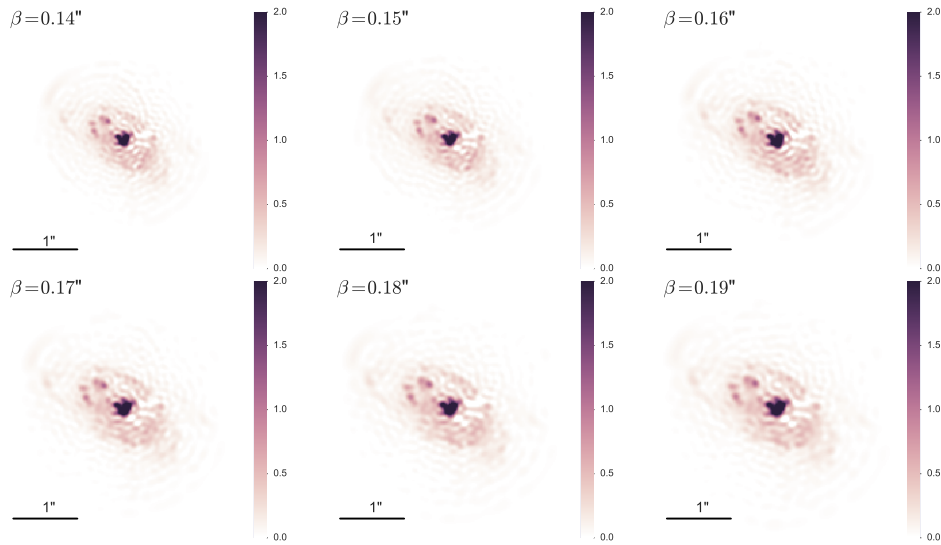
Figure 5.3: Reconstructed source surface brightness profiles as a function of shapelet scale β for filter F814W. The source reconstructions of the best fit lens model configurations are shown with a given β. We see that the features become larger with larger choices of β.
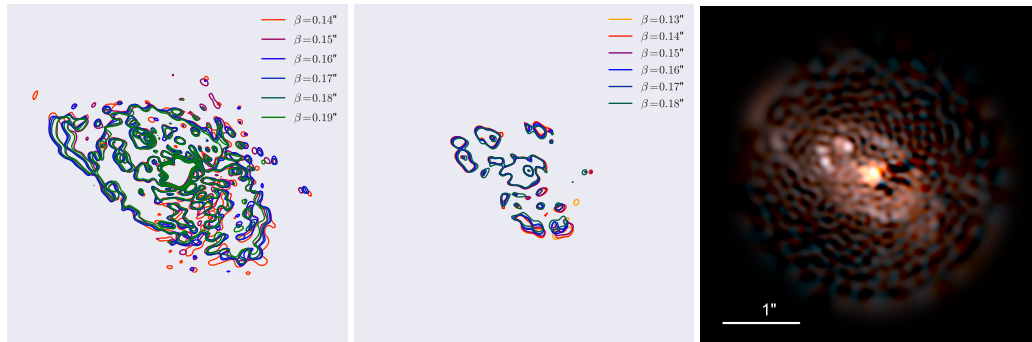
Figure 5.4: Left: Intensity contours of the reconstructed source surface profiles rescaled to fiducial value $\beta = 0.2"$ for the different shapelet scales β in filter F814W of Figure 5.3. The contour lines overlay well. The lens model does adopt to the choice of β such that the source reconstruction catches the best scales. Middle: Same as left for the filter F555W. The same behavior can be seen as for F814W. Right: Color composite model of the filters F814W and F555W for a chosen joint lens model.

of a power-law lens model formally sets the internal part of the MST. The parameters will fit preferentially those models, whose shape, modulo an artificial MST, are the most similar to the underlying mass distribution. The only effect visible in the modeling of the imaging data is on the source scale. The inferred source scale will be different from the one of the true lens model. Any assumed mass distribution which can not be rescaled according to Equation (5.19) can thus potentially lead to biased inferences, in particular on the slope of the mass profile. This also can result in significant biases in the inferred lensing potential and lens kinematics. In particular, it was stated by (286) that the assumption of a power-law lens model can potentially lead to a significant bias in the inference of the time delay distance.

Three approaches to handle the concerns of (286) in performing cosmographic estimates are:

1. One assumes that the true lens model can be described within the functional form of the chosen parameterization. This is the approach done by (180). In this case we end up with the potentially biased inference discussed in (286), a situation we want to avoid as good as possible.

2. One choses a more flexible lens model than a single power-law mass profile. This approach was followed in (282) in response to (286). Different profile parameterizations may lead to different preferred source scales. It is not guaranteed that a more sophisticated lens model parameterization infers an unbiased result in the cosmographic inference.

3. Perform simplifications and approximations that lead to greater robustness against known degeneracies. For instance accommodating MST through careful handling of the source size inference.

In this chapter, we chose the third option mentioned above. This option requires the least assumptions on the lens model and a prior is placed on the source size, rather through the functional form of the lens model. In Appendix 5.A.4 we specifically state the process in a Bayesian inference way to make clear our steps and approximations and show that a renormalization of the imaging likelihood for different imposed source scales β is needed to explore the impact of plausible internal MST on the cosmological inference.

5.3.4 Adding lens kinematics

Additional constraints on the lens model can come from kinematic data at a different scale than the Einstein ring. This becomes of particular importance when weakening the constraining power of the lens model, as described in Section 5.3.3. Lens models with different source scales predict different lens kinematics. The prediction depends on the stellar velocity anisotropy β_{ani} which can not be known from the existing data and the external convergence κ_{ext} which has to be inferred separately.

As long as the relative likelihood of additional kinematic data (Equation 5.18) can not compete with the relative likelihood of the different shapelet scales β (on the 10-σ level between the chosen source scales, see Section 5.3.2), the combined likelihood will be dominated by the lens model assumption. Only when re-normalizing the likelihood of the imaging data for different scales β, the kinematic data can have a significant impact in the determination of the lens profile and in particular the lens potential for time-delay cosmography.

5.4 Combined likelihood analysis

In this section, we discuss how we combine the different data sets and their likelihoods. We showed in the previous section that biases can emerge from choices in the lens and source modeling. These aspects have to be taken into account when the data sets are combined.

5.4.1 Combining imaging and time delay data

In a first step, we do a joint analysis of the independent measurements of the time delay and imaging data. The combined likelihood is

$$P(\boldsymbol{d}_{\text{ACS}}, \boldsymbol{\Delta t} | \boldsymbol{\eta}, D_{\Delta t}^{\text{model}}) = P(\boldsymbol{d}_{\text{ACS}} | \boldsymbol{\eta}) P(\boldsymbol{\Delta t} | D_{\Delta t}^{\text{model}}, \boldsymbol{\eta}) \tag{5.21}$$

with the independent likelihoods of Equation (5.16) and (5.17). We do not yet combine the kinematic data at the likelihood level. We sample all the lens model parameters and the time delay distance $D_{\Delta t}^{\text{model}}$. We keep the lens light parameters fixed at the final position of the particle swarm process in the MCMC process to achieve a more efficient sampling of the

relevant parameters. We included the full flexibility of the lens light parameters on a subset of the MCMC chains and come to the conclusion that the additional covariance of the lens light model on the cosmographic analysis is very minor, i.e. the impact on the uncertainty on H_0 is below 0.1%.

From Bayes theorem, the likelihood of the parameters given the data is (modulo a normalization):

$$P(\boldsymbol{\eta}, D_{\Delta t}^{\text{model}} | \boldsymbol{d}_{\text{ACS}}, \boldsymbol{\Delta t}) \propto P(\boldsymbol{d}_{\text{ACS}}, \boldsymbol{\Delta t} | \boldsymbol{\eta}, D_{\Delta t}^{\text{model}}) P(\boldsymbol{\eta}) P(D_{\Delta t}^{\text{model}}). \tag{5.22}$$

We apply flat priors on the parameters $\gamma' \in [1, 2.8]$, $\theta_E \in [0.1", 10"]$, $q \in [0.5, 1]$, $\theta_{E,\text{clump}} \in [0", 1"]$, $\gamma_{\text{ext}} \in [0, 0.3]$ and $D_{\Delta t}^{\text{model}} \in [0, 10'000]$ Mpc.

At this stage, we want to emphasize that there are 3 data points in the time delay measurement compared to several thousands of high signal-to-noise pixels in the imaging comparison. In principle, the provided time delay measurement can not only determine $D_{\Delta t}^{\text{model}}$, which is independent of the imaging data but also can partially constrain the lens model. In practice, any even minor bias introduced in the image modeling can out-weigh the constraining power of the two additional time delay measurements.

In the following, we present the results of the analysis of filter F814W. The results of the equivalent analysis of filter F555W can be found in Appendix 5.A.3. To sample the posterior distribution of the parameter space we use `CosmoHammer` (202). We fix the shapelet scale β at [0.14", 0.15", 0.16", 0.17", 0.18", 0.19"] and do a separate inference of the parameters for each choice of β. Figure 5.5 shows the posterior distribution of some of the parameters for the different choices of β. The inferred parameter constraints for different β values do not overlap. We see that γ_{ext} is very narrowly determined for a given shapelet scale β but varies from 0.07 up to 0.11 depending on the position in the degeneracy plane. We want to stress that the external convergence κ_{ext} estimated by (180) is based on an external shear prior of $\gamma_{\text{ext}} = 0.089 \pm 0.006$.
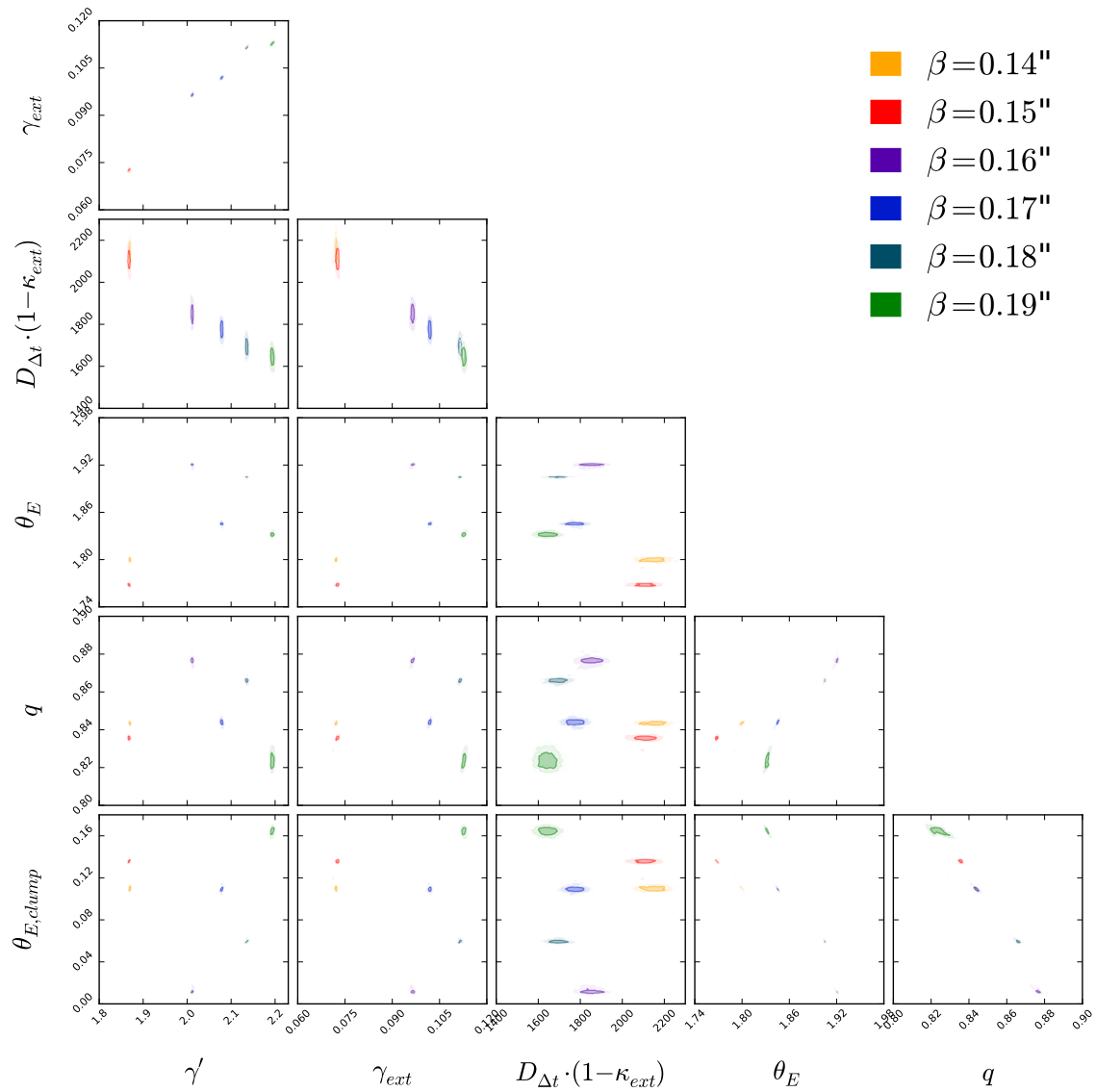
Figure 5.5: Posterior distribution (1-2-3 sigma contours) of lens model parameters and time delay distance of the combined analysis of imaging data of F814W and time delay measurements. Different colors correspond to different choices of the shapelet scale β. The posterior samples for different β values mutually disagree in almost all parameters presented.

5.4.2 Constraints from kinematic data

To investigate the potential constraining power of the velocity dispersion data, we are interested in how distinguishable different positions within the MST are in terms of their predicted central velocity dispersions. To do so, we fix the cosmology and the external convergence κ_{ext} to fiducial values. This allows us to evaluate the predicted LOS central velocity dispersion σ^{P} (Equation 5.11) for all the posterior samples of Figure 5.5. We assume a random realization of r_{ani} with a flat prior in the range $[0.5, 5]\theta_{\text{eff}}$ for all the posterior positions.

In Figure 5.6 we illustrate the predicted σ^{P} samples vs the predicted time delay distance $D_{\Delta t}$. We see that the samples can not be fully distinguished with the current velocity dispersion measurement and the assumed anisotropy prior. The relative distance in the predicted velocity dispersion σ^{P} between the different samples are all within 4σ (model given data).

There are three factors which affects the distinction of the source scales by kinematic data. (1) The uncertainty in the spectroscopic measurement, analysis and modeling of ± 20km s^{-1} which is about 6%. This is visually the most obvious contribution in Figure 5.6, marked by the gray band. The mean values of the predicted samples of the different source scales differ by about one sigma of this estimated uncertainty. (2) The anisotropic uncertainty in the lens galaxy kinematics. This is the main driver of the spread in the predictions of the velocity dispersion within each source scale sample. This scatter has a relative spread of 10% given $P_{[0.5,5]}(r_{\text{ani}})$. (3) The predicted velocity dispersion depend highly on the observational conditions and configuration. The PSF and the slit size of the spectrograph results in a convolution and averaging over a wide range of radial scales. The predicted velocity dispersion for different concentrations of the mass in the lens galaxy (i.e. power-law slope γ') differ the most in the very center of the lens. At the Einstein radius itself, the different lens models predict basically the same kinematics. With the PSF of 0.7" and a slit size of 0.81" × 0.7", power-law mass profiles with slopes in the range $\gamma' \in [1.8, 2.2]$ differ by about 100 km s^{-1} in their predicted velocity dispersion σ^{P}. A smaller slit and seeing conditions of FWHM 0.1" can double this relative difference and therefore could improve the constraining power of the kinematic data significantly.

The combined effect of non-perfect data and non-perfect modeling of the kinematic data with
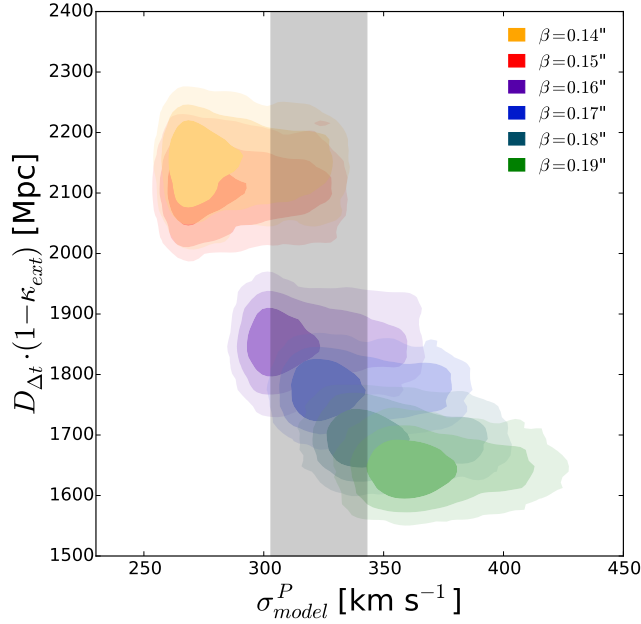
Figure 5.6: Estimated LOS central velocity dispersions σ^{P} vs. time delay distances $D_{\Delta t}$ of the sample of lens models from Figure 5.5 (in the same colors) for a kinematic anisotropy prior of $P_{[0.5,5]}(r_{\mathrm{ani}})$. The 1-2-3 sigma contours are shown. The external convergence κ_{ext} was explicitly set to zero and the cosmology has been fixed to the Planck mean values in this particular plot. The gray band reflects the 1-σ uncertainty range of the LOS velocity dispersion estimates from the data. This shows that velocity dispersion estimates add important information on the lens model constraints.

prior $P_{[0.5,5]}(r_{\mathrm{ani}})$ can be translated in a relative error in the time delay distance $D_{\Delta t}$ of about 7.5% from Figure 5.6. Only kinematic data of the lens galaxy and its analysis can reduce this error budget.

In Section 5.3.2 we showed that the individual image likelihoods of the different β samples differ by more than 10σ. Before including the velocity dispersion measurement in our cosmographic analysis, we re-normalize the image likelihood such that it is independent of the source scale β (see Section 5.3.3). This re-normalization is done by taking the same number of MCMC posterior samples from the different source scales β when doing further inferences with the lens model parameters.

5.4.3 Source scale and kinematic anisotropy priors

The combination and inference coming from the different data sets relies on priors on the source scale of the background galaxy and on the anisotropic behaviour of the stellar kinematics in the lens galaxy. In particular, the inference of the Hubble constant H_0 is related to the inference of the angular diameter distance $D_{\Delta t}$ as

$$H_0 \propto D_{\Delta t}^{-1}. \tag{5.23}$$

In Figure 5.6, we see a significant dependence between the size of the source galaxy ($\propto \beta$) and $D_{\Delta t}$. Furthermore the interpretation of the kinematic data is also dependent on the anisotropic behaviour of the lens galaxy.

Choices of the priors on the source size $P(\beta)$ and aniosotropic kinematic $P(\beta_{\mathrm{ani}}(r))$ must be chosen with care based on information gained from other work as these priors potentially have a significant impact on the infered parameter posterior (i.e. H_0). In the following, we discuss two different priors in the kinematic anisotropy and the source scale.

Source size prior $P(\beta)$

A simple form of the source size prior which does not impose any specific form of knowledge about β is a uniform prior in the range $[0", 10"]$. We refer to this prior as $P_{\mathrm{flat}}(\beta)$. This prior ignores any knowledge about the population of galaxies. The model parameter β is directly related to the brightness L of the source as

$$\beta^2 \propto L. \tag{5.24}$$

The number density of galaxies as a function of luminosity is a well measured quantity (luminosity function, LF) and its faint end slope for the blue galaxy population can be well described with a single power-law slope as

$$\frac{dn}{dL} \propto L^{\alpha_{\mathrm{LF}}} \tag{5.25}$$

with $\alpha_{LF} = -1.30$ (292). In this form, the expected source size can be stated as

$$P_{LF}(\beta) = \frac{dn}{d\beta} = \frac{dn}{dL}\frac{dL}{d\beta} \propto \beta^{2\alpha_{LF}+1}. \tag{5.26}$$

This prior is weakly dependent on β such that smaller source sizes are prefered. We chose $P_{flat}(\beta)$ as our default prior and explore the impact with $P_{LF}(\beta)$ in section 5.5.4.

Anisotropic kinematic prior $P(\beta_{ani}(r))$

Studies of early type (lens) galaxies have been made by e.g. (293; 294) which reveal similar properties compared to local early type galaxies. We consider two priors which cover the same range in the mean anisotropic behaviour and their predicted velocity dispersion σ^P. (1) The prior used in Figure 5.6 is flat in r_{ani} (equation 5.7) in the range $[0.5,5]\theta_{eff}$. This prior should cover the expected scale where the transition between isotropic and radial velocity dispersion should occur in an uniform way and is exactly the same prior used in (180). We refer to this prior as $P_{[0.5,5]}(r_{ani})$.

(2) We model a global contribution of the anisotropic behaviour in the form

$$\beta_{ani} = 1 - \frac{\bar{\sigma}_z^2}{\bar{\sigma}_r^2} \equiv 1 - \frac{1}{b} \tag{5.27}$$

in the range $[1, 1.5]$. This reflects the same range in allowed σ^P values for a given mass model. We refer to this prior as $P_{[1,1.5]}(b)$. $b = 1$ indicates a isotropic velocity dispersion and $b = 1.5$, for which the velocity dispersion ellipsoid is very elongated along the radial direction with $\beta_{ani} = 0.33$, corresponds to $r_{ani} = 0.5\theta_{eff}$ with the same mean anisotropy within the aperture. This is the same functional form of the prior as used in (295) to analyze a spiral lens galaxy althought with less range into a pure radial dispersion.

5.5 Cosmological inference

In this section, we study the cosmological constraints from strong lensing using data from images, time delays, central velocity dispersion of the lensing galaxy and independent external convergence estimates. We first show that the data can be used to constrain the angular diam-

eter relation. Based on the constraints on the angular diameter distances, we then introduce the likelihood that allows us to infer the parameters within the flat ΛCDM cosmological model.

5.5.1 Angular diameter distance posteriors

We can combine the posterior samples of Figure 5.5 with the independent velocity dispersion measurement to calculate the angular diameter distance relations D_d and D_s/D_{ds} (Equation 5.11 and 1.77) as

$$\frac{D_s}{D_{ds}} = \frac{(\sigma^P)^2}{(1 - \kappa_{ext})} \frac{1}{H(\gamma', \theta_E, \beta_{ani}(r), \theta_{eff})} \tag{5.28}$$

and

$$D_d = \frac{D_{\Delta t}^{model}}{(1 + z_d)(1 - \kappa_{ext})} \frac{D_{ds}}{D_s}. \tag{5.29}$$

To take into account the errors in σ_ν, κ_{ext} and r_{ani}, we importance sample the posteriors from the independent measurements (σ_ν and κ_{ext}) and for r_{ani} we uniformly sample in the range [0.5,5] times θ_{eff} (see e.g. 296; 179; 180, for similar use).

The D_d vs D_s/D_{ds} plane as shown in Figure 5.7 inherits the cosmological information of this analysis coming from the combined data and consistently translates the uniform prior in the source scale into the cosmological inference. This plane covers a wide range but the constrained region is more narrow. (297) did a very similar analysis in term of folding in the velocity dispersion measurement. In our case, we get a degeneracy in the two-dimensional plane coming from the MST whereas (297) and the forecasting of (298) assume independence in the two quantities. We over-plot the posterior samples of WMAP DR9 (299) and Planck15 (1) converted to the angular diameter distances of the lens system. We find that at least the posterior samples of one chosen source scale parameter β is consistent within 2σ with the CMB experiment posteriors in a flat ΛCDM cosmology for the low redshift angular diameter distance relations. Without the renormalization of the imaging likelihood (see Section 5.3.3), this statement can not be made.
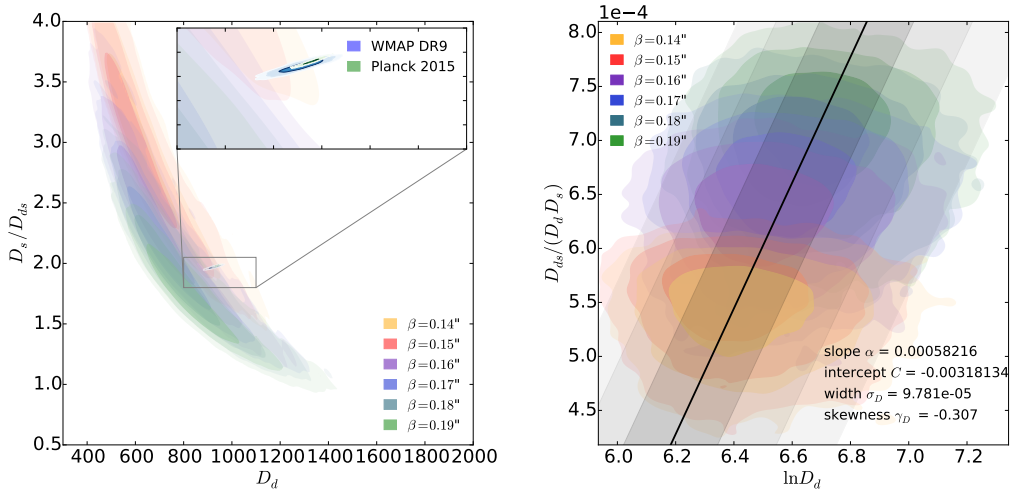
Figure 5.7: The constraints of the angular diameter distance relation for discrete positions in the MSD plane for filter F814W (same analysis for filter F555W is shown in Figure 5.11 in the appendix). The chosen priors in the source scale and the kinematic anisotropy of the lensing galaxy are $P_{\text{flat}}(\beta)$ and $P_{[0.5,5]}(r_{\text{ani}})$. Different colors indicate different imposed source scales. On the left panel: D_d vs D_s/D_{ds}. Also over-plotted are the posteriors of the WMAP DR9 and Planck 2015 ΛCDM posteriors mapped in the same angular diameter distance relation. On the right panel: Re-mapping of the angular diameter relations into a $\ln D_{\text{d}}$ vs $D_{\text{ds}}/(D_{\text{d}}D_{\text{s}})$ plane. The linear fit is indicated by the thick black line and the (1,2,3)-σ upper and lower limits of the projected distribution are plotted in different gray scale. The parameters of the fit are indicated in the figure.

5.5.2   An analytic likelihood for cosmology

So far, we have discretized the degeneracy plane by uniformly sample β in steps of 0.01". Effectively this means that while all the other parameters are sampled through standard MCMC methods, the β direction is sampled on a grid. This separation is needed to allow us to do the re-normalization of the likelihood as described in Section 5.4.2. Sampling the β-grid finely is computationally expensive. In the following, we show how we can analytically describe the posterior distribution and fill the gaps in β without additional sampling.

To do so, we first map the D_{d} vs $D_{\mathrm{s}}/D_{\mathrm{ds}}$ plane of Figure 5.7 (left panel) into a $\ln D_{\mathrm{d}}$ vs $D_{\mathrm{ds}}/(D_{\mathrm{d}}D_{\mathrm{s}})$ plane (right panel). We see a linear relations between the posterior samples in a monotonic and equally spaced increasing fashion as a function of β. We fit with linear regression the function

$$\frac{D_{\mathrm{ds}}}{D_{\mathrm{d}}D_{\mathrm{s}}} = \alpha \ln(D_{\mathrm{d}}) + C \tag{5.30}$$

with α being the slope and C being the intercept. The legend of Figure 5.7 (right panel) shows the best fit values, which we discuss in more detail later. The linear fit is a good description of the combined samples of different source scalings. The same is shown for the filter F555W analysis in Appendix 5.A.3. The spread of the distribution orthogonal to the linear relation is not well fit by a Gaussian distribution, but we find a skewed normal distribution provides a good description.

The one-dimensional likelihood $P(\boldsymbol{d}_{\mathrm{RXJ}}, \boldsymbol{\pi})$ of the strong lens system data $\boldsymbol{d}_{\mathrm{RXJ}}$ given a cosmological model $\boldsymbol{\pi}$ is given by the one-dimensional probability density of the samples relative to the fitted line:

$$P(\boldsymbol{d}_{\mathrm{RXJ}}, \boldsymbol{\pi}) = \phi_\gamma \left(x = \frac{D_{\mathrm{ds}}}{D_{\mathrm{d}}D_{\mathrm{s}}}, \mu = \alpha \ln(D_{\mathrm{d}}) + C, \sigma_D, \gamma_D \right), \tag{5.31}$$

where σ_D is the standard deviation, γ_D the skewness and ϕ_γ is the re-parameterized skewed normal distribution function described in Appendix 5.A.5. How the different source scale priors on β fold in the likelihood is described in Appendix 5.A.6 and equation 5.52. In this section, we apply a flat prior in β, $P_{\mathrm{flat}}(\beta)$, and a flat prior in r_{ani}, $P_{[0.5,5]}(r_{\mathrm{ani}})$, (see section

5.4.3). The inferences for the other combinations of the choices of priors are presented in Section 5.5.4.

For the analysis of the HST band F814W we fit the values $C = -3.18 \cdot 10^{-3}$, $\alpha = 5.82 \cdot 10^{-4}$, $\sigma_D = 9.78 \cdot 10^{-5}$ and $\gamma_D = -0.307$. For band F555W the fits result in $C = -4.70 \cdot 10^{-3}$, $\alpha = 8.20 \cdot 10^{-4}$, $\sigma_D = 1.32 \cdot 10^{-4}$ and $\gamma_D = -0.333$. Fitting the combined samples of the band F814W and F555W leads to $C = -3.43 \cdot 10^{-3}$, $\alpha = 6.22 \cdot 10^{-4}$, $\sigma_D = 1.04 \cdot 10^{-4}$ and $\gamma_D = -0.307$. The units of these parameters are given in respect with the angular diameter distances in Mpc.

The simple form of the likelihood enables a fast and consistent combination of different strong lensing systems also in combination with other cosmological probes.

5.5.3 Cosmological parameter constraints

The constraints on the angular diameter distance relations can be turned into constraints on the cosmological parameters of the background evolution. In the following we assume a flat ΛCDM cosmology. The homogeneous expansion can be described in terms of the matter density Ω_m and the Hubble constant H_0. We use the likelihood of Equation (5.31) with the values of α, C, σ_D and γ_D from the analysis of F814W and F555W separately. First, we sample the parameters Ω_m and H_0 simultaneously with uniform priors of $\Omega_m \in [0,1]$ and $H_0 \in [0,200]$. Figure 5.8 shows the posterior distributions for the filter F814W (left panel) and F555W (middle panel) for the priors $(P_{\text{flat}}(\beta), P_{[0.5,5]}(r_{\text{ani}}))$ separately. The degeneracy in Ω_m is strong but H_0 can be determined fairly well. A good approximation of the degeneracy shown in the H_0-Ω_m-plane can be described by

$$H_0 = H_0^* \left[1 + \frac{1}{2}(\Omega_m - \Omega_m^*) \right]^{-1} \pm \sigma_{H_0^*} \left(\frac{H_0}{H_0^*} \right) \tag{5.32}$$

where H_0^* is the value for H_0 at fixed Ω_m^* and $\sigma_{H_0^*}$ is the marginalized error at fixed Ω_m^*. This form allows us to more directly compare with other results from the literature.

For a fixed value of $\Omega_m = 0.3$, we infer a Hubble constant of $H_0 = 85.0^{+6.0}_{-6.3}$ km s^{-1}Mpc^{-1} for the F814W and $H_0 = 88.5^{+7.5}_{-7.4}$ km s^{-1}Mpc^{-1} for the F555W analysis.
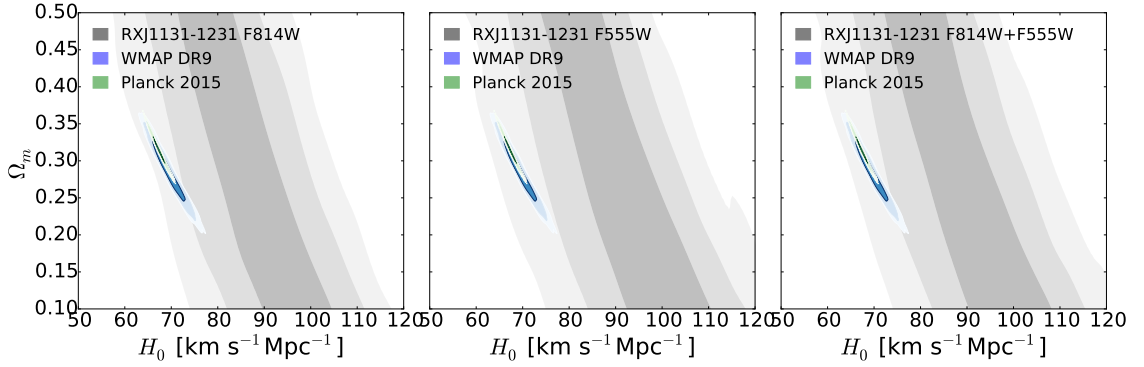
Figure 5.8: Posterior sampling of the cosmological parameters for the filters F814W (left), F555W (middle) and combined with equal weight of the likelihoods of the two images (right). The posterior distribution of WMAP DR9 and Planck 2015 are over-plotted. The chosen priors in the source scale and the kinematic anisotropy of the lensing galaxy are $P_{\text{flat}}(\beta)$ and $P_{[0.5,5]}(r_{\text{ani}})$.

From the analysis of each filter separately, we get an uncertainty coming from the imaging data only to be below 1% in the resulting inference of H_0. Given the fact that our estimates for the two filters F814W and F555W is about 4.0% different while using exactly the same analysis and the same time-delay and kinematic data for all other parameters involved, we conclude that the imaging data inference is partially driven by unknown systematics in the modeling and the data. To marginalize out potential systematics in the analysis, we combine the two analyses on the angular diameter posterior level. The two-dimensional posteriors are shown in the right panel of Figure 5.8. In this way, we get a Hubble constant of $H_0 = 86.6^{+6.8}_{-6.9}$ km s^{-1}Mpc^{-1}. The full posteriors for both samples are shown in Figure 5.9.

5.5.4 Prior dependence

In this section we investigate the dependence of the cosmological inference from the choice of priors of the source scale β and the anisotropic kinematics of the lensing galaxy β_{ani}. In Section 5.4.3 we stated for each parameter two different priors, each of them being quoted to be uninformative and probing the same range in the physics. In table 5.1 the likelihood parameters and the resulting H_0 inference for fixed $\Omega_m = 0.3$ in flat ΛCDM are stated. We see a strong prior dependence on the posterior distribution which can result in a mean shift in H_0 of more than 10 km s^{-1}Mpc^{-1}. The source scale prior $P(\beta)$ can result in a weak mean shift of about 1-2 km s^{-1}Mpc^{-1}without a change in the uncertainty. This means that
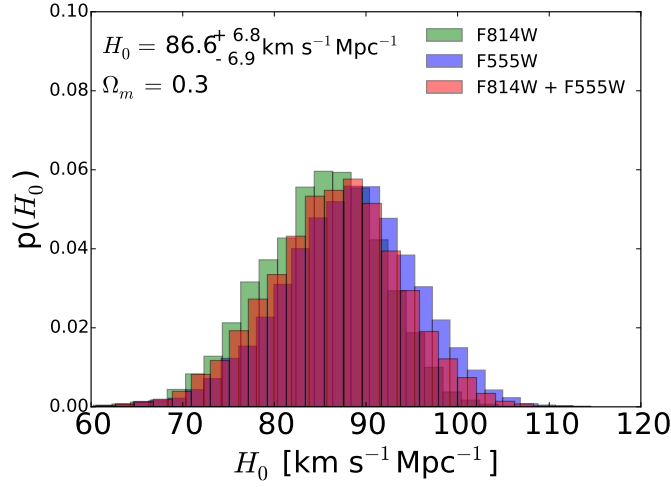
Figure 5.9: Posterior distribution for the value of H_0 for a fixed $\Omega_m = 0.3$ for filter F814W (green), F555W (blue) and the combined samples (red). The chosen priors in the source scale and the kinematic anisotropy of the lensing galaxy are $P_{\text{flat}}(\beta)$ and $P_{[0.5,5]}(r_{\text{ani}})$.

$P(\beta_{\text{ani}})$	$P(\beta)$	C	α	σ_D	γ_D	C_β	α_β	H_0[a]
$P_{[0.5,5]}(r_{\text{ani}})$	$P_{\text{flat}}(\beta)$	$-3.43 \cdot 10^{-3}$	$6.22 \cdot 10^{-4}$	$1.04 \cdot 10^{-4}$	-0.307	-	-	$86.6^{+6.8}_{-6.9}$
$P_{[0.5,5]}(r_{\text{ani}})$	$P_{\text{LF}}(\beta)$	$-3.43 \cdot 10^{-3}$	$6.22 \cdot 10^{-4}$	$1.04 \cdot 10^{-4}$	-0.307	-0.0012	263.8	$84.3^{+6.7}_{-7.0}$
$P_{[1,1.5]}(b)$	$P_{\text{flat}}(\beta)$	$-3.64 \cdot 10^{-3}$	$6.30 \cdot 10^{-4}$	$1.41 \cdot 10^{-4}$	-0.089	-	-	$75.7^{+8.3}_{-7.8}$
$P_{[1,1.5]}(b)$	$P_{\text{LF}}(\beta)$	$-3.64 \cdot 10^{-3}$	$6.30 \cdot 10^{-4}$	$1.41 \cdot 10^{-4}$	-0.089	-0.0014	264.2	$74.5^{+8.0}_{-7.8}$

[a] For fixed $\Omega_m = 0.3$ in flat ΛCDM.

Table 5.1: Likelihood and posteriors for different choices of priors. The H_0 inference is for fixed $\Omega_m = 0.3$. $P_{[0.5,5]}(r_{\text{ani}})$ indicates a flat prior in r_{ani} in the range $[0.5, 5]\theta_{\text{eff}}$ in the parameterization of equation 5.7 and $P_{[1,1.5]}(b)$ indicates a flat prior in b of equation 5.27 of the anisotropic behavior of the lens galaxy. $P_{\text{flat}}(\beta)$ reflects a flat prior in the source scale and $P_{\text{LF}}(\beta)$ reflects a prior of the galaxy luminosity function (see section 5.4.3). The parameters describe the likelihood function stated in equation 5.31 and 5.52.

the information content in the imprinted priors are roughly the same and the systematic uncertainty is subdominant to the quoted total uncertainty. The situation changes for the kinematic prior $P(\beta_{\text{ani}})$. The flat prior approach for the two different parameterizations shifts the mean infered value of H_0 by more than 1σ. The precision is also affected: The prior $P_{[0.5,5]}(r_{\text{ani}})$ results in a significantly higher precision inference than $P_{[1,1.5]}(b)$. This implies that $P_{[0.5,5]}(r_{\text{ani}})$ inherits more information for the specific task of measuring H_0 than $P_{[1,1.5]}(b)$. If this prior is not representative of the distribution of early type galaxies, the inference with $P_{[0.5,5]}(r_{\text{ani}})$ can be significantly biased compared with $P_{[1,1.5]}(b)$.

5.6 Joint uncertainties and comparison with other work

In this Section we analyze the impact of the different data sets on the cosmological inference and we compare our method and results with the literature.

5.6.1 Uncertainties from the different data sets

We assign uncertainty estimates on the inference of H_0 coming from the independent data sets, namely the time delays, the HST ACS images, the line-of-sight analysis of wide field data and the spectra of the lens galaxy for the kinematic estimate [2]. We do so by forecasting a perfect modeling result for all data sets except the one in question. We then proceed in exactly the same way as presented in Section 5.5. This leads to an inference of the cosmological parameters only affected by the uncertainties coming from one single data set. We perform this analysis with the default priors $P_{[0.5,5]}(r_{\text{ani}})$ and $P_{\text{flat}}(\beta)$.

In Table 5.2 the estimated uncertainties from the different data sets are summarized and the 1-σ uncertainties on H_0 for fixed Ω_m is stated. The Gaussian approximation of all these errors leads to a total uncertainty of 9.4% on H_0. The estimate of the uncertainty coming from the full sampling results in 7.9%. This analysis does not include further potential systematics and does not question the priors chosen.

Our approach on the error analysis is different than the one chosen by (180). We do not quote an error on the lens model itself, as this inference is dependent on different data sets. We quote an error on the lens model modulo a MST for the image reconstruction and separately an error on the kinematic estimate, which potentially can fully break the degeneracy.

We clearly see that the dominant contribution in the final uncertainty can be related to the kinematic data and its modeling. As discussed in Section 5.4.2, high resolution spectroscopy can provide data which can better constrain different positions in the MST and therefore significantly reduce the uncertainty on the angular diameter distance relation. The second most dominant uncertainty come from the line-of-sight contribution.

---

[2] In this analysis we ignore the dependence of the line-of-sight analysis on the shear term from the ACS image reconstruction.

5.6.2 Comparison with other work

Cosmographic inference has been published by (180) with the same lens model parameterization and by (282) in combination with a composite (dark matter and baryonic matter separated) lens model, in response to the work of (286). The values and uncertainties on the Hubble constant are $H_0 = 78.7^{+4.3}_{-4.5}$ km s^{-1}Mpc^{-1}for a value of $\Omega_m = 0.27$ in (180), a 5.5% error, and $H_0 = 80.0^{+4.5}_{-4.7}$ km s^{-1}Mpc^{-1}, a 5.75% error, with $\Omega_m = 0.27$ for a flat ΛCDM universe.

One difference between the work of (180; 282) and the one presented in this chapter arise from the explicit treatment of the MSD and related degeneracies in our work and its link to the source surface brightness reconstruction method. This allows us to overcome (at least partially) systematics from the source reconstruction method and the mass profile assumption. On the other hand, this weakens the constraining power of the image reconstruction. This explains our larger uncertainties compared to (180; 282). Furthermore, their stated values on H_0 are Ω_m-independent in the flat scenario while our values do depend on Ω_m (see our Figure 5.8 vs. Figure 8 in (180)). This comes from the different description of the cosmological likelihood. The likelihood in (180) is described fully in terms of the time-delay distance $D_{\Delta t}$ where else our likelihood has an additional dependence on D_d. In that sense, their stated H_0 value is independent of Ω_m but ours requires a prior on Ω_m.

A second difference is that we work in a 2D-plane of angular diameter distance relations (Figure 5.7) without the need of cosmological priors to define our angular diameter distance likelihood. This results in a different shape of the posterior distribution in the Ω_m-H_0 plane (Figure 5.8) and the inferred projected H_0 posteriors have a strong Ω_m dependence.

The best comparison with the work of (180; 282) should be done when comparing the inference with the same kinematic prior $P_{[0.5,5]}(r_{\text{ani}})$ (first or second row in Table 5.1). We want to stress that we use explicit priors on the source scale. The cosmological inference is dependent on this prior as the constraining power of the kinematic data is weak. Therefore a shift of about 1σ in our stated uncertainty on the inference of H_0 is not surprising.

Comparing our results with the CMB experiments, we get a 2.5 σ shift for $P_{[0.5,5]}(r_{\text{ani}})$ and a 1σ shift for $P_{[1,1.5]}(b)$ in the ΛCDM parameter inference. We conclude that the angular

Description	Uncertainty
Time delays	1.6%
HST ACS image reconstruction	2.8%
Line-of-sight contribution	4.7%
Lens kinematics [a]	7.5%
Total (Gaussian)	9.4%
Total (full sampling [b])	7.9%

[a] The quoted uncertainty includes the uncertainty in the unisotropy radius r_{ani} with a prior of $[0.5, 5]\theta_{eff}$.
[b] The uncertainty in the full sampling is given as half of the 68% confidence interval divided by the mean posterior value.

Table 5.2: Error budget on H_0 for a fixed Ω_m.

diameter distance at last scattering and the inferred angular diameter distance relation at lower redshift from this analysis are consistent with a flat ΛCDM cosmology. Our analysis depends on uninformative priors on the kinematics of the lens galaxy β_{ani} and the source reconstruction scale β. Further systematics can potentially also occur and are not included in this analysis.

5.7 Summary

In this chapter we applied the newly developed source reconstruction technique of (29) to the strong lens system RXJ1131-1231 to extract cosmographic information. We showed how different source reconstruction scales probe different regimes in the MST even when the lens model is not fully transformable through the MST.

This chapter is built on the modeling and the data of (180) and the systematics analysis of (286). We incorporate a re-normalization of the imaging likelihood such that we have explicit priors on the source scale before combining with the kinematic data.

We introduced a cosmographic inference analysis which enables us to combine imaging, time-delay and kinematic data without relying on any cosmological priors. We came up with a likelihood function only based on the angular diameter distance relations, which can be described in analytic terms.

We find that the choice of priors on lens model parameters and source size are subdominant

for the statistical errors for H_0 measurements of this systems. The choice of prior for the source is sub-dominant at present (2% uncertainty on H_0) but may be relevant for future studies. More importantly, we find that the priors on the kinematic anisotropy of the lens galaxy have a significant impact on our cosmological inference. When incorporating all the above modeling uncertainties, we find $H_0 = 86.6^{+6.8}_{-6.9}$ km s^{-1}Mpc^{-1}(for $\Omega_m = 0.3$), when using kinematic priors similar to other studies. When we use a different kinematic prior motivated by Barnabè et al. (2012) (295) but covering the same anisotropic range, we find $H_0 = 74.5^{+8.0}_{-7.8}$ km s^{-1}Mpc^{-1}. This means that the choice of kinematic modeling and priors have a significant impact on cosmographic inferences. Further systematics in the data and modeling can also occur. The way forward is either to get better velocity dispersion measures which would down weight the impact of the priors or to construct physically motivated priors for the velocity dispersion model.

This inference analysis was achieved with a single strong lens system in two imaging bands. Combining the information of multiple systems with comparable data can add vital constraints about the late time expansion history of the universe, also in terms of extensions of the standard cosmological model.

5.A   Appendix of chapter 5

5.A.1   Numerical computation of the luminosity-weighted LOS velocity dispersion

The computation of the luminosity-weighted LOS velocity dispersion within an aperture under certain seeing conditions σ^{P} (Equation 5.10) involves numerically challenging projection integrals and convolutions. In this section, we describe our approach to achieve a numerically stable and fast computation with a Monte-Carlo ray-tracing approach, similarly used by e.g. (300) to render convolved Galaxy light profiles. This method is based on drawing positions representing the total light distribution of the galaxy.

For the light in the galaxy, we take a Hernquist profile (290)

$$I(r) = \frac{I_0 a}{2\pi r (r + a)^3} \tag{5.33}$$

where I_0 is the total flux and a relaxed to the effective radius of the galaxy by $a = 0.551\theta_{\text{eff}}$. The radial distribution function of flux is then

$$P(r)dr = \frac{2r}{(r+a)^3}dr. \tag{5.34}$$

The cumulative distribution function is

$$P(< r) = \int_0^r \frac{2r'}{(r'+a)^3}dr' = \frac{r^2}{(a+r^2)}. \tag{5.35}$$

A sample of $P(r)$ can then be drawn from the distribution

$$P(r) = \frac{a\sqrt{\mathcal{U}}\left(\sqrt{\mathcal{U}}+1\right)}{1-\mathcal{U}}, \tag{5.36}$$

where $\mathcal{U}$ is the uniform distribution in $[0,1]$.

In the following, we describe the steps starting from a representative sample of the flux in the galaxy to get to the estimate of the aperture averaged velocity dispersion:

1. Draw a representative sample of radii r_i drawn from the three-dimensional light distribution of the Hernquist profile (Equation 5.36).

2. Project the radius r_i on a random two-dimensional plane and compute its projected radius R_i and the projected coordinates (x_i, y_i). This sample represents the projected light profile of the galaxy.

3. Displace the two-dimensional coordinates (x_i, y_i) with a random realization according to the seeing distribution to (x_i', y_i'). We assume the PSF is a two-dimensional Gaussian distribution. This sample represents the convolved, projected two-dimensional light distribution of the galaxy.

4. Select samples, whose displaced position is on the aperture $(x_i', y_i') \in \mathcal{A}$. This selects a sample representative for the luminosity and radial weighting within the aperture.

5. Evaluate $\sigma_s^2(r_i, R_i)$, the projected (but unweighted) velocity dispersion for the remaining samples.
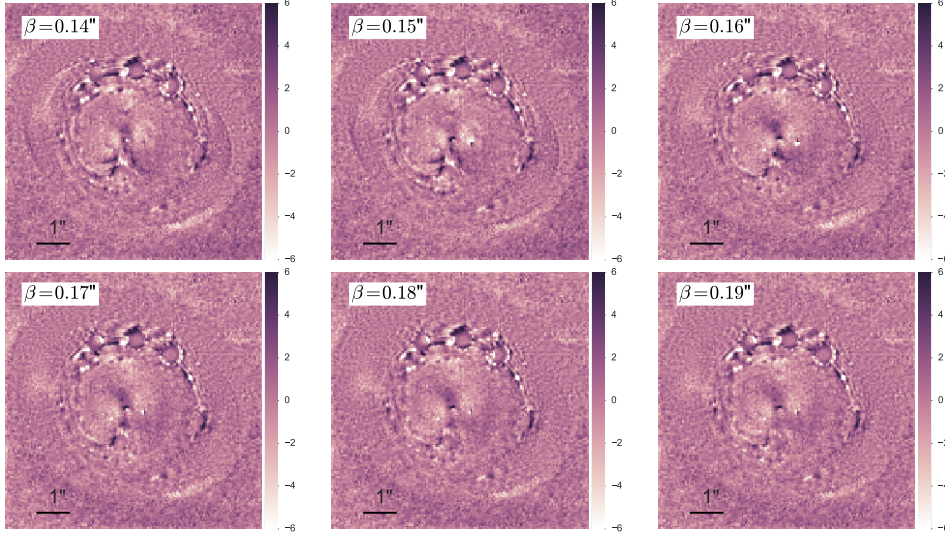
Figure 5.10: The normalized residual maps for the best fit reconstruction for the different choices of the shapelet scale β for the F814W image. The residuals differ significantly for the different choices of β. From the imaging data only, a scale $\beta = 0.19''$ is favored over a scale $\beta = 0.14''$ by more than $30\,\sigma$. This statement is entirely lens model dependent.

6. Take the sample average of the velocity dispersion $\langle \sigma_s^2(r_i, R_i) \rangle$. This average (once converged) corresponds to $(\sigma^{\mathrm{P}})^2$ with the assumption of a Gaussian velocity dispersion.

About 100 samples evaluated in the aperture gives already an accuracy in σ^{P} of about 1%. In this chapter, the computation is done with 1000 samples.

5.A.2   Residual maps

In Figure 5.10 the normalized residuals corresponding to the source models with different source scales β in Section 5.3.2 are shown. The residual maps differ significantly between the best fit values of the different shapelet scales β. This reflects the fact that extended structure in the Einstein ring can give constraints on the local slope of the mass profile and the given mass model can not adopt equally well to different source scales as it is can not be rescaled according to the mass-sheet transform. The inferred lens models can be understood as the best fit power-law profiles at different positions within the MST.
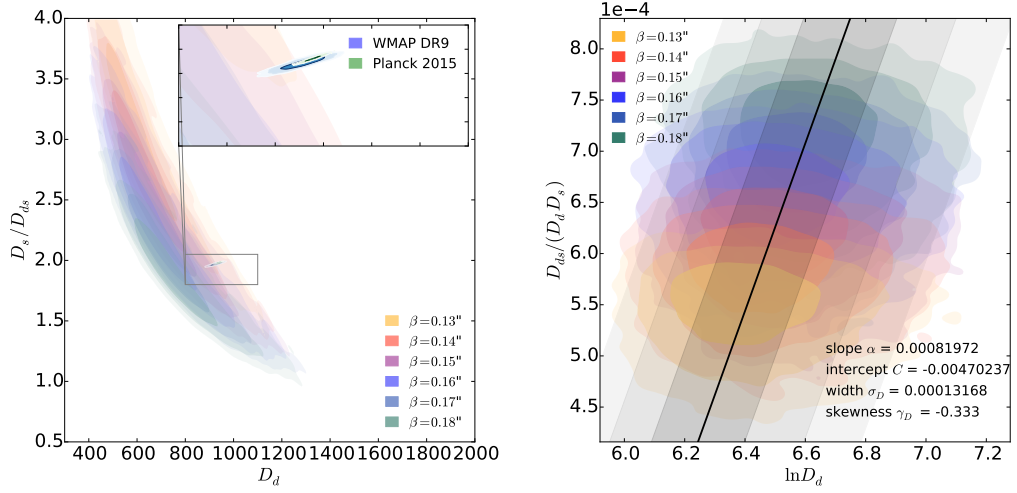
Figure 5.11: The constraints of the angular diameter distance relation for discrete positions in the MSD plane for filter F555W (same as Figure 5.7 for filter F814W). Different colors indicate different imposed source scales. On the left panel: D_d vs D_s/D_{ds}. Also over-plotted are the posteriors of the WMAP DR9 and Planck 2015 ΛCDM posteriors mapped in the same angular diameter distance relation. On the right panel: Re-mapping of the angular diameter relations into a $\ln D_d$ vs $D_{ds}/(D_d D_s)$ plane. The linear fit is indicated by the thick black line and the $(1,2,3)$-σ upper and lower limits of the projected distribution are plotted in gray scale. The parameters of the fit are indicated in the figure.

5.A.3 Analysis on WFC1 F555W

In the chapter, we did focus on the analysis of the WFC1 F814W filter band. Here we present the same analysis for filter F555W. Figure 5.12 shows the posterior distribution of the lens model parameters and time delay distance for F555W. Figure 5.11 shows the constraints on the angular diameter distance relation. The values describing the distribution can be found in the main text.

5.A.4 Bayesian description and renormalization of the imaging likelihood

One of the steps presented in this chapter is the renormalization of the imaging likelihood for different source scales β. In Section 5.3.3 we provided heuristic arguments for this approach in the case of time delay cosmography. In the following Section, we provide a Bayesian interpretation and justification of our choice in performing this calculation.

Let us assume that there is a complete model that is able to fully describe the lens, with
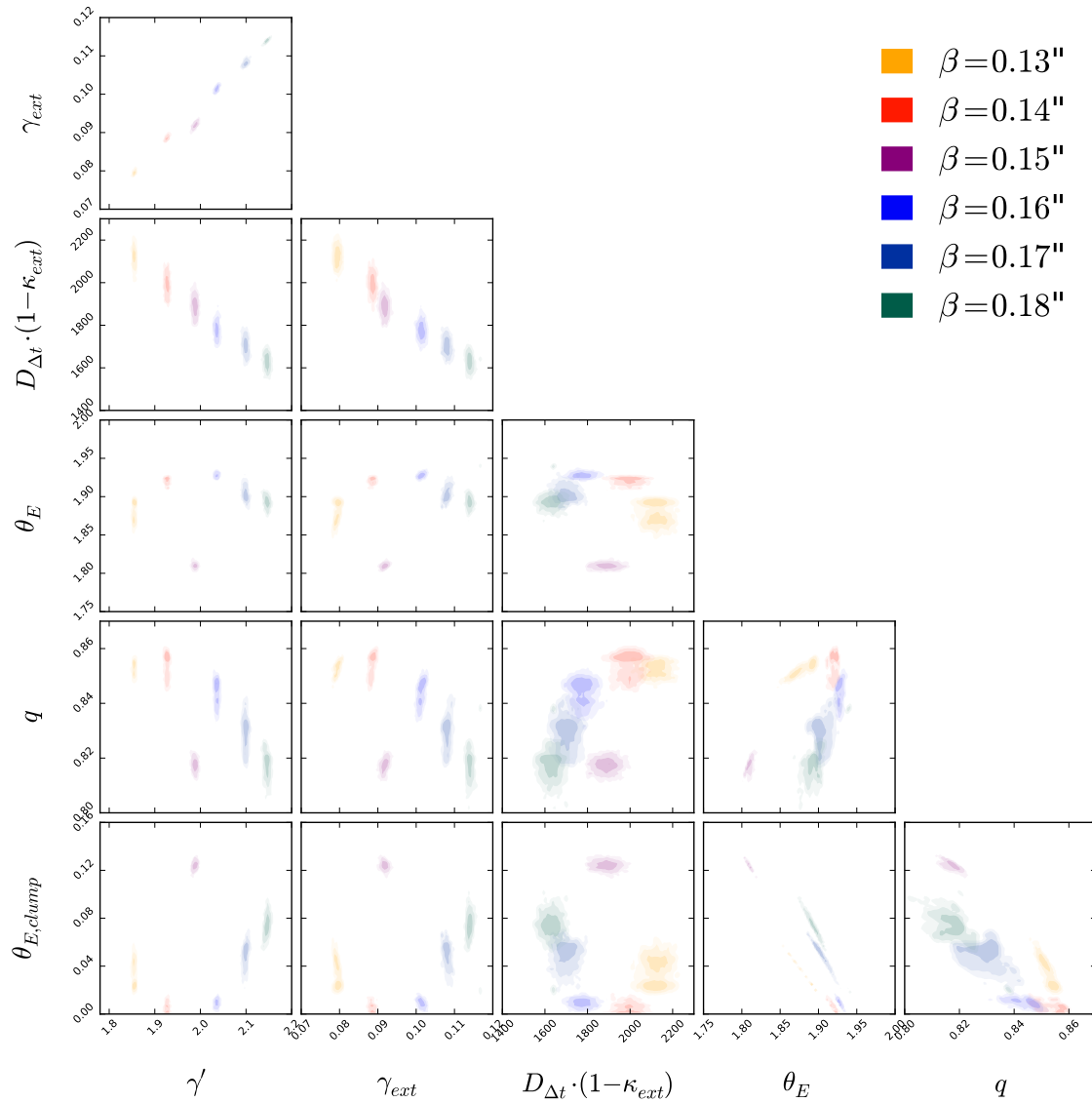
Figure 5.12: Posterior distribution (1-2-3 sigma contours) of lens model parameters and time delay distance of the combined analysis of imaging data of F555W and time delay measurements. Different colors correspond to different choices of the shapelet scale β. (same as Figure 5.5 for filter F814W).

parameters α. However, when we fit the data, in our modeling process, we use a restricted subset of the model containing only the parameters $\hat{\alpha}$ and that the missing degrees of freedom are captured by the parameters θ. To complete our notations, the source scale is given as β, the cosmological parameters as π. We also denote the image data as D_I, the kinematic data as D_σ and any other independent data of the time delays and the lens environment as D_o.

Our goal is to estimate the cosmological parameters π given the data, which is $P(\pi|D_I, D_\sigma, D_o)$. We can state, using Bayes rule

$$P(\pi|D_I, D_\sigma, D_o) = \int P(\pi|\hat{\alpha}, \theta, \beta, D_I, D_\sigma, D_o) P(\hat{\alpha}, \theta, \beta|D_I, D_\sigma, D_o) d\hat{\alpha} d\theta d\beta. \tag{5.37}$$

Independence of D_I, D_σ and D_o results in

$$P(\pi|D_I, D_\sigma, D_o) = \int P(\pi|\hat{\alpha}, \theta, \beta, D_I, D_\sigma, D_o) P(\hat{\alpha}, \theta, \beta|D_I) P(\hat{\alpha}, \theta, \beta|D_\sigma, D_o) d\hat{\alpha} d\theta d\beta. \tag{5.38}$$

The internal part of the MST is encapsulated in the term $P(\hat{\alpha}, \theta, \beta|D_I)$. One way to think about MST is that the source scale cannot be measured from imaging data alone. In other words, given image data and marginalizing over all possible lens models, one should recover the source size prior. The Bayesian expression for the MST is then

$$\int P(\hat{\alpha}, \theta, \beta|D_I) d\hat{\alpha} d\theta = P(\beta), \tag{5.39}$$

which can also be written as

$$P(\hat{\alpha}, \theta, \beta|D_I) = P(\hat{\alpha}, \theta|D_I, \beta) P(\beta|D_I) = P(\hat{\alpha}, \theta|D_I, \beta) P(\beta). \tag{5.40}$$

Incorporating this into the earlier expression we get

$$P(\pi|D_I, D_\sigma, D_o) = \int P(\pi|\hat{\alpha}, \theta, \beta, D_I, D_\sigma, D_o) P(\hat{\alpha}, \theta|D_I, \beta) P(\beta) P(\hat{\alpha}, \theta, \beta|D_\sigma, D_o) d\hat{\alpha} d\theta d\beta.$$
$$\tag{5.41}$$

This can be simplified further by considering the dependencies of the variables. For instance

$P(\pi | \hat{\alpha}, \theta, \beta, D_I, D_\sigma, D_o)$ simplifies to $P(\pi | \hat{\alpha}, \theta, D_\sigma, D_o)$, since all the information from D_I and β are captured by $\hat{\alpha}$ and θ. Further more the parameter β is not directly dependent on the velocity dispersion D_σ and related quantities D_o through the lens model. This relation of parameters and conditional data leads to

$$P(\pi | D_I, D_\sigma, D_o) = \int P(\pi | \hat{\alpha}, \theta, D_\sigma, D_o) P(\hat{\alpha}, \theta | D_I, \beta) P(\beta) P(\hat{\alpha}, \theta | D_\sigma, D_o) d\hat{\alpha} d\theta d\beta. \quad (5.42)$$

Until now, no approximations are made in the Bayesian analysis. The split of $\alpha \to (\hat{\alpha}, \theta)$ has been useful in working out the impact of the internal MST in our Bayesian analysis. However, to move the analysis further, we will have to make some simplifying assumptions about the further, i.e. beyond MST, impact of the unknown lens model parameters θ. (301) showed that time delays (and hence the cosmological inference) depends mostly on the slope of the density profile in the annulus over which the lens images are observed, which is part of $\hat{\alpha}$ in our model. From this we assume that $\hat{\alpha}$ is a good approximation of the overall lens model α and the relative deviation θ is small in terms of the impact on the cosmographic analysis (π). We approximate $\alpha \approx \hat{\alpha}$ at this stage, which leads to

$$P(\pi | D_I, D_\sigma, D_o) \approx \int P(\pi | \hat{\alpha}, D_o) P(\hat{\alpha} | D_I, \beta) P(\beta) P(\hat{\alpha} | D_\sigma, D_o) d\hat{\alpha} d\beta. \quad (5.43)$$

This equation is the formal expression of the steps that we perform in our analysis of combining imaging, time-delay, kinematic and environment data in our cosmographic analysis. The imaging data D_I folds in the analysis through the term $P(\hat{\alpha} | D_I, \beta)$. This term is conditional on the source scale β. This conditional likelihood is effectively computed by a renormalization of the imaging likelihood to a given source scale β.

5.A.5 Skewed normal distribution

The the skewed normal distribution is defined with a parameter α as

$$\phi_\gamma(x) = 2\phi(x)\Phi(\alpha x) \quad (5.44)$$

with $\phi(x)$ being the standard normal probability density function and $\Phi(x)$ its cumulative distribution. Location and scale can be added with

$$x \rightarrow \frac{x - \xi}{\omega}. \tag{5.45}$$

The mean μ of this distribution is given by

$$\mu = \xi + \omega \delta \sqrt{\frac{2}{\pi}} \tag{5.46}$$

where

$$\delta = \frac{\alpha}{\sqrt{1 + \alpha^2}}. \tag{5.47}$$

The variance σ^2 is

$$\sigma^2 = \omega^2 \left(1 - \frac{2\delta^2}{\pi}\right) \tag{5.48}$$

and the skewness γ as

$$\gamma = \frac{4 - \pi}{2} \frac{\left(\delta\sqrt{2/\pi}\right)^3}{\left(1 - 2\delta^2/\pi\right)^{3/2}}. \tag{5.49}$$

The skewed normal distribution $\phi_\gamma(x, \xi, \omega, \alpha)$ can be re-parameterized to $\phi_\gamma(x, \mu, \sigma, \gamma)$ by inverting the equations (5.46)-(5.49).

5.A.6 Source size prior

To account for an arbitrary prior in β in the Bayesian inference, one has to marginalize as

$$P(\boldsymbol{d}_{\mathrm{RXJ}}|\boldsymbol{\pi}) = \int P(\boldsymbol{d}_{\mathrm{RXJ}}|\boldsymbol{\pi}, \beta) P(\beta) d\beta = \int P(\boldsymbol{d}_{\mathrm{RXJ}}|\boldsymbol{\pi}) P(\boldsymbol{\pi}|\beta) P(\beta) d\beta. \tag{5.50}$$

β does not appear as a parameter in the likelihood of Equation 5.31. From Figure 5.7 one sees that the different source scale posteriors are equally spaced in the $D_{ds}/(D_d D_s)$ axis. The likelihood defined in Equation 5.31 is an approximation for a flat prior in the source scale β.

We approximate $P(\boldsymbol{\pi}|\beta)$ by a delta function in the parameter $D_{ds}/(D_d D_s)$ as

$$P(\boldsymbol{\pi}|\beta) \approx \delta\left(\alpha_\beta \frac{D_{ds}}{D_d D_s} + C_\beta - \beta\right) \tag{5.51}$$

where α_β is the slope of the $D_{ds}/(D_d D_s)$ vs β and C_β the intercept. In this form, the prior on β can be added to the likelihood of Equation 5.31 as

$$P(\boldsymbol{d}_{\mathrm{RXJ}}, \boldsymbol{\pi}) = \phi_\gamma\left(x = \frac{D_{\mathrm{ds}}}{D_{\mathrm{d}} D_{\mathrm{s}}}, \mu = \alpha \ln(D_{\mathrm{d}}) + C, \sigma_D, \gamma_D\right)\left(\alpha_\beta \frac{D_{ds}}{D_d D_s} + C_\beta\right)^{2\alpha_{\mathrm{LF}}+1}. \tag{5.52}$$

6 Strong lensing constraints on dark matter from substructure

This chapter will appear in a similar form in Birrer, Amara & Refregier in prep.

The physical properties of dark matter affects cosmological structures on small, sub-galactic scales. This has been used to set constraints on the mass of thermal relic warm dark matter (WDM) using the Lyman-α correlation function (see e.g. 20, for latest results) and the counts of dwarf satellites in the Milky Way (see e.g. 21; 22). Strong lensing is another probe of sub-galactic scales (151; 154; 155). For example, strong lensing was used to detect luminous and dark substructure in strong lens systems (157; 158). This has led to the detections of individual clumps down to masses of about $2 \times 10^8 M_\odot$ (159; 160; 302). To measure the properties of WDM one needs however to go beyond the clump-by-clump analysis and to statically probe the population of more numerous smaller clumps.

Statistical approaches to quantify substructure have been done in simple lens configurations, for instance by analyzing flux ratios in multiple imaged lensed quasars (see e.g., 161; 163; 164; 165). Deviating flux ratios have been reported relative to a prediction based on simple smooth lens models, known as flux-ratio anomalies. This approach requires multiple strong lens systems as the statistics is rather poor for a single lens system.

In this chapter, we take a different statistical approach to quantify substructure in strong gravitational lenses. It is based on an extended forward modeling scheme that relies on rendering of the subhalo population for different dark matter models, coupled to image

simulations. These generated mocks are tuned to be similar to the data (see e.g. 303; 304; 305; 306; 307; 308; 309; 310; 311; 312; 313; 314; 315, for other application of forward modeling in cosmology). A simple feature of a forward modeling approach is that the same analysis tools are run on both simulations and the data. This allows us to make statistically significant inference statements, even for cases where the full likelihood is not tractable, by understanding the statistical behaviors coming from the simulated mocks. A key step is to find diagnostic measures that are sensitive to the signal we wish to target, which is substructure in our case.

The chapter is structured as follows: In section 6.1, we describe the smooth lens model and source reconstruction of the lens RXJ1131-1231. In section 6.2, we describe the substructure modeling and how we create realistic mock lenses for different dark matter models. We then describe how we perform a substructure analysis and how we compare the statistical features in section 6.3. In section 6.4, we present our dark matter model constraints based on the analysis of the lens RXJ1131-1231 based on hundreds of different simulations. In section 6.5 we discuss the results, compare them to the literature and discuss possible extensions to this work. We summarize this chapter in section 6.6. Further technical details about the analysis are provided in the appendices of this chapter.

Throughout this chapter, we assume a flat ΛCDM cosmological model with $H_0 = 69.31 \text{kms}^{-1}\text{Mpc}^{-1}$, $\Omega_b = 0.049$, $\Omega_m = 0.315$, $\sigma_8 = 0.829$ and $n_s = 0.968$.

6.1 Smooth lens model

Before we measure the signatures of substructure, we first need to model the smooth global features of the lens. This is needed to separate global effect from small scale effects since the former does not depend on dark matter properties.

To fit the smooth lens model, we apply the method presented in chapter 3 on the data used in chapter 5. The main difference in the lens modeling is that, for this chapter, we reconstruct the two HST ACS filter bands F814W and F555W simultaneously and combine the likelihood of the lens model given the imaging data. For the global source structure, we use shapelets up to order $n_{\max} = 50$ with scale parameter $\beta = 0.18"$ centered at the position of the quasar. In chapter 5, we discussed and quantified the impact of the source scale on time-delay cosmography

through the internal part of the mass-sheet degeneracy (286). We do not expect a significant impact of this known degeneracy for the substructure analysis, as the lens profile re-scaling by the mass-sheet degeneracy only affects large scales.

In modeling the global feature that do not come from substructure, care needs to be taken in two respects. The first is that the smooth lens model needs to be sufficiently flexible to capture the scales that are larger than the scales where dark matter substructure has an impact. In appendix 6.A.1, we show how the choice of the smooth model can affect the substructure analysis. The second consideration is that the substructure signal depends on the source surface brightness variations. This means that we need to ensure that the simulations have the right level of small scale variation in the source to give reliable statistics in the mocks. We also need to allow the model to have sufficient small scale features so as not lead to biases. In appendix 6.A.1, we present illustrative examples and a discussion of how these model biases affect substructure signatures.

For the smooth lens model, we use the same model parameters as in chapter 5. It is composed of a ellipsoidal power law mass distribution, a SIS profile centered on the visible lens substructure and external shear parameters. Additionally, we use shapelets potential perturbations as introduced in chapter 3. We chose the Einstein radius of the lens as the scale parameter of the lens model shapelet. In total, we use 21 additional shapelet parameters (corresponding to $n_{\max} = 5$) which enables us to model the lens mass distribution down to 0.4" resolution.

In the source, we identify, in the residual map of the best fit reconstruction, multiple distinct features that correspond to source structure on smaller scales. We traced back the residual features to the source plane and identify 15 distinct regions with flux being emitted by small scale structure. For each of these specific regions, we position additional source surface brightness shapelets with order $n_{\max, \text{clump}} = 3$ and scale β_{clump} being the magnified resolution limit of the data. This procedure allows us to fit the data well an extensive expansion of the basis set. The source reconstruction is performed independently for the two imaging bands, but with the same configuration and order of the basis sets.

We use the framework of chapter 3 to fit the high dimensional non-linear parameter space by a Particle-Swarm Optimization (205). The best fit lens model results in a reduced χ^2 value

of $\chi^2_{\mathrm{red}} = 1.03$. This model is the starting point for the substructure analysis on the data and allows to explicitly quantify lensing structure emerging from scales below 0.4".

6.2 Simulations

We now compare the statistical features in the image residuals with the features of simulations based on different dark matter models. For our analysis, we target deflection angle perturbations arising from substructure within the lens relative to a smooth lens model. To investigate the statistical significance of a signal, we use multiple realizations of the same physical model (WDM, CDM) and for different halo masses. We chose semi-analytic descriptions to compare a large number of realizations of given models. The details on how we produce the set of simulations and how we incorporate WDM in the semi-analytic description is given in this section.

6.2.1 Dark matter substructure model

To compute a sample of expected deflection perturbations, we use a semi-analytic model based on EPS merger trees and a subhalo evolution and disruption prescriptions, tuned to N-body CDM simulations.

Power spectrum

We follow (8) in describing a power spectrum for WDM. The relative transfer function $r_T(k)$, defined as the ratio of WDM to CDM model can be defined as

$$r_T(k) = [P(k)_{\lambda\mathrm{WDM}} / P(k)_{\lambda\mathrm{CDM}}]^{1/2} \tag{6.1}$$

In case of WDM, the relative transfer function can be described as (17)

$$r_T(k) = \left[1 + (\alpha k)^{2\nu}\right]^{-5/\nu}, \tag{6.2}$$

where α is the scale of the free streaming break of the WDM particle and ν was fixed by (8) to $\nu = 1.12$ and

$$\alpha = 0.049 \left(\frac{m_\chi}{1\text{keV}}\right)^{-1.11} \left(\frac{\Omega_\chi}{0.25}\right)^{0.11} \left(\frac{h}{0.7}\right)^{1.22} h^{-1}\text{Mpc}. \tag{6.3}$$

m_χ is the DM thermal relict mass and Ω_χ is the normalized dark matter density. Of importance for structure formation is the mass variance $S(M)$, which can be computed as

$$S(M) \equiv \sigma^2(M) = \frac{1}{2\pi^2} \int_0^\infty 4\pi k^2 P(k) W^2(k|M) dk, \tag{6.4}$$

which depends on the window function W used. For a top-hat filter in real space, the corresponding mass M for a given filter size R is given by

$$M = \frac{4\pi}{3} \bar{\rho} R^3. \tag{6.5}$$

A sharp-k filter can capture the features of WDM well, as the suppressed high k modes lead to a flattening of $S(M)$. The sharp-k filter

$$W(k|M) = \begin{cases} 1 & \text{if } k \le k_s(M) \\ 0 & \text{if } k > k_s(M) \end{cases} \tag{6.6}$$

defines the mass M according to the filter scale k_s and a clear physical relation on the $k_s(M)$ relation is missing. We follow (316) and relate k_s to the mass radius R as defined in equation 6.5 through a scale parameter a as $k_s = a/R$, where $a = 2.5$ was set by (316) to predict the same turnover in the halo mass function.

Merger tree

To generate stochastic merger trees based on the power spectrum and mass variance, we follow (316), which is based on the original merger tree for CDM structure of (46). Using a sharp-k filter to compute $S(M)$ will change the halo mass function at high mass. To avoid

over-predicting high mass halos, the barrier for collapse has to be increased by a factor of 1.197 (see 316). Furthermore, WDM further changes the collapse threshold below a characteristic mass (317). (316) found, based on work of (317), that the threshold of collapse $\delta_c(M, t)$ from WDM and CDM can be fitted. We adopt for this their equation 7-10 of (317). We do not, for simplicity, incorporate the non-Gaussian walks of a moving barrier in WDM, as proposed by (316). On the other hand, we set the resolution limit of the merger tree to the characteristic mass. This semi-analytic description allows us to produce many realizations of substructure and the in-fall histories of the substructure.

Mass-concentration relation

The concentration of the dark matter halos depend on the mass assembly history. Halos that assemble earlier are more concentrated (e.g. 200; 318; 319). Taking a mass-concentration-redshift relation based on CDM simulations may not reflect the fact that WDM halos form later and therefore are expected to have lower concentrations. We use the $c(M, z)$ relation of (320) for CDM cosmologies, namely

$$c = A \left(\frac{M}{M_{\text{pivot}}} \right)^B (1 + z)^C.$$ (6.7)

Throughout this work, we set $A = 5.22$, $B = -0.072$, $C = -0.42$ and $M_{\text{pivot}} = 2 \cdot 10^{12}$, taken from (320). In the WDMW case, we set the $c(M, z)$ relation by a mapping of the same formation histories of CDM halos as described in (321).

Substructure evolution

To describe the evolution of the subhalos from the time they get assigned to a more massive halo, the information extracted from a merger tree is insufficient, since subhalos get disrupted by tidal stripping. We use the semi-analytic description of (322) to describe the average mass-loss rate of a dark matter subhalo, work based on (323). The parameterized form of the subhalo decay $\dot{m}$ depends on the instantaneous mass ratio m/M as

$$\dot{m} = -A \frac{m}{\tau_{\text{dyn}}} \left(\frac{m}{M} \right)^\zeta,$$ (6.8)

where A and ζ describe the normalization and mass dependents of the subhalo decay, respectively. We use the values derived from (324), $A = 1.54$ (corresponding to $\tau_0 = 2.0$) and $\zeta = 0.07$. We use differential evolution to solve for the final subhalo mass. We do not introduce scatter in this relation (as done by 322) as we expect a realistic scatter from the merger tree realizations.

Dynamical friction plays also a role in subhalo disruption. We use the description of (325), which provide a fitting formula from simulations

$$n_{\text{dyn}} \equiv \frac{\tau_{\text{merge}}}{\tau_{\text{dyn}}} = A \frac{(M/m)^b}{\ln(1 + M/m)} \exp[c\eta] \tag{6.9}$$

where $\eta = j/j_c(E)$ is the orbital circularity. We ignore the dependence on orbital energy. The fitting parameters are tuned to simulations and given by (325) as $A = 0.216$, $b = 1.3$ and $c = 1.9$. (123) showed that the circularity η of subhallos at infall can be well fit by

$$P(\eta) \propto \eta^{1.22}(1 - \eta)^{1.22}. \tag{6.10}$$

We draw from this distribution for every subhalo at in-fall to compute the dynamical friction time. We use $\tau_{\text{dyn}} = 0.1 H^{-1}$ for this computation, which results in the scale factor of disruption a_{dyn}

$$a_{\text{dyn}} = a_{\text{infall}} \exp\left[n_{\text{dyn}}/10\right]. \tag{6.11}$$

Subhalos whose dynamical friction time a_{dyn} is shorter than that needed to survive to the redshift of interest (in our case to the redshift of the lens a_{lens}) are assumed to be fully disrupted.

For the evolution of the subhalo internal structure, we use the description of (326; 327), used as well for comparison in (328). The relation $r_{\text{max}}/r_{\text{max,infall}}$ and $v_{\text{max}}/v_{\text{max,infall}}$ is parameterized as a function of mass loss $x = m/m_{\text{infall}}$ as

$$y(x) = \frac{2^\alpha x^\beta}{(1 + x)^\alpha} \tag{6.12}$$

where α and β are fitting coefficients and $y(x)$ is either the ratio of radii or maximal velocity. (326; 327) provides $(\alpha, \beta) = (0.4, 0.3)$ for v_{max} and $(\alpha, \beta) = (-0.3, 0.4)$ for r_{max}.

175

The semi-analytic prescription used to compute the disruption of halos does not have a specific concentration dependence. In the case of WDM, where the concentration of halos is expected to be lower than for CDM, ignoring the concentration as a parameter in the disruption process might slightly under-predict the strength of disruption. For all of our conclusions, the bounds on the thermal relict mass of a WDM particle are conservative with respect to the subhalo disruption effect.

The spacial positioning of the subhalos is computed based on the orbit at infall. The orbit is computed based on the orbital circularity distribution (see equation 6.10) and provides a radial distribution function from which we draw from. Our tests show that the radial distribution of substructure mass follows the NFW profile of the parent halo.

The prescriptions and assumptions mentioned here allow us to generate mock halos of different masses with substructure with different WDM thermal relict masses. Figure 6.1 shows a set of realizations of halos in the range 12.0-13.5 $\log(M/M_\odot)$ and thermal relict masses 1.0-3.0 keV. The enhanced amount of substructure of higher WDM mass is visible on the figure.

6.2.2   Mock image generation

In our approach, our aim is to separate the effects of global features from those coming from substructure. In generating mock images, we have therefore decided to use a hybrid approach. For the global features, we use the deflections coming from the smooth lens model $\vec{\alpha}_{\text{smooth}}$ that we infer from the data. This allows us to generate mock strong lens systems that are in the same regime as that for the data. The small scale features, which is where dark matter properties have an impact, are modeled using the halos generated by the merger tree procedure described in section 6.2. To do this, we first calculate the full deflection angles for the mock halos $\vec{\alpha}_{dm}$. We then generate a smoothed field $\vec{\alpha}_{dm}^s$ through a convolution with a Gaussian kernel g of width $\phi_{\min}$,

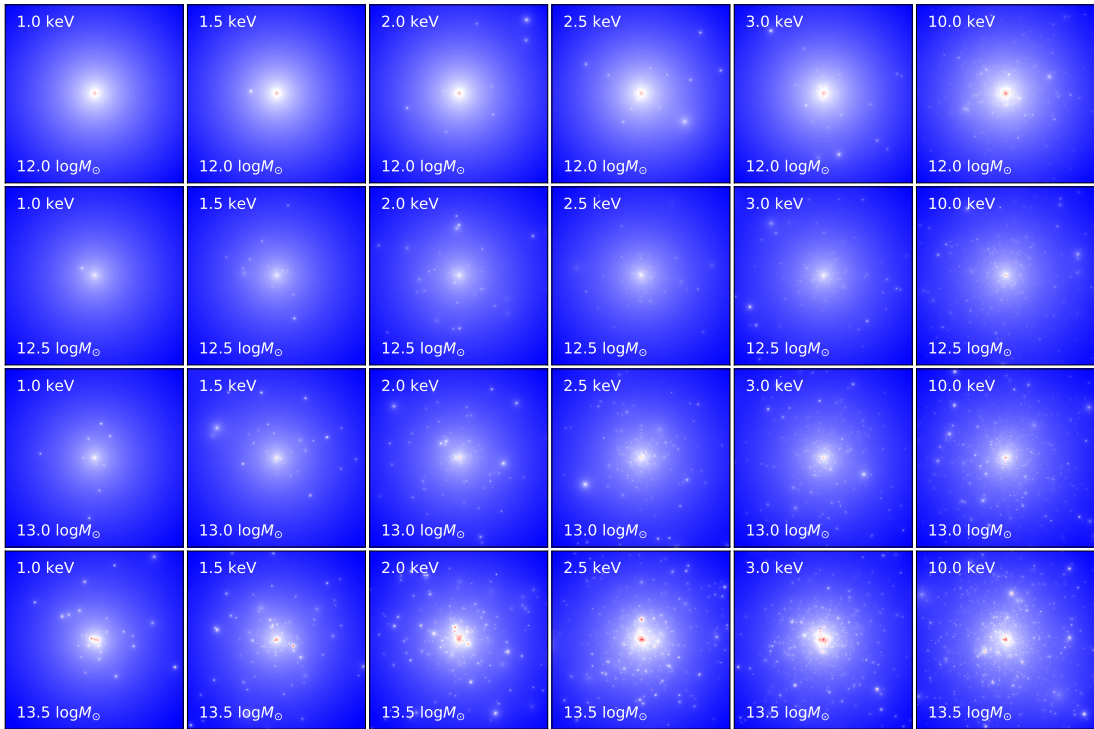$$\vec{\alpha}_{dm}^s = \vec{\alpha}_{dm} * g(\phi_{\min}). \tag{6.13}$$

Figure 6.1: The projected mass of a set of different stochastic semi-analytic substructure realizations. From left to right: Increasing dark matter thermal relict mass from 1 keV to 10 keV. From top to bottom: Increasing parent halo mass from $10^{12} M_{\odot}$-$10^{13.5} M_{\odot}$. The size of the region is identical to the HST image being modeled, i.e. 4" by 4". The color scale is fixed for different dark matter models but change with halo mass.
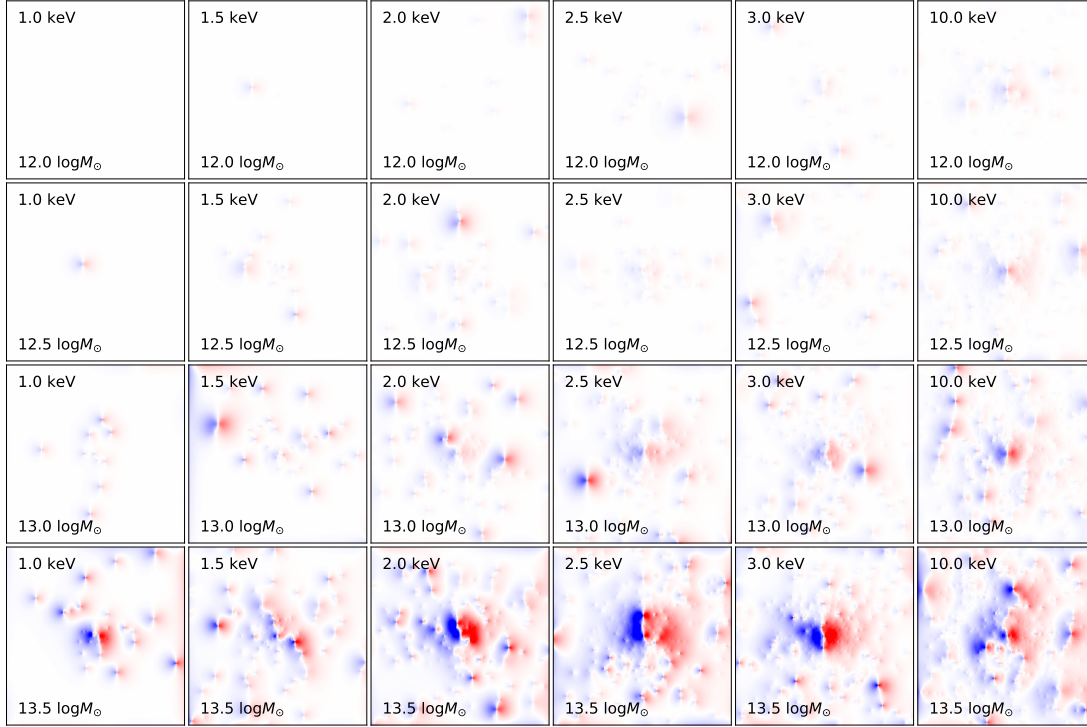
Figure 6.2: The deflection perturbation $\Delta\alpha_{1,dm}$ of the same set of different stochastic semi-analytic substructure realizations as in figure 6.1. From left to right: Increasing dark matter thermal relict mass. From top to bottom: Increasing parent halo masses. The the size is again 4" by 4". The color scale is chosen such that full color (red or blue) indicate deflection perturbation of 0.04".

We then approximate the perturbations to the deflection angles $\Delta\vec{\alpha}_{dm}$ as

$$\Delta\vec{\alpha}_{dm} = \vec{\alpha}_{dm} - \vec{\alpha}_{dm}^s. \tag{6.14}$$

We chose the kernel width $\phi_{\min}$ to match the smooth lens model shapelet deflection scale, $\theta_{\min} = 0.4$". The result is a deflection angle map that has no net mass beyond scales of $\theta_{\min}$. As an example, figure 6.2 shows the deflection perturbation map of one of the components of $\Delta\alpha_{dm}$ for the cases shown in figure 6.1.

The deflection used to generate the mock images is therefore

$$\vec{\alpha}_{mock} = \vec{\alpha}_{\text{smooth}} + \Delta\vec{\alpha}_{dm}. \tag{6.15}$$

We also use the same source model as inferred from the data, the same PSF model with its errors, same weight map and add Poisson and Gaussian noise on the mock image.

6.3 Substructure model comparison

In our forward modeling approach, we develop an analysis method that can be applied to data and mock data in the same way. Instead of trying to identify individual structures within the lens, we use the scanning strategy described below. The feature maps that we generate from these scans can then be studied to discriminate between different models of dark matter.

6.3.1 Substructure scanning procedure

Our scanning procedure works by fitting a lens model that includes a single subclump perturber, with corresponding deflection $\vec{\alpha}_{\mathrm{clump}}$, at a position (x_i, y_i) as well as the smooth global model. The combined deflection angle is then

$$\vec{\alpha}_{\mathrm{pert}}(x_i, y_i) = \vec{\alpha}_{\mathrm{smooth}} + \vec{\alpha}_{\mathrm{clump}}(x_i, y_i), \tag{6.16}$$

where again $\vec{\alpha}_{\mathrm{smooth}}$ the deflection angles for the smooth global model.

We have chosen the perturber to be a truncated SIS profile,

$$\alpha_{\mathrm{clump}}(r) = \begin{cases} \theta_{\mathrm{E,clump}} & \text{if } r \leq r_{\mathrm{trunc}} \\ \theta_{\mathrm{E,clump}} \left(2 - \frac{r}{r_{\mathrm{trunc}}}\right) & \text{if } r_{\mathrm{trunc}} < r \leq 2 r_{\mathrm{trunc}} \\ 0 & \text{if } r > 2 r_{\mathrm{trunc}} \end{cases} \tag{6.17}$$

where $\theta_{\mathrm{E,clump}}$ is the Einstein radius of the clump and r_{trunc} its truncation radius. For our analysis, we set $\theta_{\mathrm{E,clump}} = 0.01"$ and $r_{\mathrm{trunc}} = 0.1"$. The total mass enclosed in the truncation radius is $M_{\mathrm{clump}} = 10^{8.35} M_\odot$.

We then move the position of the subclump perturber in increments of $0.05"$ over the image in both the horizontal and vertical directions. At each position (x_i, y_i), we make a best fit reconstruction of the lensed image I (this can be a fit to data or to simulated data) using the

perturbed model ($\vec{\alpha}_{\text{pert}}$). We call this reconstructed image I_i^{pert}. We also perform fits using the smooth model $\vec{\alpha}_{\text{smooth}}$ leading to reconstructed images I_i^{smooth}. When performing a fit we again use the source basis sets as described in section 6.1. In addition, we add source surface brightness shapelets of order $n_{\text{max}} = 5$ at the source position

$$(x_i', y_i') = (x_i, y_i) - \vec{\alpha}_{\text{smooth}}(x_i, y_i), \tag{6.18}$$

which corresponds to the source plane position being mapped through the subclump pertruber. We set the scale β of these additional shapelets to the average magnification within the truncated SIS profile at (x_i, y_i).

From the set of images that we generate, we can define three different residual measure (R), related to the relative agreement of the images I, I_i^{pert} and I_i^{smooth} as

$$R_i^{\text{pert}} = \sum \frac{1}{\sigma^2} \left(I - I_i^{\text{pert}} \right)^2, \tag{6.19}$$

$$R_i^{\text{smooth}} = \sum \frac{1}{\sigma^2} \left(I - I_i^{\text{smooth}} \right)^2, \tag{6.20}$$

and

$$R_i^{\text{sens}} = \sum \frac{1}{\sigma^2} \left(I_i^{\text{pert}} - I_i^{\text{smooth}} \right)^2, \tag{6.21}$$

where the sum is over the pixels of the image and σ is the noise associated with each pixel. We further introduce a relative excess distance ΔR_i defined as

$$\Delta R_i = R_i^{\text{pert}} - R_i^{\text{smooth}}. \tag{6.22}$$

Figure 6.3 shows the scanning results, in terms of ΔR values, for the HST data (left column), a model with a halo mass of $10^{13.5} M_\odot$ (second column), a model with a halo mass of $10^{13.0} M_\odot$ (third column). The last column shows the sensitivity map, R_i^{smooth}. The different rows indicate
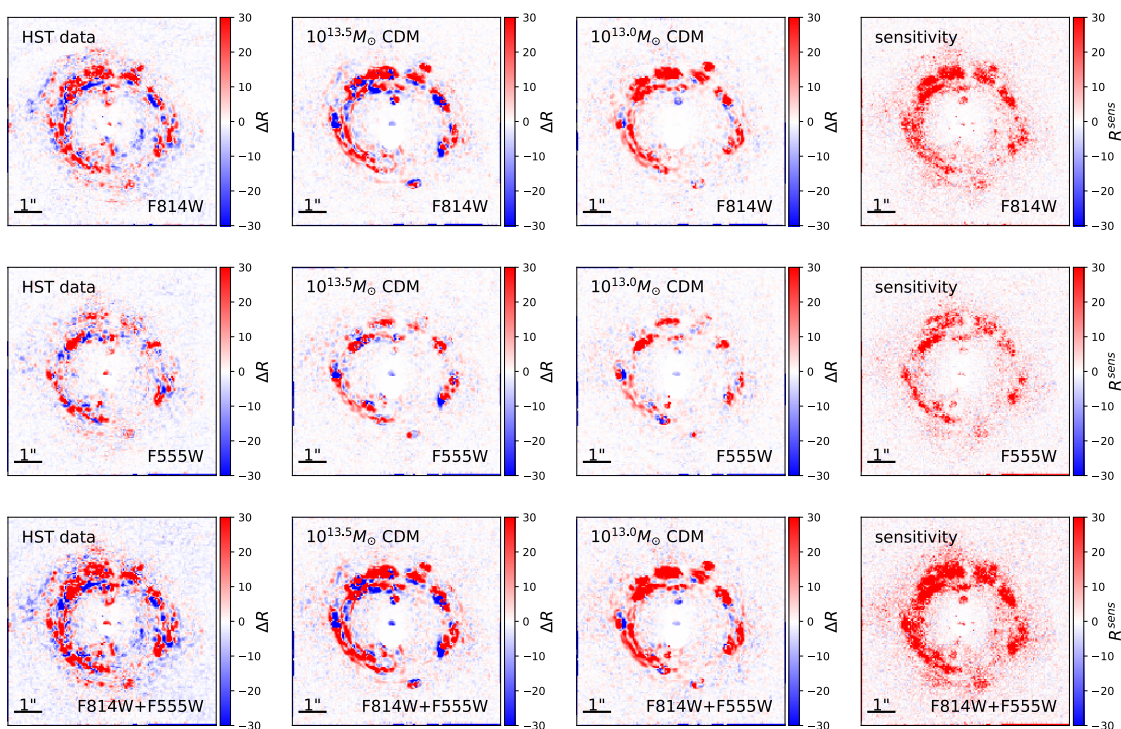
Figure 6.3: The scanning results for ΔR_i of the HST data (left column) and two selected CDM semi-analytic realizations with halo masses $10^{13.5} M_\odot$ and $10^{13} M_\odot$ (middle two columns) and the sensitivity map (right column). The different rows indicate the analysis of filter F814W (top), F555W (middle) and combined F814W+F555W (bottom). Each pixel in the plot reflects ΔD_i when placing the perturber at the position of the pixel.

the results for different filters. Filter F814W results are shown in the top row, F555W in the middle and the combined results F814W+F555W at the bottom.

6.3.2 Model comparison based on summary statistics

Given the scan maps that we are able to generate for both data and mocks, the challenge is to construct a statistical test that will allow us to select from different dark matter models. The framework of Approximate Bayesian Computing (ABC) (329; 330) allows to construct a posterior distribution even for cases where the likelihood of a model, given the data, is not accessible. ABC requires distance measures (metric distances) between the output of different simulations, based on summary statistics. In general, the summary statistic can take many forms but need to map to one or several metric distances.

Our goal is to construct measures that are sensitive to dark matter properties. This means

that comparisons of different realization of the same dark matter model should result in small distances D. One the other hand, comparing realizations of different dark matter models should result in large distances.

We find that a good distance measure can be constructed based on the cumulative distribution $P(x) = N(\Delta R_i < x)$ and the spherical averaged two-point correlation function $C(dr) = \langle \Delta R_i(r) \Delta R_i(r+dr) \rangle_r$, both based on the relative excess distance ΔR_i (6.22) of the scanning procedure.

The specific metric we adopt is the product of the two metrics involving the cumulative distribution D_N and the correlation function D_C

$$D(\Delta R_i^1, \Delta R_i^2) = D_C(\Delta R_i^1, \Delta R_i^2) \cdot D_N(\Delta R_i^1, \Delta R_i^2),$$
(6.23)

where

$$D_N(\Delta R_i^1, \Delta R_i^2) = \sum_x \left[P(\Delta R_i^1, x) - P(\Delta R_i^2, x) \right]^2$$
(6.24)

and

$$D_C(\Delta R_i^1, \Delta R_i^2) = \sum_{dr} \left[C(\Delta R_i^1, dr) - C(\Delta R_i^2, dr) \right]^2$$
(6.25)

are the quadratic distances between the two distributions in respect of $P(x)$ and $C(dr)$. The sum in D_N is uniform spaced in the range ΔR_i in $[10, 300]$ in incremental increases of $\Delta R_i = 10$. The sum in D_C is uniformly spaced in the range $0" - 2.5"$ in incremental increases of $\Delta r = 0.05"$.

The expression 6.23 provides, without normalization, an equal weight on both diagnostics. We emphasis that the expression in equation 6.23 is in no means a likelihood.

The cumulative distribution is an indicator that quantifies the strength of the substructure signal and the correlation function to quantify the spatial signature. Figure 6.4 shows the cumulative distribution $P(x)$ for different realizations of the same dark matter models, split in four different halo masses and six different dark matter models. The cumulative distribution

can well distinguish different halo masses. Higher halo masses relate to more and more massive substructures.

Figure 6.5 shows the one-dimensional correlation function $C(dr)$ for different realizations of the same dark matter models with the same splits in halo masses as in figure 6.4. We see clear discrimination of the correlation function of different dark matter model samples. In both Figures, the thick black line indicates the statistics of the data.

6.4   Dark Matter model constraints

In this work, we focus on thermal relict mass constraints for dark matter. From the diagnostics of section 6.3 and Figure 6.4 and 6.5 we see that the mass of the parent halo has a significant impact on our distance metric and needs to be included as a free parameter in our analysis. In this chapter therefore, we will focus on the two parameters space of relic mass and parent halo mass M_h. We do this by making 20 realizations per point in parameter space for a grid of $[1.0, 1.5, 2.0, 2.5, 3.0, 10]$ keV in thermal relict mass and $[12.0, 12.5, 13.0, 13.5]$ $\log(M/M_\odot)$ for the parent halo mass. This leads to a total of 480 realizations. This current number of realizations is a little too low for a converged likelihood with the ABC formalism but it already allows us to construct a conservative inference of our parameters.

Figure 6.6 shows the 1- and 2-σ posterior distribution when performing a distance threshold cut so the best 100 realizations are selected. This is an ABC estimated posterior with a weak (i.e. conservative) threshold.

Even with this conservative procedure, we can exclude parent halo masses below $10^{12.9} M_\odot$ with $> 2\sigma$ confidence. This alone shows that we are detecting a significant amount of substructure and that the statistical information contained in this analysis allows us to make measurements that are not sample variance limited. We see that the ABC method preferentially selects models with high keV thermal relict masses. The thermal relict constraints are partially degenerate with the parent halo mass. Nevertheless, even for high parent halo masses, e.g. $10^{13.5} M_\odot$, the pattern is inconsistent with WDM with $m_{TH} <$2keV on the $2-\sigma$ level and in agreement with a CDM scenario. This result is comparable in strength and in agreement with the other probes of the small scale structure. The latest constraints from the Lyman-α forest (20) results in a
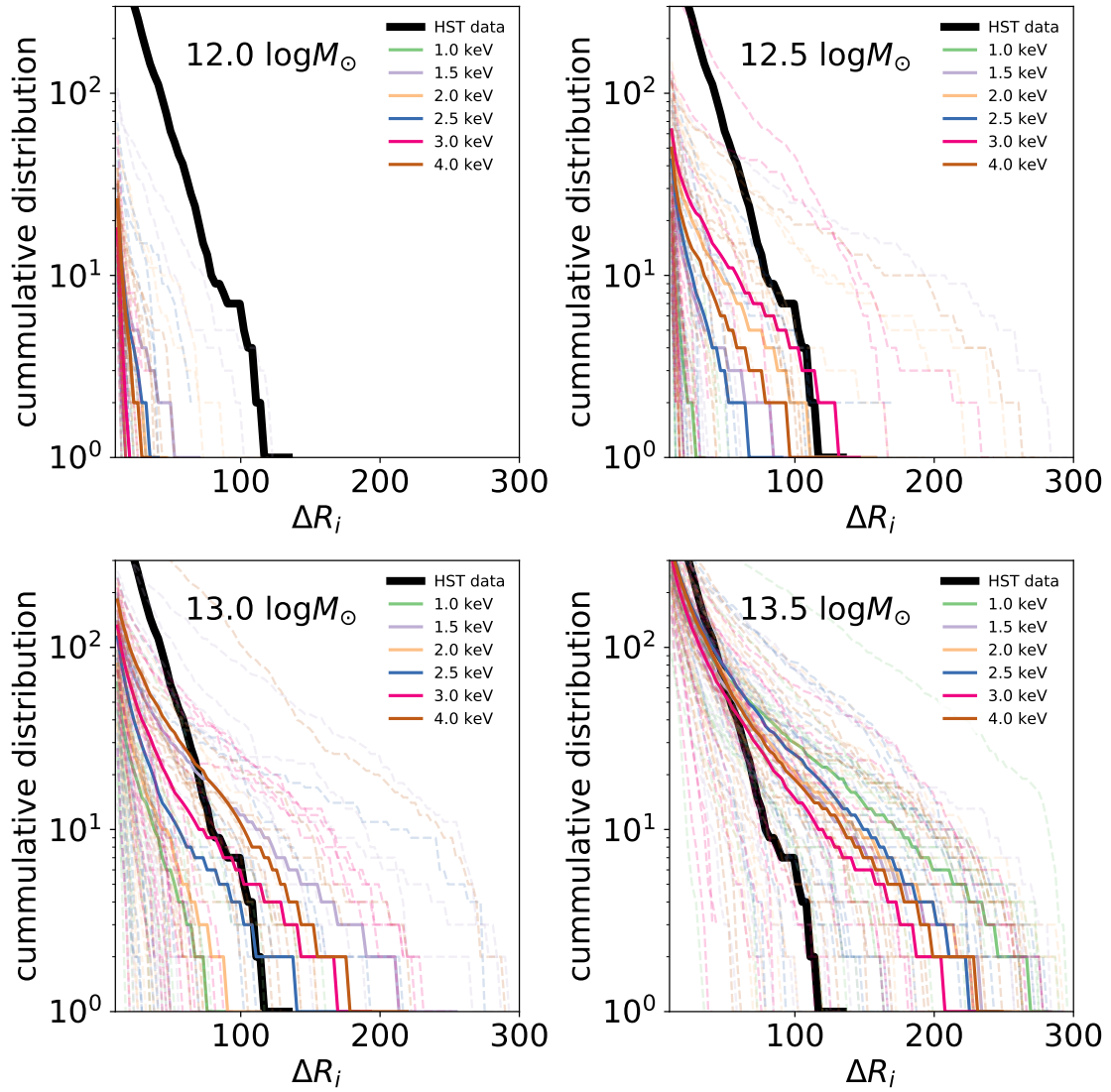
Figure 6.4: Cumulative distribution $P(x) = N(\Delta R_i < x)$ for different realization. Top left to bottom right: Increasing parent halo mass. Bold black line indicates the distribution of the data. Continuous lines are the mean distribution for the different dark matter models. Dashed lines are individual realizations of the mocks.
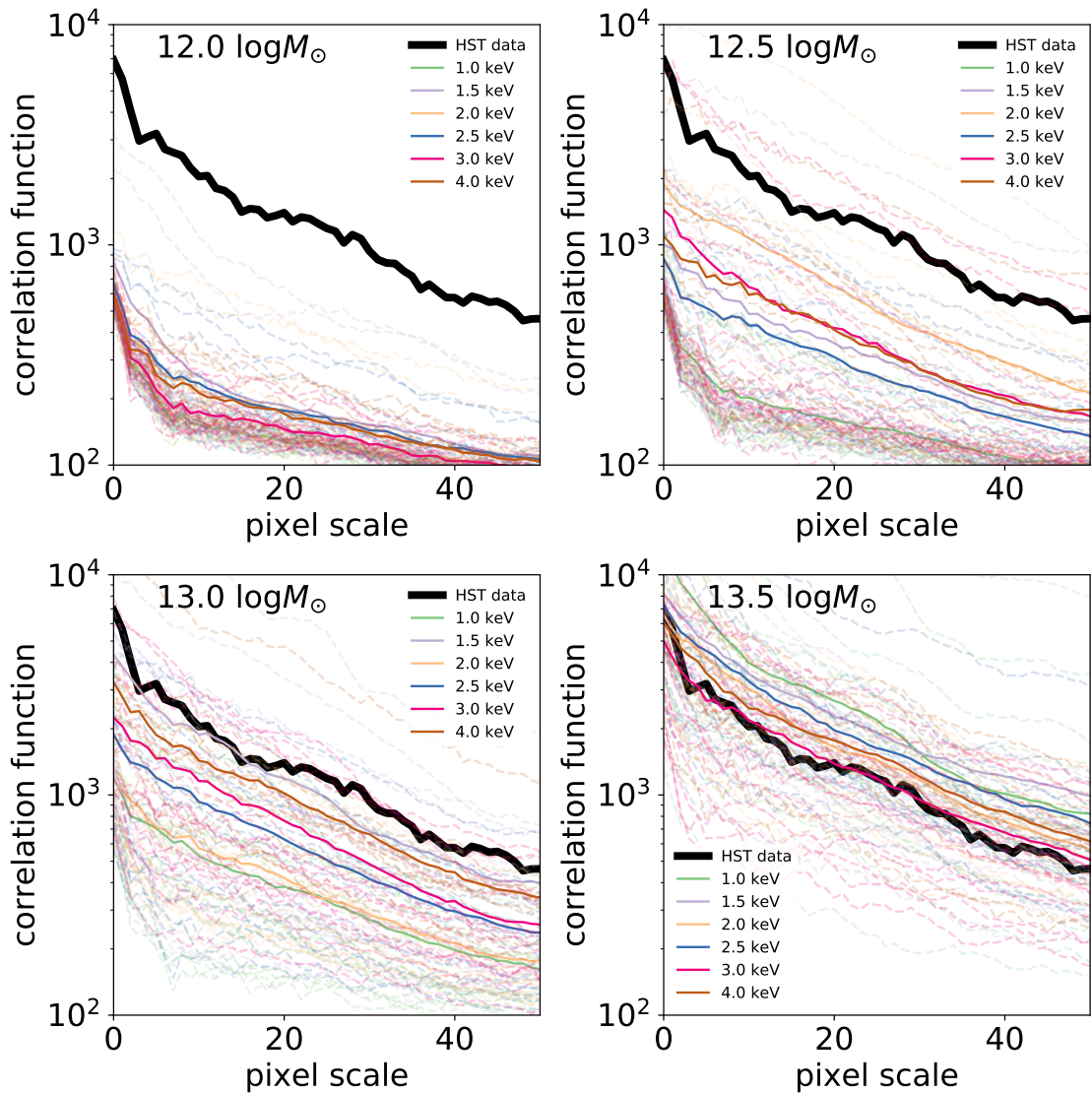
Figure 6.5: Correlation function of the excess residuals $C(dr)$ for different realizations. Top left to bottom right: Increasing parent halo mass. Bold black line indicates the distribution of the data. Continuous lines are the mean distribution for the different dark matter models. Dashed lines are individual realizations of the mocks.

lower bound on the thermal relic mass $m_{TH} = 3.3$keV at the 2σ confidence level. The limits obtained from dwarf galaxy counts disfavors particle masses below $m_{TH} = 2.3$keV at the 2σ confidence level(21; 22).

6.5 Discussion

Our method is able to statistically capture substructure features in strong lensing systems and allows us to discriminate between different dark matter models. Our substructure scanning probes the mass range of $M_{clump} = 10^{8.35} M_{\odot}$ withing 0.1" (see section 6.3.1). This filter matches qualitatively a free streaming mass of a 2keV particle [1].

The model shows that multiple substructures within the same mass range have a significant impact on the lensing statistics. This means that probing substructure on the one-by-one basis at this mass limit may not be feasible. However, when pushing the limits to lower substructure masses, a statistical approach like ours is able to account for the effects of hundreds of subclumps simultaneously. For instance, this can be important for ALMA. ALMA can provide higher resolution data than HST images, which leads to greater potential for measuring smaller lens perturbations (see e.g. 302; 331; 332) if the information from these higher resolution data can be properly tapped.

Our method can be extended to probe other statistics of lensing substructure, such as mass-concentration relations and disruption efficiencies and might provide a window to probe suggested solutions of the cusp-core discrepancies. Such studies are in reach with the current HST data and are only limited by the predictive power of the physical models and the sample variance.

Our current constraints are mainly limited by the statistics of the single lens and the moderate sample size from our simulations. This results in a conservative estimate of the likelihood. Better data (more lenses, better quality data) can discriminate models with high significance. Tackling the partial degeneracy between parent halo mass and dark matter thermal relict mass can also be done by incorporating additional, independent, priors on the halo mass. For

---

[1]The mean cosmic mass enclosed within a free-streaming volume for a 2keV particle is roughly $10^{9.2} M_{\odot}$. However, the effects that we are sensitive to, and are probed by our filter, come from the central regions of collapsed clumps, which contain a fraction of the overall free-streaming mass.
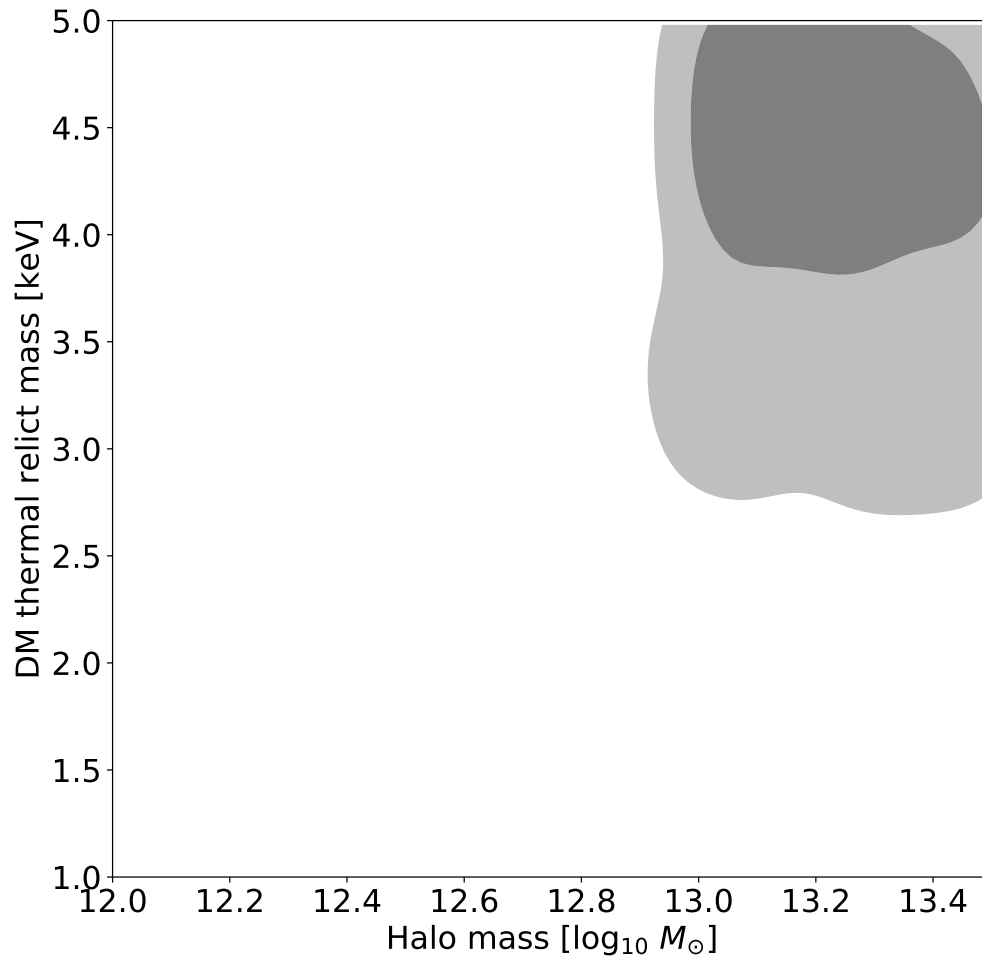
Figure 6.6: $1 - \sigma$ (dark region) and $2 - \sigma$ (light region) posterior distribution estimated by ABC method on the thermal relict mass vs halo mass plane from the lens RXJ1131-1231. The sample number is limited and the details in the posterior distribution is not fully converged.

instance coming from abundance matching or galaxy-galaxy lensing. What we do have from the strong lensing measures is an accurate measure of the total mass (dark matter + baryonic mass) within the Einstein, which for RXJ1131-1231 is $M_{<\theta_E} = 10^{11.9} M_\odot$. The host galaxy of the lens is a very massive elliptical early-type galaxy. A significant fraction of the mass withing θ_E comes from baryonic matter. These observations allow us to set a conservative lower limit on the expected halo mass of the lens RXJ1131-1231 to be $10^{13} M_\odot$ from abundance matching and forward modeling of the galaxy population through cosmic time (e.g. 78; 28), as shown in chapter 2.

Systematics in the lens modeling can be another limiting factor. In this work, we focus on the source reconstruction scale and the intermediate lens model scale descriptions. Further effects that we do not include could mimic lensing substructure effects, such as micro-lensing by stars of the lensing galaxy, luminous structure of dwarf galaxies, substructure displaced along the line-of-sight or dust extinction. For the current constraints, we do not expect a major impact of such effects.

6.6   Summary

In this chapter, we presented a statistical analysis of the substructure content in the strong lens RXJ1131-1231. We report a lower limit for the mass of thermal relic dark matter of 2keV at 2σ confidence level. This is comparable to earlier limits derived from Lyman-α clustering and MW dwarf counts. Our statistical method significantly improves upon clump-by-clump strong lensing analyses. It can easily be extended to multiple lens systems.

6.A   Appendix of chapter 6

6.A.1   Systematics in the modeling

Spurious effects in the reconstruction modeling may lead to effects in the substructure scanning that may be interpreted as caused by substructure. Here, we investigate on the resolution of the smooth model and on the source surface brightness reconstruction.
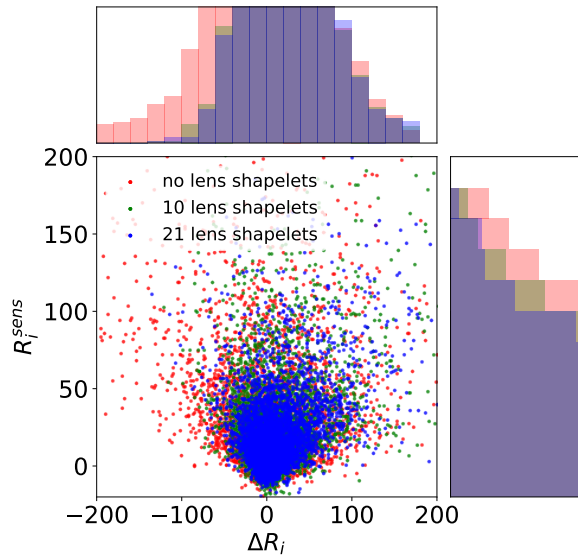
Figure 6.7: The distributions $\left(\Delta R_i, R_i^{\text{sens}}\right)$ for the data with different resolution in the smooth lens model. Positive y-axis refers to the expected improve in the fit in the presence of the perturber in the mock data. The projected samples are displayed above and on the right.

Substructure deflection perturbations

The smooth lens model described in section 6.1 allows us to use multiple shapelet potentials. This basis set enables to model the intermediate scales of the lens model. To test the impact on the substructure scanning analysis, we model the data with different number of lens model shapelets in addition to the smooth model. We fit the data with 10 ($n_{\text{max}} = 3$) and 21 ($n_{\text{max}} = 5$) additional shapelets. Figure 6.7 shows the $\left(\Delta R_i, R_i^{\text{sens}}\right)$ distributions of those models, including a model without any additional lens shapelets. We see that the statistics changes dramatically when adding the first 10 shapelets. This means that the most of the "signal" in the smooth model can be attributed to features arising from scales captured by the first 10 lens shapelets. Increasing the shapelet number to 21 does not lead to a significant change in the statistics. Our underlying assumption is that the features present in the scanning of the model with 21 shapelets arise from scales smaller than being captured by the lens shapelets.

Source surface brightness

The specific source reconstruction model may also have an effect on the substructure scanning statistics. To investigate this dependence, we take the lens model with 21 additional (our default model) lens shapelets and perform the scanning based on three different source reconstructions. First, we only reconstruct the source with the global shapelets with $n_{\mathrm{max}} = 50$. Second, we add the nested shapelet description at the position of the lens perturber (see section 6.1). Third, additionally to the nested shapelets, we add 15 fixed high resolution clumps to the source model (also described in section 6.1). Figure 6.8 shows the scanning statistics of those three reconstruction models. Substructure perturbation can change the magnification locally significantly. When the source reconstruction can not resolve the existing scales, substructure can significantly help in reconstructing the image because it can demagnify those regions such that the source reconstruction description better match the scales involved. This effect is an artifact. A substructure detection method requires to be able to describe the smallest scales involved in the source surface brightness. We tested our method by further enhancing the nested shapelets and increasing the additionally modeled source clumps. Neither of those pushes to smaller scales did significantly change the substructure scanning statistics. The underlying assumption in our inference is that we are able to match the smallest scales in the source reconstruction description relevant to match the observations.

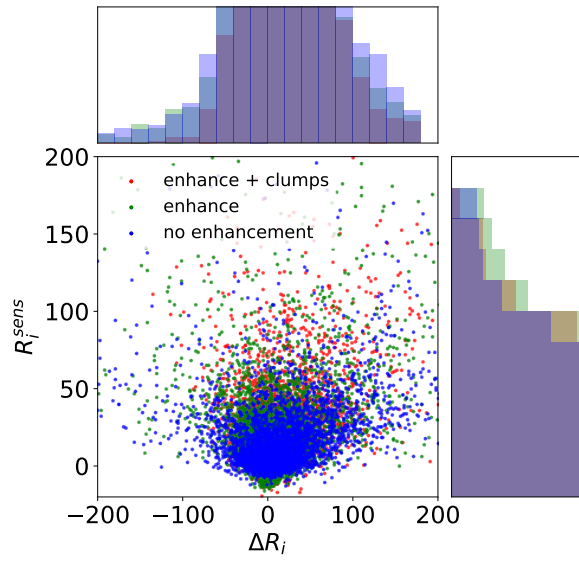Figure 6.8: The distributions $\left(\Delta R_i, R_i^{\text{sens}}\right)$ for the data with different resolutions and descriptions of the source surface brightness reconstruction. Negative x-axis refers to improved fit when adding a perturber. Positive y-axis refers to the expected improve in the fit in the presence of the perturber in the mock data. The projected samples are displayed above and on the right.

7 Conclusions

The aim of this thesis was to probe the fundamental predictions of the standard model of cosmology with the main emphasis on the small cosmological scales.

To link the luminous structure with the underlying dark matter structure, we developed a semi-analytic phenomenological model of galaxy evolution embedded in cosmological dark matter structure formation by incorporating pre-defined baryonic prescriptions into a dark matter hierarchical merger tree. Specifically the model is based on the simple gas-regulator model introduced by (91) coupled with the empirical quenching rules of (83; 85).

Having worked on the connection between galaxies and halos, we then turned to study sub-galactic scales using strong gravitational lensing. For this purpose, we developed a strong lensing modeling technique based on versatile basis sets for the lens and source planes. The method uses high performance Monte Carlo algorithms, allows for an adaptive build up of complexity and bridges the gap between parametric and pixel based reconstruction methods. The source reconstruction technique is explicitly source scale dependent, numerically stable and computationally fast compared with pixelized source grid reconstruction techniques.

Furthermore, this thesis presents a simple method to accurately infer line-of-sight (LOS) integrated lensing effects for galaxy scale strong lens systems through image reconstruction. The approach is able to separate weak lensing LOS effects from the main strong lens deflector. We performed a halo rendering based on a mass-to-light description to predict the LOS integrated

weak lensing effect on strong lens systems. We demonstrated the merit of combining a halo rendering approach with strong lens image reconstruction in gaining further insights in the galaxy-halo connection and the non-negligible impact of the LOS structure on the inference from strong lens modeling. Joint constraints of multiple strong lens systems may add valuable information to the galaxy-halo connection and may allow independent weak lensing shear measurement calibrations.

To probe the late time expansion of the universe and the nature of dark energy, we consistently incorporated known lensing degeneracies in the lens reconstruction modeling and propagated the uncertainty to time-delay cosmographic inferences. The method is based on assigning priors on the physical size of the source galaxy and does not crucially rely on lens model assumptions. A specific strong lens source reconstruction technique applied on a specific strong lens system may preferentially reconstruct certain source sizes better than others and can therefore indirectly bias the results of the lens mass inference. Control over source size is crucial to avoid biasing. The method presented in this thesis allows for full control over the source size.

To probe the small scale structure formation and the nature of dark matter, we analyzed the substructure content of strong lenses in a statistical manner, relying on extensive forward modeling and the use of Approximate Bayesian Computing. To create simulations with realistic lensing substructure content, we extended semi-analytic methods to predict the small scale substructure content for different dark matter particle models. This framework allows us to predict the expected substructure content in number, spacial location and internal structure for different dark matter thermal relict masses.

Matching the source resolution and intermediate scales in the lens model of simulation and data is of crucial importance. Any artefact arising from either non-resolved source surface brightness variations or larger scale misfittings in the lens model can give rise to relative residuals in the substructure quantification and can be misinterpreted as an overestimation of the substructure content. To overcome this problem, specific enhancement of the source resolution at the position where lensing substructure is explored is conducted.

Applying our developments to current data, in particular from high resolution HST images,

led to the following main results: (1) We showed that a simple phenomenological model can be build to incorporate the baryonic processes important for galaxy evolution in a hierarchical growth scenario of dark matter structure. (2) For the strong lens system RXJ1131-1231, we reconstructed a reliable high resolution image of the source in two bands with resolved small scale features. (3) We model the lens COSMOS0038+4133 through image reconstruction and its LOS with halo rendering and measure the LOS integrated shear to a precision of ± 0.003. (4) The Hubble parameter H_0 estimate of the lens system RXJ1131-1231 is highly dependent on the dynamical model and the source scale prior and is consistent with current cosmic microwave background experiments. (5) The substructure content in the strong lens RXJ1131-1231 is consistent with a cold dark matter particle scenario. Bounds on the free streaming mass of a possible dark matter particle can be set to 2keV.

The precision gain by these new techniques and results shed new light on the distribution of dark matter, its physical nature and its relation to visible matter. Moreover, the number of strong lensing systems being discovered is growing rapidly and the quality of the observations which will be significantly improved by the next generations of instruments. Together, these developments offer great prospects for the high-precision study of dark matter and other sectors of the standard cosmological model with strong gravitational lensing.

Bibliography

[1] Planck and P. A. R. Ade, *Planck 2015 results. xiii. cosmological parameters, arXiv* **astro-ph.CO** (Feb, 2015) , [1502.01589v2].

[2] L. Anderson, É. Aubourg, S. Bailey, F. Beutler, V. Bhardwaj, M. Blanton et al., *The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: baryon acoustic oscillations in the data releases 10 and 11 galaxy samples, Monthly Notices of the Royal Astronomical Society* **441** (Jun, 2014) 24.

[3] C. Heymans, E. Grocutt, A. Heavens, M. Kilbinger, T. Kitching, F. Simpson et al., *Cfhtlens tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments, Monthly Notices of the Royal Astronomical Society* **432** (Jan, 7) 2433–2433–2453–2453.

[4] T. D. E. S. Collaboration, T. Abbott, F. B. Abdalla, S. Allam, A. Amara, J. Annis et al., *Cosmology from cosmic shear with des science verification data, eprint arXiv* **1507** (Jul, 2015) arXiv:1507.05552.

[5] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant, The Astronomical Journal* **116** (Sep, 1998) 1009.

[6] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro et al., *Measurements of ω and λ from 42 high-redshift supernovae,* The Astrophysical Journal **517** (Jun, 1999) 565.

Bibliography

[7] T. de Haan, B. A. Benson, L. E. Bleem, S. W. Allen, D. E. Applegate, M. L. N. Ashby et al., *Cosmological constraints from galaxy clusters in the 2500 square-degree spt-sz survey, arXiv* **astro-ph.CO** (Mar, 2016) , [1603.06522v1].

[8] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese and A. Riotto, *Constraining warm dark matter candidates including sterile neutrinos and light gravitinos with wmap and the lyman- forest, Physical Review D* **71** (Mar, 2005) 063534.

[9] U. Seljak, A. Slosar and P. McDonald, *Cosmological parameters from combining the lyman- forest with cmb, galaxy clustering and sn constraints, Journal of Cosmology and Astroparticle Physics* **10** (Oct, 2006) 014.

[10] G. Kauffmann, S. D. M. White and B. Guiderdoni, *The formation and evolution of galaxies within merging dark matter haloes, Monthly Notices of the Royal Astronomical Society* **264** (Sep, 1993) 201.

[11] A. Klypin, A. V. Kravtsov, O. Valenzuela and F. Prada, *Where are the missing galactic satellites?, The Astrophysical Journal* **522** (Sep, 1999) 82.

[12] B. Moore, T. Quinn, F. Governato, J. Stadel and G. Lake, *Cold collapse and the core catastrophe, Monthly Notices of the Royal Astronomical Society* **310** (Dec, 1999) 1147.

[13] A. Kravtsov, *The dark matter annihilation signal from dwarf galaxies and subhalos, Advances in Astronomy* **2010** (Jan, 2010) –.

[14] M. Boylan-Kolchin, J. S. Bullock and M. Kaplinghat, *Too big to fail? the puzzling darkness of massive milky way subhaloes, arXiv* **astro-ph.CO** (Feb, 2011) , [1103.0007v2].

[15] A. D. Popolo, *On the dark matter haloes inner structure and galaxy morphology, Astrophysics and Space Science* **361** (Jul, 2016) 222.

[16] D. J. Sand, T. Treu, G. P. Smith and R. S. Ellis, *The dark matter distribution in the central regions of galaxy clusters, arXiv* **astro-ph** (Oct, 2003) , [astro-ph/0310703v1].

[17] P. Bode, J. P. Ostriker and N. Turok, *Halo formation in warm dark matter models, The Astrophysical Journal* **556** (Jul, 2001) 93.

[18] K. Abazajian, *Linear cosmological structure limits on warm dark matter, Physical Review D* **73** (Mar, 2006) 063513.

[19] D. N. Spergel and P. J. Steinhardt, *Observational evidence for self-interacting cold dark matter, Physical Review Letters* **84** (Apr, 2000) 3760.

[20] M. Viel, G. D. Becker, J. S. Bolton and M. G. Haehnelt, *Warm dark matter as a solution to the small scale crisis: New constraints from high redshift lyman- forest data, Physical Review D* **88** (Aug, 2013) 043502.

[21] E. Polisensky and M. Ricotti, *Constraints on the dark matter particle mass from the number of milky way satellites, Physical Review D* **83** (Feb, 2011) 043506.

[22] R. Kennedy, C. Frenk, S. Cole and A. Benson, *Constraining the warm dark matter particle mass with milky way satellites, Monthly Notices of the Royal Astronomical Society* **442** (Aug, 2014) 2487.

[23] K. Bechtol, A. Drlica-Wagner, E. Balbinot, A. Pieres, J. D. Simon, B. Yanny et al., *Eight new milky way companions discovered in first-year dark energy survey data, The Astrophysical Journal* **807** (Jul, 2015) 50.

[24] A. Drlica-Wagner, K. Bechtol, E. S. Rykoff, E. Luque, A. Queiroz, Y.-Y. Mao et al., *Eight ultra-faint galaxy candidates discovered in year two of the dark energy survey, The Astrophysical Journal* **813** (Nov, 2015) 109.

[25] T. F.-L. Collaboration, :, M. Ackermann, A. Albert, B. Anderson, L. Baldini et al., *Dark matter constraints from observations of 25 milky way satellite galaxies with the fermi large area telescope, eprint arXiv* **1310** (Oct, 2013) 828.

[26] A. Drlica-Wagner, A. Albert, K. Bechtol, M. Wood, L. Strigari, M. Sánchez-Conde et al., *Search for gamma-ray emission from des dwarf spheroidal galaxy candidates with fermi-lat data, The Astrophysical Journal Letters* **809** (Aug, 2015) L4.

[27] T. M. Undagoitia and L. Rauch, *Dark matter direct-detection experiments, Journal of Physics G: Nuclear and Particle Physics* **43** (Jan, 2016) 013001.

[28] S. Birrer, S. Lilly, A. Amara, A. Paranjape and A. Refregier, *A simple model linking galaxy and dark matter evolution, The Astrophysical Journal* **793** (Sep, 2014) 12.

[29] S. Birrer, A. Amara and A. Refregier, *Gravitational lens modeling with basis sets, The Astrophysical Journal* **813** (Nov, 2015) 102.

[30] S. Birrer, A. Amara and A. Refregier, *The mass-sheet degeneracy and time-delay cosmography: analysis of the strong lens rxj1131-1231, Journal of Cosmology and Astroparticle Physics* **08** (Aug, 2016) 020.

[31] S. Birrer, C. Welschen, A. Amara and A. Refregier, *Line-of-sight effects in strong lensing: Putting theory into practice, eprint arXiv* **1610** (Oct, 2016) arXiv:1610.01599.

[32] S. Dodelson, *Modern cosmology, Modern cosmology / Scott Dodelson. Amsterdam (Netherlands): Academic Press. ISBN 0-12-219141-2* (Jan, 2003) .

[33] J. A. Peacock, *Cosmological physics, Cosmological Physics* (Jan, 1999) 704.

[34] A. Friedmann, *Über die krümmung des raumes, Zeitschrift fÃ¼r Physik* **10** (Jan, 1922) 377.

[35] A. Friedmann, *Über die möglichkeit einer welt mit konstanter negativer krümmung des raumes, Zeitschrift für Physik* **21** (Dec, 1924) 326.

[36] G. Lemaître, *Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques, Annales de la Société Scientifique de Bruxelles* **47** (Jan, 1927) 49.

[37] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, M. Blanton, A. S. Bolton et al., *The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: baryon acoustic oscillations in the data release 9 spectroscopic galaxy sample, Monthly Notices of the Royal Astronomical Society* **427** (Dec, 2012) 3435.

[38] V. Bonvin, F. Courbin, S. H. Suyu, P. J. Marshall, C. E. Rusu, D. Sluse et al., *H0licow v. new cosmograil time delays of he0435-1223, arXiv* **astro-ph.CO** (Jul, 2016) , [1607.01790v1].

[39] A. G. Riess, L. M. Macri, S. L. Hoffmann, D. Scolnic, S. Casertano, A. V. Filippenko et al., *A 2.4hubble constant, The Astrophysical Journal* **826** (Jul, 2016) 56.

[40] W. H. Press and P. Schechter, *Formation of galaxies and clusters of galaxies by self-similar gravitational condensation, Astrophysical Journal* **187** (Feb, 1974) 425.

[41] J. R. Bond, S. Cole, G. Efstathiou and N. Kaiser, *Excursion set mass functions for hierarchical gaussian fluctuations, Astrophysical Journal* **379** (Oct, 1991) 440.

[42] R. K. Sheth, H. J. Mo and G. Tormen, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes, Monthly Notices of the Royal Astronomical Society* **323** (May, 2001) 1.

[43] R. K. Sheth and G. Tormen, *Large-scale bias and the peak background split, Monthly Notices of the Royal Astronomical Society* **308** (Sep, 1999) 119.

[44] R. S. Somerville and T. S. Kolatt, *How to plant a merger tree, Monthly Notices of the Royal Astronomical Society* **305** (May, 1999) 1.

[45] S. Cole, C. G. Lacey, C. M. Baugh and C. S. Frenk, *Hierarchical galaxy formation, Monthly Notices of the Royal Astronomical Society* **319** (Nov, 2000) 168.

[46] H. Parkinson, S. Cole and J. Helly, *Generating dark matter halo merger trees, Monthly Notices of the Royal Astronomical Society* **383** (Jan, 2008) 557.

[47] J. Zhang, O. Fakhouri and C.-P. Ma, *How to grow a healthy merger tree, Monthly Notices of the Royal Astronomical Society* **389** (Oct, 2008) 1521.

[48] A. Einstein, *Die grundlage der allgemeinen relativitätstheorie, Annalen der Physik* (Jan, 1916) .

[49] A. Eddington, *The total eclipse of 1919 may 29 and the influence of gravitation on light, The Observatory* (Jan, 1919) .

[50] C. McCully, C. R. Keeton, K. C. Wong and A. I. Zabludoff, *A new hybrid framework to efficiently model lines of sight to gravitational lenses, Monthly Notices of the Royal Astronomical Society* **443** (Oct, 2014) 3631.

[51] V. Springel and L. Hernquist, *Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation, Monthly Notice of the Royal Astronomical Society* **339** (Feb, 2003) 289.

[52] A. A. Klypin, S. Trujillo-Gomez and J. Primack, *Dark matter halos in the standard cosmological model: Results from the bolshoi simulation, The Astrophysical Journal* **740** (Oct, 2011) 102.

[53] C. Lacey and J. Silk, *Tidally triggered galaxy formation. i - evolution of the galaxy luminosity function, Astrophysical Journal* **381** (Nov, 1991) 14.

[54] S. D. M. White and C. S. Frenk, *Galaxy formation through hierarchical clustering, Astrophysical Journal* **379** (Sep, 1991) 52.

[55] R. S. Somerville and J. R. Primack, *Semi-analytic modelling of galaxy formation: the local universe, Monthly Notices of the Royal Astronomical Society* **310** (Dec, 1999) 1087.

[56] G. Kauffmann, J. M. Colberg, A. Diaferio and S. D. M. White, *Clustering of galaxies in a hierarchical universe - i. methods and results at z=0, Monthly Notices of the Royal Astronomical Society* **303** (Feb, 1999) 188.

[57] V. Springel, N. Yoshida and S. D. M. White, *Gadget: a code for collisionless and gasdynamical cosmological simulations, New Astronomy* **6** (Apr, 2001) 79.

[58] J. C. Helly, S. Cole, C. S. Frenk, C. M. Baugh, A. Benson and C. Lacey, *Galaxy formation using halo merger histories taken from n-body simulations, Monthly Notice of the Royal Astronomical Society* **338** (Feb, 2003) 903.

[59] S. Hatton, J. E. G. Devriendt, S. Ninin, F. R. Bouchet, B. Guiderdoni and D. Vibert, *Galics-i. a hybrid n-body/semi-analytic model of hierarchical galaxy formation, Monthly Notice of the Royal Astronomical Society* **343** (Jul, 2003) 75.

[60] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao et al., *Simulations of the formation, evolution and clustering of galaxies and quasars, Nature* **435** (Jun, 2005) 629.

[61] D. J. Croton, V. Springel, S. D. M. White, G. D. Lucia, C. S. Frenk, L. Gao et al., *The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies, Monthly Notices of the Royal Astronomical Society* **365** (Jan, 2006) 11.

[62] X. Yang, H. J. Mo and F. C. van den Bosch, *Constraining galaxy formation and cosmology with the conditional luminosity function of galaxies, Monthly Notice of the Royal Astronomical Society* **339** (Mar, 2003) 1057.

[63] J. A. Peacock and R. E. Smith, *Halo occupation numbers and galaxy bias, Monthly Notices of the Royal Astronomical Society* **318** (Nov, 2000) 1144.

[64] U. Seljak, *Analytic model for galaxy and dark matter clustering, Monthly Notices of the Royal Astronomical Society* **318** (Oct, 2000) 203.

[65] I. Zehavi, Z. Zheng, D. H. Weinberg, M. R. Blanton, N. A. Bahcall, A. A. Berlind et al., *Galaxy clustering in the completed sdss redshift survey: The dependence on color and luminosity, The Astrophysical Journal* **736** (Jul, 2011) 59.

[66] T. G. Brainerd, R. D. Blandford and I. Smail, *Weak gravitational lensing by galaxies, Astrophysical Journal v.466* **466** (Aug, 1996) 623.

[67] E. S. Sheldon, D. E. Johnston, J. A. Frieman, R. Scranton, T. A. McKay, A. J. Connolly et al., *The galaxy-mass correlation function measured from weak lensing in the sloan digital sky survey, The Astronomical Journal* **127** (May, 2004) 2544.

[68] A. Leauthaud, A. Finoguenov, J.-P. Kneib, J. E. Taylor, R. Massey, J. Rhodes et al., *A weak lensing study of x-ray groups in the cosmos survey: Form and evolution of the mass-luminosity relation, The Astrophysical Journal* **709** (Jan, 2010) 97.

[69] A. A. Berlind, J. Frieman, D. H. Weinberg, M. R. Blanton, M. S. Warren, K. Abazajian et al., *Percolation galaxy groups and clusters in the sdss redshift survey: Identification, catalogs, and the multiplicity function, The Astrophysical Journal Supplement Series* **167** (Nov, 2006) 1.

[70] X. Yang, H. J. Mo, F. C. van den Bosch, A. Pasquali, C. Li and M. Barden, *Galaxy groups in the sdss dr4. i. the catalog and basic properties, The Astrophysical Journal* **671** (Dec, 2007) 153.

[71] A. Leauthaud, J. Tinker, K. Bundy, P. S. Behroozi, R. Massey, J. Rhodes et al., *New constraints on the evolution of the stellar-to-dark matter connection: A combined*

analysis of galaxy-galaxy lensing, clustering, and stellar mass functions from z = 0.2 to z =1, The Astrophysical Journal **744** (Jan, 2012) 159.

[72] A. P. Hearin, A. R. Zentner, A. A. Berlind and J. A. Newman, *Sham beyond clustering: new tests of galaxy-halo abundance matching with galaxy groups, Monthly Notices of the Royal Astronomical Society* **433** (Jul, 2013) 659.

[73] R. M. Reddick, R. H. Wechsler, J. L. Tinker and P. S. Behroozi, *The connection between galaxies and dark matter structures in the local universe, The Astrophysical Journal* **771** (Jul, 2013) 30.

[74] J. L. Tinker, A. Leauthaud, K. Bundy, M. R. George, P. Behroozi, R. Massey et al., *Evolution of the stellar-to-dark matter relation: Separating star-forming and passive galaxies from z = 1 to 0, The Astrophysical Journal* **778** (Dec, 2013) 93.

[75] S. More, F. C. van den Bosch and M. Cacciato, *Satellite kinematics - i. a new method to constrain the halo mass-luminosity relation of central galaxies, Monthly Notices of the Royal Astronomical Society* **392** (Jan, 2009) 917.

[76] C. Firmani and V. Avila-Reese, *Galaxy downsizing evidenced by hybrid evolutionary tracks, The Astrophysical Journal* **723** (Nov, 2010) 755.

[77] X. Yang, H. J. Mo, F. C. van den Bosch, Y. Zhang and J. Han, *Evolution of the galaxy-dark matter connection and the assembly of galaxies in dark matter halos, The Astrophysical Journal* **752** (Jun, 2012) 41.

[78] P. S. Behroozi, R. H. Wechsler and C. Conroy, *The average star formation histories of galaxies in dark matter halos from z = 0-8, The Astrophysical Journal* **770** (Jun, 2013) 57.

[79] Z. Lu, H. J. Mo, Y. Lu, N. Katz, M. D. Weinberg, F. C. van den Bosch et al., *An empirical model for the star formation history in dark matter haloes, Monthly Notices of the Royal Astronomical Society* **439** (Feb, 2014) 1294.

[80] J. Brinchmann, S. Charlot, S. D. M. White, C. Tremonti, G. Kauffmann, T. Heckman et al., *The physical properties of star-forming galaxies in the low-redshift universe, Monthly Notices of the Royal Astronomical Society* **351** (Jul, 2004) 1151.

[81] K. G. Noeske, B. J. Weiner, S. M. Faber, C. Papovich, D. C. Koo, R. S. Somerville et al., *Star formation in aegis field galaxies since z=1.1: The dominance of gradually declining star formation, and the main sequence of star-forming galaxies, The Astrophysical Journal* **660** (May, 2007) L43.

[82] E. Daddi, M. Dickinson, G. Morrison, R. Chary, A. Cimatti, D. Elbaz et al., *Multiwavelength study of massive galaxies at z 2. i. star formation and galaxy growth, The Astrophysical Journal* **670** (Nov, 2007) 156.

[83] Y. Peng, S. J. Lilly, K. Kovač, M. Bolzonella, L. Pozzetti, A. Renzini et al., *Mass and environment as drivers of galaxy evolution in sdss and zcosmos and the origin of the schechter function, The Astrophysical Journal* **721** (Sep, 2010) 193.

[84] G. Rodighiero, E. Daddi, I. Baronchelli, A. Cimatti, A. Renzini, H. Aussel et al., *The lesser role of starbursts in star formation at z = 2, The Astrophysical Journal Letters* **739** (Oct, 2011) L40.

[85] Y. jie Peng, S. J. Lilly, A. Renzini and M. Carollo, *Mass and environment as drivers of galaxy evolution. ii. the quenching of satellite galaxies as the origin of environmental effects, The Astrophysical Journal* **757** (Sep, 2012) 4.

[86] N. Bouché, A. Dekel, R. Genzel, S. Genel, G. Cresci, N. M. F. Schreiber et al., *The impact of cold gas accretion above a mass floor on galaxy scaling relations, The Astrophysical Journal* **718** (Aug, 2010) 1001.

[87] R. Davé, K. Finlator and B. D. Oppenheimer, *An analytic model for the evolution of the stellar, gas, and metal content of galaxies, arXiv* **astro-ph.CO** (Aug, 2011) , [1108.0426v2].

[88] M. R. Krumholz and A. Dekel, *Metallicity-dependent quenching of star formation at high redshift in small galaxies, The Astrophysical Journal* **753** (Jul, 2012) 16.

[89] A. Dekel and N. Mandelker, *An analytic solution for the minimal bathtub toy model: challenges in the star-formation history of high-z galaxies, eprint arXiv* **1402** (Feb, 2014) 2283.

[90] P. Dayal, A. Ferrara and J. S. Dunlop, *The physics of the fundamental metallicity relation, Monthly Notices of the Royal Astronomical Society* **430** (Apr, 2013) 2891.

[91] S. J. Lilly, C. M. Carollo, A. Pipino, A. Renzini and Y. Peng, *Gas regulation of galaxies: The evolution of the cosmic specific star formation rate, the metallicity-mass-star-formation rate relation, and the stellar content of halos, The Astrophysical Journal* **772** (Aug, 2013) 119.

[92] S. Tacchella, M. Trenti and C. M. Carollo, *A physical model for the 0 < z < 8 redshift evolution of the galaxy ultraviolet luminosity and stellar mass functions, The Astrophysical Journal Letters* **768** (May, 2013) L37.

[93] A. Dekel, A. Zolotov, D. Tweed, M. Cacciato, D. Ceverino and J. R. Primack, *Toy models for galaxy formation versus simulations, Monthly Notices of the Royal Astronomical Society* **435** (Oct, 2013) 999.

[94] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw et al., *Seven-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological interpretation, The Astrophysical Journal Supplement* **192** (Feb, 2011) 18.

[95] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, *The statistics of peaks of gaussian random fields, Astrophysical Journal* **304** (May, 1986) 15.

[96] G. Bruzual and S. Charlot, *Stellar population synthesis at the resolution of 2003, Monthly Notices of the Royal Astronomical Society* **344** (Oct, 2003) 1000.

[97] F. Mannucci, G. Cresci, R. Maiolino, A. Marconi and A. Gnerucci, *A fundamental relation between mass, star formation rate and metallicity in local and high-redshift galaxies, Monthly Notices of the Royal Astronomical Society* **408** (Nov, 2010) 2115.

[98] H. J. Mo, S. Mao and S. D. M. White, *The formation of galactic discs, Monthly Notices of the Royal Astronomical Society* **295** (Apr, 1998) 319.

[99] R. Feldmann, *Lessons from cosmic history: the case for a linear star formation - h2 relation, Monthly Notices of the Royal Astronomical Society* **433** (Aug, 2013) 1910.

[100] J. Schaye, C. D. Vecchia, C. M. Booth, R. P. C. Wiersma, T. Theuns, M. R. Haas et al., *The physics driving the cosmic star formation history, Monthly Notices of the Royal Astronomical Society* **402** (Mar, 2010) 1536.

[101] V. Springel and L. Hernquist, *The history of star formation in a cold dark matter universe, Monthly Notice of the Royal Astronomical Society* **339** (Feb, 2003) 312.

[102] A. J. Benson, R. G. Bower, C. S. Frenk, C. G. Lacey, C. M. Baugh and S. Cole, *What shapes the luminosity function of galaxies?, The Astrophysical Journal* **599** (Dec, 2003) 38.

[103] G. D. Lucia, G. Kauffmann and S. D. M. White, *Chemical enrichment of the intracluster and intergalactic medium in a hierarchical galaxy formation model, Monthly Notices of the Royal Astronomical Society* **349** (Apr, 2004) 1101.

[104] F. Governato, B. Willman, L. Mayer, A. Brooks, G. Stinson, O. Valenzuela et al., *Forming disc galaxies in cdm simulations, Monthly Notices of the Royal Astronomical Society* **374** (Feb, 2007) 1479.

[105] B. D. Oppenheimer and R. Davé, *Mass, metal, and energy feedback in cosmological simulations, Monthly Notices of the Royal Astronomical Society* **387** (Jun, 2008) 577.

[106] C. Scannapieco, P. B. Tissera, S. D. M. White and V. Springel, *Effects of supernova feedback on the formation of galaxy discs, Monthly Notices of the Royal Astronomical Society* **389** (Sep, 2008) 1137.

[107] R. G. Bower, A. J. Benson and R. A. Crain, *What shapes the galaxy mass function? exploring the roles of supernova-driven winds and active galactic nuclei, Monthly Notices of the Royal Astronomical Society* **422** (Jun, 2012) 2816.

[108] F. Governato, L. Mayer, J. Wadsley, J. P. Gardner, B. Willman, E. Hayashi et al., *The formation of a realistic disk galaxy in -dominated cosmologies, The Astrophysical Journal* **607** (Jun, 2004) 688.

[109] R. G. Bower, A. J. Benson, R. Malbon, J. C. Helly, C. S. Frenk, C. M. Baugh et al., *Breaking the hierarchy of galaxy formation, Monthly Notices of the Royal Astronomical Society* **370** (Aug, 2006) 645.

[110] C. M. Booth and J. Schaye, *Cosmological simulations of the growth of supermassive black holes and feedback from active galactic nuclei: method and tests, Monthly Notices of the Royal Astronomical Society* **398** (Sep, 2009) 53.

[111] A. C. Fabian, K. A. Arnaud, M. W. Bautz and Y. Tawara, *Asca observations of cooling flows in clusters of galaxies, Astrophysical Journal* **436** (Nov, 1994) L63.

[112] H. Böhringer, K. Matsushita, E. Churazov, Y. Ikebe and Y. Chen, *The new emerging model for the structure of cooling cores in clusters of galaxies, Astronomy and Astrophysics* **382** (Feb, 2002) 804.

[113] W. Ishibashi and A. C. Fabian, *Active galactic nucleus feedback and triggering of star formation in galaxies, Monthly Notices of the Royal Astronomical Society* **427** (Dec, 2012) 2998.

[114] J. E. Gunn and J. R. Gott, *On the infall of matter into clusters of galaxies and some effects on their evolution, Astrophysical Journal* **176** (Aug, 1972) 1.

[115] I. G. McCarthy, A. Babul, R. G. Bower and M. L. Balogh, *Towards a holistic view of the heating and cooling of the intracluster medium, Monthly Notices of the Royal Astronomical Society* **386** (May, 2008) 1309.

[116] A. S. Font, R. G. Bower, I. G. McCarthy, A. J. Benson, C. S. Frenk, J. C. Helly et al., *The colours of satellite galaxies in groups and clusters, Monthly Notices of the Royal Astronomical Society* **389** (Oct, 2008) 1619.

[117] M. Prescott, I. K. Baldry, P. A. James, S. P. Bamford, J. Bland-Hawthorn, S. Brough et al., *Galaxy and mass assembly (gama): the red fraction and radial distribution of satellite galaxies, Monthly Notices of the Royal Astronomical Society* **417** (Oct, 2011) 1374.

[118] K. Kovač, S. J. Lilly, C. Knobel, T. J. Bschorr, Y. Peng, C. M. Carollo et al., *zcosmos 20k: satellite galaxies are the main drivers of environmental effects in the galaxy population at least to z ∼ 0.7, Monthly Notices of the Royal Astronomical Society* **438** (Feb, 2014) 717.

[119] C. Knobel, S. J. Lilly, K. Kovač, Y. Peng, T. J. Bschorr, C. M. Carollo et al., *The colors of central and satellite galaxies in zcosmos out to z = 0.8 and implications for quenching, The Astrophysical Journal* **769** (May, 2013) 24.

[120] R. I. Epstein, *Proto-galactic perturbations, Monthly Notices of the Royal Astronomical Society (ISSN 0035-8711)* **205** (Oct, 1983) 207.

[121] C. Lacey and S. Cole, *Merger rates in hierarchical models of galaxy formation, Monthly Notices of the Royal Astronomical Society (ISSN 0035-8711)* **262** (Jun, 1993) 627.

[122] M. Boylan-Kolchin, C.-P. Ma and E. Quataert, *Dynamical friction and galaxy merging time-scales, Monthly Notices of the Royal Astronomical Society* **383** (Jan, 2008) 93–101.

[123] A. R. Zentner, A. A. Berlind, J. S. Bullock, A. V. Kravtsov and R. H. Wechsler, *The physics of galaxy clustering. i. a model for subhalo populations, The Astrophysical Journal* **624** (May, 2005) 505.

[124] S. Borgani, K. Dolag, G. Murante, L.-M. Cheng, V. Springel, A. Diaferio et al., *Hot and cooled baryons in smoothed particle hydrodynamic simulations of galaxy clusters: physics and numerics, Monthly Notices of the Royal Astronomical Society* **367** (Apr, 2006) 1641.

[125] F. Jiang and F. C. van den Bosch, *Generating merger trees for dark matter haloes: a comparison of methods, Monthly Notices of the Royal Astronomical Society* (Mar, 2014) 482.

[126] N. Y. Gnedin, *Effect of reionization on structure formation in the universe, The Astrophysical Journal* **542** (Oct, 2000) 535.

[127] T. Okamoto, L. Gao and T. Theuns, *Mass loss of galaxies due to an ultraviolet background, Monthly Notices of the Royal Astronomical Society* **390** (Nov, 2008) 920.

[128] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. A. Prieto, D. An et al., *The seventh data release of the sloan digital sky survey, The Astrophysical Journal Supplement* **182** (Jun, 2009) 543.

[129] D. Elbaz, E. Daddi, D. L. Borgne, M. Dickinson, D. M. Alexander, R.-R. Chary et al., *The reversal of the star formation-density relation in the distant universe, Astronomy and Astrophysics* **468** (Jun, 2007) 33.

Bibliography

[130] M. Pannella, C. L. Carilli, E. Daddi, H. J. McCracken, F. N. Owen, A. Renzini et al., *Star formation and dust obscuration at z 2: Galaxies at the dawn of downsizing, The Astrophysical Journal Letters* **698** (Jun, 2009) L116.

[131] D. P. Stark, M. A. Schenker, R. Ellis, B. Robertson, R. McLure and J. Dunlop, *Keck spectroscopy of 3 < z < 7 faint lyman break galaxies: The importance of nebular emission in understanding the specific star formation rate and stellar mass density, The Astrophysical Journal* **763** (Feb, 2013) 129.

[132] M. T. Sargent, M. Béthermin, E. Daddi and D. Elbaz, *The contribution of starbursts and normal galaxies to infrared luminosity functions at z < 2, The Astrophysical Journal Letters* **747** (Mar, 2012) L31.

[133] E. Neistein and A. Dekel, *Merger rates of dark matter haloes, Monthly Notices of the Royal Astronomical Society* **388** (Aug, 2008) 1792.

[134] S. M. Weinmann, A. Pasquali, B. D. Oppenheimer, K. Finlator, J. T. Mendel, R. A. Crain et al., *A fundamental problem in our understanding of low-mass galaxy evolution, Monthly Notices of the Royal Astronomical Society* **426** (Nov, 2012) 2797.

[135] R. Davé, B. D. Oppenheimer and K. Finlator, *Galaxy evolution in cosmological simulations with outflows - i. stellar masses and star formation rates, Monthly Notices of the Royal Astronomical Society* **415** (Jul, 2011) 11.

[136] A. M. Hopkins and J. F. Beacom, *On the normalization of the cosmic star formation history, The Astrophysical Journal* **651** (Nov, 2006) 142.

[137] L. J. Tacconi, R. Genzel, R. Neri, P. Cox, M. C. Cooper, K. Shapiro et al., *High molecular gas fractions in normal massive star-forming galaxies in the young universe, Nature* **463** (Feb, 2010) 781.

[138] J. E. Geach, I. Smail, S. M. Moran, L. A. MacArthur, C. del P Lagos and A. C. Edge, *On the evolution of the molecular gas fraction of star-forming galaxies, The Astrophysical Journal Letters* **730** (Apr, 2011) L19.

210

[139] I. K. Baldry, K. Glazebrook and S. P. Driver, *On the galaxy stellar mass function, the mass-metallicity relation and the implied baryonic mass function, Monthly Notices of the Royal Astronomical Society* **388** (Aug, 2008) 945.

[140] L. Pozzetti, M. Bolzonella, E. Zucca, G. Zamorani, S. Lilly, A. Renzini et al., *zcosmos - 10k-bright spectroscopic sample. the bimodality in the galaxy stellar mass function: exploring its evolution with redshift, Astronomy and Astrophysics* **523** (Nov, 2010) 13.

[141] I. K. Baldry, S. P. Driver, J. Loveday, E. N. Taylor, L. S. Kelvin, J. Liske et al., *Galaxy and mass assembly (gama): the galaxy stellar mass function at z < 0.06, Monthly Notices of the Royal Astronomical Society* **421** (Mar, 2012) 621.

[142] S. J. Lilly, Y. Peng, A. Renzini and C. M. Carollo, *A simple continuity approach to galaxy evolution, Galaxy Mergers in an Evolving Universe* **477** (Oct, 2013) 11.

[143] P. S. Behroozi, R. H. Wechsler and C. Conroy, *On the lack of evolution in galaxy star formation efficiency, The Astrophysical Journal Letters* **762** (Jan, 2013) L31.

[144] A. Leauthaud, M. R. George, P. S. Behroozi, K. Bundy, J. Tinker, R. H. Wechsler et al., *The integrated stellar content of dark matter halos, The Astrophysical Journal* **746** (Feb, 2012) 95.

[145] J.-H. Woo, A. Schulze, D. Park, W.-R. Kang, S. C. Kim and D. A. Riechers, *Do quiescent and active galaxies have different m bh-* relations?, The Astrophysical Journal* **772** (Jul, 2013) 49.

[146] B. D. Oppenheimer, R. Davé, D. Kereš, M. Fardal, N. Katz, J. A. Kollmeier et al., *Feedback and recycled wind accretion: assembling the z = 0 galaxy mass function, Monthly Notices of the Royal Astronomical Society* **406** (Aug, 2010) 2325.

[147] B. M. B. Henriques, S. D. M. White, P. A. Thomas, R. E. Angulo, Q. Guo, G. Lemson et al., *Simulations of the galaxy population constrained by observations from z = 3 to the present day: implications for galactic winds and the fate of their ejecta, Monthly Notices of the Royal Astronomical Society* **431** (Jun, 2013) 3373.

[148] R. Davé, N. Katz, B. D. Oppenheimer, J. A. Kollmeier and D. H. Weinberg, *The neutral hydrogen content of galaxies in cosmological hydrodynamic simulations, arXiv* **astro-ph.CO** (Feb, 2013) , [1302.3631v1].

[149] K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, S. Bailey et al., *The baryon oscillation spectroscopic survey of sdss-iii, The Astronomical Journal* **145** (Jan, 2013) 10.

[150] K. T. Inoue, R. Takahashi, T. Takahashi and T. Ishiyama, *Constraints on warm dark matter from weak lensing in anomalous quadruple lenses, arXiv* **astro-ph.CO** (Sep, 2014) , [1409.1326v2].

[151] R. B. Metcalf and P. Madau, *Compound gravitational lensing as a probe of dark matter substructure within galaxy halos, The Astrophysical Journal* **563** (Dec, 2001) 9.

[152] N. Dalal and C. S. Kochanek, *Direct detection of cold dark matter substructure, The Astrophysical Journal* **572** (Jun, 2002) 25.

[153] J. Yoo, C. S. Kochanek, E. E. Falco and B. A. McLeod, *Halo structures of gravitational lens galaxies, The Astrophysical Journal* **642** (May, 2006) 22.

[154] C. R. Keeton and L. A. Moustakas, *A new channel for detecting dark matter substructure in galaxies: Gravitational lens time delays, The Astrophysical Journal* **699** (Jul, 2009) 1720.

[155] L. A. Moustakas, K. Abazajian, A. Benson, A. S. Bolton, J. S. Bullock, J. Chen et al., *Strong gravitational lensing probes of the particle nature of dark matter, Astro2010: The Astronomy and Astrophysics Decadal Survey* **2010** (Jan, 2009) 214.

[156] T. Treu, *Strong lensing by galaxies, Annual Review of Astronomy and Astrophysics* **48** (Sep, 2010) 87.

[157] L. V. E. Koopmans, *Gravitational imaging of cold dark matter substructures, Monthly Notices of the Royal Astronomical Society* **363** (Nov, 2005) 1136.

[158] S. Vegetti and L. V. E. Koopmans, *Bayesian strong gravitational-lens modelling on*

adaptive grids: objective detection of mass substructure in galaxies, Monthly Notices of the Royal Astronomical Society **392** (Jan, 2009) 945.

[159] S. Vegetti, L. V. E. Koopmans, A. Bolton, T. Treu and R. Gavazzi, *Detection of a dark substructure through gravitational imaging, Monthly Notices of the Royal Astronomical Society* **408** (Nov, 2010) 1969.

[160] S. Vegetti, D. J. Lagattuta, J. P. McKean, M. W. Auger, C. D. Fassnacht and L. V. E. Koopmans, *Gravitational detection of a low-mass dark satellite galaxy at cosmological distance, Nature* **481** (Jan, 2012) 341.

[161] R. B. Metcalf and H. Zhao, *Flux ratios as a probe of dark substructures in quadruple-image gravitational lenses, The Astrophysical Journal* **567** (Mar, 2002) L5.

[162] C. S. Kochanek and N. Dalal, *Tests for substructure in gravitational lenses, The Astrophysical Journal* **610** (Jul, 2004) 69.

[163] A. Amara, R. B. Metcalf, T. J. Cox and J. P. Ostriker, *Simulations of strong gravitational lensing with substructure, Monthly Notices of the Royal Astronomical Society* **367** (Apr, 2006) 1367.

[164] R. B. Metcalf and A. Amara, *Small-scale structures of dark matter and flux anomalies in quasar gravitational lenses, Monthly Notices of the Royal Astronomical Society* **419** (Feb, 2012) 3414.

[165] D. Xu, D. Sluse, L. Gao, J. Wang, C. Frenk, S. Mao et al., *How well can cold dark matter substructures account for the observed radio flux-ratio anomalies, Monthly Notices of the Royal Astronomical Society* **447** (Mar, 2015) 3189.

[166] R. Takahashi and K. T. Inoue, *Weak lensing by intergalactic ministructures in quadruple lens systems: simulation and detection, Monthly Notices of the Royal Astronomical Society* **440** (May, 2014) 870.

[167] M. Oguri and P. J. Marshall, *Gravitationally lensed quasars and supernovae in future wide-field optical imaging surveys, Monthly Notices of the Royal Astronomical Society* **405** (Jul, 2010) 2579.

Bibliography

[168] C. S. Kochanek, *The implications of lenses for galaxy structure, Astrophysical Journal* **373** (Jun, 1991) 354.

[169] J.-P. Kneib, R. S. Ellis, I. Smail, W. J. Couch and R. M. Sharples, *Hubble space telescope observations of the lensing cluster abell 2218, Astrophysical Journal v.471* **471** (Nov, 1996) 643.

[170] C. R. Keeton, *Computational methods for gravitational lensing, eprint arXiv* (Feb, 2001) 2340.

[171] E. Jullo, J.-P. Kneib, M. Limousin, Á. Elíasdóttir, P. J. Marshall and T. Verdugo, *A bayesian approach to strong lensing modelling of galaxy clusters, New Journal of Physics* **9** (Dec, 2007) 447.

[172] S. Wallington, C. S. Kochanek and R. Narayan, *Lensmem: A gravitational lens inversion algorithm using the maximum entropy method, Astrophysical Journal v.465* **465** (Jul, 1996) 64.

[173] S. J. Warren and S. Dye, *Semilinear gravitational lens inversion, The Astrophysical Journal* **590** (Jun, 2003) 673.

[174] T. Treu and L. V. E. Koopmans, *Massive dark matter halos and evolution of early-type galaxies to z 1, The Astrophysical Journal* **611** (Aug, 2004) 739.

[175] S. Dye and S. J. Warren, *Decomposition of the visible and dark matter in the einstein ring 0047-2808 by semilinear inversion, The Astrophysical Journal* **623** (Apr, 2005) 31.

[176] B. J. Brewer and G. F. Lewis, *Strong gravitational lens inversion: A bayesian approach, The Astrophysical Journal* **637** (Feb, 2006) 608.

[177] S. H. Suyu, P. J. Marshall, M. P. Hobson and R. D. Blandford, *A bayesian analysis of regularized source inversions in gravitational lensing, Monthly Notices of the Royal Astronomical Society* **371** (Sep, 2006) 983.

[178] R. B. Wayth and R. L. Webster, *Lensview: software for modelling resolved gravitational lens images, Monthly Notices of the Royal Astronomical Society* **372** (Nov, 2006) 1187.

214

[179] S. H. Suyu, P. J. Marshall, M. W. Auger, S. Hilbert, R. D. Blandford, L. V. E. Koopmans et al., *Dissecting the gravitational lens b1608+656. ii. precision measurements of the hubble constant, spatial curvature, and the dark energy equation of state, The Astrophysical Journal* **711** (Mar, 2010) 201.

[180] S. H. Suyu, M. W. Auger, S. Hilbert, P. J. Marshall, M. Tewes, T. Treu et al., *Two accurate time-delay distances from strong lensing: Implications for cosmology, The Astrophysical Journal* **766** (Apr, 2013) 70.

[181] A. S. Tagore and C. R. Keeton, *Statistical and systematic uncertainties in pixel-based source reconstruction algorithms for gravitational lensing, Monthly Notices of the Royal Astronomical Society* **445** (Nov, 2014) 694.

[182] R. Blandford, G. Surpi and T. Kundić, *Modeling galaxy lenses, Gravitational Lensing: Recent Progress and Future Goals* **237** (Jan, 2001) 65.

[183] P. Saha and L. L. R. Williams, *A portable modeler of lensed quasars, arXiv* **astro-ph** (Feb, 2004) , [astro-ph/0402135v1].

[184] M. Bradač, T. Erben, P. Schneider, H. Hildebrandt, M. Lombardi, M. Schirmer et al., *Strong and weak lensing united, Astronomy and Astrophysics* **437** (Jul, 2005) 49.

[185] P. Saha, J. Coles, A. V. Macciò and L. L. R. Williams, *The hubble time inferred from 10 time delay lenses, The Astrophysical Journal* **650** (Oct, 2006) L17.

[186] S. H. Suyu and R. D. Blandford, *The anatomy of a quadruply imaged gravitational lens system, Monthly Notices of the Royal Astronomical Society* **366** (Feb, 2006) 39.

[187] M. J. Jee, H. C. Ford, G. D. Illingworth, R. L. White, T. J. Broadhurst, D. A. Coe et al., *Discovery of a ringlike dark matter structure in the core of the galaxy cluster cl 0024+17, The Astrophysical Journal* **661** (Jun, 2007) 728.

[188] S. H. Suyu, P. J. Marshall, R. D. Blandford, C. D. Fassnacht, L. V. E. Koopmans, J. P. McKean et al., *Dissecting the gravitational lens b1608+656. i. lens potential reconstruction, The Astrophysical Journal* **691** (Jan, 2009) 277.

[189] J. P. Coles, J. I. Read and P. Saha, *Gravitational lens recovery with glass: measuring the mass profile and shape of a lens, Monthly Notices of the Royal Astronomical Society* **445** (Dec, 2014) 2181.

[190] J. Merten, *Mesh-free free-form lensing i: Methodology and application to mass reconstruction, eprint arXiv* **1412** (Dec, 2014) 5186.

[191] R. B. Metcalf and M. Petkova, *Glamer - i. a code for gravitational lensing simulations with adaptive mesh refinement, Monthly Notices of the Royal Astronomical Society* **445** (Dec, 2014) 1942.

[192] A. Refregier, *Shapelets - i. a method for image analysis, Monthly Notice of the Royal Astronomical Society* **338** (Jan, 2003) 35.

[193] A. Refregier and D. Bacon, *Shapelets - ii. a method for weak lensing measurements, Monthly Notice of the Royal Astronomical Society* **338** (Jan, 2003) 48.

[194] R. Massey and A. Refregier, *Polar shapelets, Monthly Notices of the Royal Astronomical Society* **363** (Oct, 2005) 197.

[195] A. S. Tagore and N. Jackson, *On the use of shapelets in modelling resolved, gravitationally lensed images, arXiv* **astro-ph.CO** (May, 2015) , [1505.00198v1].

[196] C. R. Keeton, *On modeling galaxy-scale strong lens systems, General Relativity and Gravitation* **42** (Sep, 2010) 2151.

[197] J.-P. Kneib and P. Natarajan, *Cluster lenses, The Astronomy and Astrophysics Review* **19** (Nov, 2011) 47.

[198] P. Saha and L. L. R. Williams, *A portable modeler of lensed quasars, The Astronomical Journal* **127** (May, 2004) 2604.

[199] R. Barkana, *Fast calculation of a family of elliptical mass gravitational lens models, Astrophysical Journal v.502* **502** (Aug, 1998) 531.

[200] J. F. Navarro, C. S. Frenk and S. D. M. White, *A universal density profile from hierarchical clustering, The Astrophysical Journal* **490** (Dec, 1997) 493.

[201] J. Sersic, *Atlas de galaxias australes, . . .* (Jan, 1968) .

[202] J. Akeret, S. Seehars, A. Amara, A. Refregier and A. Csillaghy, *Cosmohammer: Cosmological parameter estimation with the mcmc hammer, Astronomy and Computing* **2** (Aug, 2013) 27.

[203] J. Goodman and J. Weare, *Ensemble samplers with affine invariance, Commun. Appl. Math. Comput. Sci.* **5** (2010) 65–80.

[204] D. Foreman-Mackey, A. Conley, W. M. Farr, D. W. Hogg, D. Long, P. Marshall et al., *emcee: The mcmc hammer, Astrophysics Source Code Library* (Mar, 2013) 1303.002.

[205] J. Kennedy, J. Kennedy and R. Eberhart, *Swarm intelligence, books.google.com* (Jan, 2001) .

[206] D. Sluse, J. Surdej, J.-F. Claeskens, D. Hutsemékers, C. Jean, F. Courbin et al., *A quadruply imaged quasar with an optical einstein ring candidate: 1rxs j113155.4-123155, Astronomy and Astrophysics* **406** (Jul, 2003) L43.

[207] J.-F. Claeskens, D. Sluse, P. Riaud and J. Surdej, *Multi wavelength study of the gravitational lens system rxs j1131-1231. ii. lens model and source reconstruction, Astronomy and Astrophysics* **451** (Jun, 2006) 865.

[208] B. J. Brewer and G. F. Lewis, *Unlensing hst observations of the einstein ring 1rxs j1131-1231: a bayesian analysis, Monthly Notices of the Royal Astronomical Society* **390** (Oct, 2008) 39.

[209] J. Bansal, P. Singh and M. Saraswat. . . , *Inertia weight strategies in particle swarm optimization, . . . (NaBIC)* (Jan, 2011) .

[210] K. Zielinski and R. Laur, *Stopping criteria for differential evolution in constrained single-objective optimization, Advances in differential evolution* (Jan, 2008) .

[211] M. Bartelmann and P. Schneider, *Weak gravitational lensing, Physics Reports* **340** (Jan, 2001) 291.

[212] A. Refregier, *Weak gravitational lensing by large-scale structure, Annual Review of Astronomy &Astrophysics* **41** (Jan, 2003) 645.

Bibliography

[213] S. Refsdal, *On the possibility of determining hubble's parameter and the masses of galaxies from the gravitational lens effect, Monthly Notices of the Royal Astronomical Society* **128** (Jan, 1964) 307.

[214] C. S. Kochanek, *Is there a cosmological constant?, Astrophysical Journal v.466* **466** (Aug, 1996) 638.

[215] T. Treu and P. J. Marshall, *Time delay cosmography, eprint arXiv* **1605** (May, 2016) arXiv:1605.05333.

[216] D. D. Xu, S. Mao, A. P. Cooper, L. Gao, C. S. Frenk, R. E. Angulo et al., *On the effects of line-of-sight structures on lensing flux-ratio anomalies in a cdm universe, Monthly Notices of the Royal Astronomical Society* **421** (Apr, 2012) 2553.

[217] P. L. Schechter, C. D. Bailyn, R. Barr, R. Barvainis, C. M. Becker, G. M. Bernstein et al., *The quadruple gravitational lens pg 1115+080: Time delays and models, The Astrophysical Journal* **475** (Feb, 1997) L85.

[218] M. Oguri, *Gravitational lens time delays: A statistical assessment of lens model dependences and implications for the global hubble constant, The Astrophysical Journal* **660** (May, 2007) 1.

[219] R. Bar-Kana, *Effect of large-scale structure on multiply imaged sources, Astrophysical Journal v.468* **468** (Sep, 1996) 17.

[220] I. Momcheva, K. Williams, C. Keeton and A. Zabludoff, *A spectroscopic study of the environments of gravitational lens galaxies, The Astrophysical Journal* **641** (Apr, 2006) 169.

[221] K. C. Wong, C. R. Keeton, K. A. Williams, I. G. Momcheva and A. I. Zabludoff, *The effect of environment on shear in strong gravitational lenses, The Astrophysical Journal* **726** (Jan, 2011) 84.

[222] M. Jaroszynski and Z. Kostrzewa-Rutkowska, *The influence of the matter along the line of sight and in the lens environment on the strong gravitational lensing, Monthly Notices of the Royal Astronomical Society* **439** (Apr, 2014) 2432.

[223] C. McCully, C. R. Keeton, K. C. Wong and A. I. Zabludoff, *Quantifying environmental and line-of-sight effects in models of strong gravitational lens systems, arXiv* **astro-ph.CO** (Jan, 2016) , [1601.05417v1].

[224] P. Schneider, *The cosmological lens equation and the equivalent single-plane gravitational lens, Monthly Notices of the Royal Astronomical Society* **292** (Dec, 1997) 673.

[225] C. R. Keeton, C. S. Kochanek and U. Seljak, *Shear and ellipticity in gravitational lenses, The Astrophysical Journal* **482** (Jun, 1997) 604.

[226] C. R. Keeton and A. I. Zabludoff, *The importance of lens galaxy environments, The Astrophysical Journal* **612** (Sep, 2004) 660.

[227] C. D. Fassnacht, R. R. Gal, L. M. Lubin, J. P. McKean, G. K. Squires and A. C. S. Readhead, *Mass along the line of sight to the gravitational lens b1608+656: Galaxy groups and implications for h0, The Astrophysical Journal* **642** (May, 2006) 30.

[228] T. Anguita, C. Faure, J.-P. Kneib, J. Wambsganss, C. Knobel, A. M. Koekemoer et al., *Cosmos 5921+0638: characterization and analysis of a new strong gravitationally lensed agn, Astronomy and Astrophysics* **507** (Nov, 2009) 35.

[229] C. D. Fassnacht, L. V. E. Koopmans and K. C. Wong, *Galaxy number counts and implications for strong lensing, Monthly Notices of the Royal Astronomical Society* **410** (Feb, 2011) 2167.

[230] R. Nakajima, G. M. Bernstein, R. Fadely, C. R. Keeton and T. Schrabback, *Improved constraints on the gravitational lens q0957+561. i. weak lensing, The Astrophysical Journal* **697** (Jun, 2009) 1793.

[231] Z. S. Greene, S. H. Suyu, T. Treu, S. Hilbert, M. W. Auger, T. E. Collett et al., *Improving the precision of time-delay cosmography with observations of galaxies along the line of sight, The Astrophysical Journal* **768** (May, 2013) 39.

[232] T. E. Collett, P. J. Marshall, M. W. Auger, S. Hilbert, S. H. Suyu, Z. Greene et al., *Reconstructing the lensing mass in the universe from photometric catalogue data, Monthly Notices of the Royal Astronomical Society* **432** (Jun, 2013) 679.

[233] H. I. Levine and A. O. Petters, *New caustic singularities in multiple lens plane gravitational lensing, Astronomy and Astrophysics* **272** (May, 1993) L17.

[234] R. Kayser and T. Schramm, *New caustic singularities in multiple lens plane gravitational lensing are not stable, Astronomy and Astrophysics (ISSN 0004-6361)* **278** (Oct, 1993) L13.

[235] A. O. Petters, *Multiplane gravitational lensing. i. morse theory and image counting., Journal of Mathematical Physics* **36** (Aug, 1995) 4263.

[236] A. O. Petters, *Multiplane gravitational lensing. ii. global geometry of caustics., Journal of Mathematical Physics* **36** (Aug, 1995) 4276.

[237] A. O. Petters, H. Levine and J. Wambsganss, *Singularity theory and gravitational lensing, Singularity theory and gravitational lensing / Arlie O. Petters* (Jan, 2001) .

[238] R. Blandford and R. Narayan, *Fermat's principle, caustics, and the classification of gravitational lens images, Astrophysical Journal* **310** (Nov, 1986) 568.

[239] I. Kovner, *The thick gravitational lens - a lens composed of many elements at different distances, Astrophysical Journal* **316** (May, 1987) 52.

[240] P. Schneider, J. Ehlers and E. E. Falco, *Gravitational lenses, Gravitational Lenses* (Jan, 1992) 112.

[241] R. Narayan and M. Bartelmann, *Lectures on gravitational lensing, eprint arXiv* (Jun, 1996) arXiv:astro–ph/9606001.

[242] E. E. Falco, M. V. Gorenstein and I. I. Shapiro, *On model-dependent bounds on h(0) from gravitational images application of q0957 + 561a,b, Astrophysical Journal* **289** (Feb, 1985) L1.

[243] P. Schneider and C. Seitz, *Steps towards nonlinear cluster inversion through gravitational distortions. 1: Basic considerations and circular clusters, Astronomy and Astrophysics (ISSN 0004-6361)* **294** (Feb, 1995) 411.

[244] P. Saha, *Lensing degeneracies revisited, The Astronomical Journal* **120** (Oct, 2000) 1654.

[245] O. Wucknitz, *Degeneracies and scaling relations in general power-law models for gravitational lenses, Monthly Notices of the Royal Astronomical Society* **332** (Jun, 2002) 951.

[246] C. Faure, T. Anguita, D. Alloin, K. Bundy, A. Finoguenov, A. Leauthaud et al., *On the evolution of environmental and mass properties of strong lens galaxies in cosmos, Astronomy & Astrophysics* **529** (May, 2011) A72.

[247] I. Zehavi, M. R. Blanton, J. A. Frieman, D. H. Weinberg, H. J. Mo, M. A. Strauss et al., *Galaxy clustering in early sloan digital sky survey redshift data, The Astrophysical Journal* **571** (May, 2002) 172.

[248] A. V. Kravtsov, A. A. Berlind, R. H. Wechsler, A. A. Klypin, S. Gottlöber, B. Allgood et al., *The dark side of the halo occupation distribution, The Astrophysical Journal* **609** (Jul, 2004) 35.

[249] Q. Guo, S. White, C. Li and M. Boylan-Kolchin, *How do galaxies populate dark matter haloes?, Monthly Notices of the Royal Astronomical Society* **404** (May, 2010) 1111.

[250] P. S. Behroozi, C. Conroy and R. H. Wechsler, *A comprehensive analysis of uncertainties affecting the stellar mass-halo mass relation for $0 < z < 4$, The Astrophysical Journal* **717** (Jul, 2010) 379.

[251] J. F. Navarro, C. S. Frenk and S. D. M. White, *The structure of cold dark matter halos, Astrophysical Journal v.462* **462** (May, 1996) 563.

[252] A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk et al., *The statistics of cdm halo concentrations, Monthly Notices of the Royal Astronomical Society* **381** (Nov, 2007) 1450.

[253] A. D. Ludlow, J. F. Navarro, R. E. Angulo, M. Boylan-Kolchin, V. Springel, C. Frenk et al., *The mass-concentration-redshift relation of cold dark matter haloes, Monthly Notices of the Royal Astronomical Society* **441** (Jun, 2014) 378.

[254] M. Bartelmann, *Arcs from a universal dark-matter halo profile., Astronomy and Astrophysics* **313** (Sep, 1996) 697.

Bibliography

[255] C. O. Wright and T. G. Brainerd, *Gravitational lensing by nfw halos, The Astrophysical Journal* **534** (May, 2000) 34.

[256] C. Faure, J.-P. Kneib, G. Covone, L. Tasca, A. Leauthaud, P. Capak et al., *First catalog of strong lens candidates in the cosmos field, The Astrophysical Journal Supplement Series* **176** (May, 2008) 19–38.

[257] O. Ilbert, P. Capak, M. Salvato, H. Aussel, H. J. McCracken, D. B. Sanders et al., *Cosmos photometric redshifts with 30-bands for 2-deg2, The Astrophysical Journal* **690** (Jan, 2009) 1236.

[258] N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis et al., *The cosmic evolution survey (cosmos): Overview, The Astrophysical Journal Supplement Series* **172** (Sep, 2007) 1.

[259] R. Jedrzejewski, W. Hack, C. Hanley, I. Busko and A. M. Koekemoer, *Multidrizzle: Automated image combination and cosmic-ray identification software, Astronomical Data Analysis Software and Systems XIV ASP Conference Series* **347** (Dec, 2005) 129.

[260] A. M. Koekemoer, H. Aussel, D. Calzetti, P. Capak, M. Giavalisco, J.-P. Kneib et al., *The cosmos survey: Hubble space telescope advanced camera for surveys observations and data processing, The Astrophysical Journal Supplement Series* **172** (Sep, 2007) 196.

[261] R. Massey, C. Stoughton, A. Leauthaud, J. Rhodes, A. Koekemoer, R. Ellis et al., *Pixel-based correction for charge transfer inefficiency in the hubble space telescope advanced camera for surveys, Monthly Notices of the Royal Astronomical Society* **401** (Jan, 2010) 371.

[262] N. Scoville, R. G. Abraham, H. Aussel, J. E. Barnes, A. Benson, A. W. Blain et al., *Cosmos: Hubble space telescope observations, The Astrophysical Journal Supplement Series* **172** (Sep, 2007) 38.

[263] P. Capak, H. Aussel, M. Ajiki, H. J. McCracken, B. Mobasher, N. Scoville et al., *The first release cosmos optical and near-ir data and catalog, The Astrophysical Journal Supplement Series* **172** (Sep, 2007) 99.

[264] E. F. Bell and R. S. de Jong, *Stellar mass-to-light ratios and the tully-fisher relation, The Astrophysical Journal* **550** (Mar, 2001) 212.

[265] M. Bernardi, R. K. Sheth, J. Annis, S. Burles, D. J. Eisenstein, D. P. Finkbeiner et al., *Early-type galaxies in the sloan digital sky survey. iii. the fundamental plane, The Astronomical Journal* **125** (Apr, 2003) 1866.

[266] E. F. Bell, D. H. McIntosh, N. Katz and M. D. Weinberg, *The optical and near-infrared properties of galaxies. i. luminosity and stellar mass functions, The Astrophysical Journal Supplement Series* **149** (Dec, 2003) 289.

[267] M. R. George, A. Leauthaud, K. Bundy, A. Finoguenov, J. Tinker, Y.-T. Lin et al., *Galaxies in x-ray groups. i. robust membership assignment and the impact of group environments on quenching, The Astrophysical Journal* **742** (Dec, 2011) 125.

[268] R. D. Blandford and R. Narayan, *Cosmological applications of gravitational lensing, In: Annual review of astronomy and astrophysics. Vol. 30 (A93-25826 09-90)* **30** (Jan, 1992) 311.

[269] F. Courbin, A. Eigenbrod, C. Vuissoz, G. Meylan and P. Magain, *Cosmograil: the cosmological monitoring of gravitational lenses, Gravitational Lensing Impact on Cosmology* **225** (Jun, 2005) 297.

[270] A. Eigenbrod, F. Courbin, C. Vuissoz, G. Meylan, P. Saha and S. Dye, *Cosmograil: The cosmological monitoring of gravitational lenses. i. how to sample the light curves of gravitationally lensed quasars to measure accurate time delays, Astronomy and Astrophysics* **436** (Jun, 2005) 25.

[271] F. Courbin, V. Chantry, Y. Revaz, D. Sluse, C. Faure, M. Tewes et al., *Cosmograil: the cosmological monitoring of gravitational lenses. ix. time delays, lens dynamics and baryonic fraction in he 0435-1223, Astronomy & Astrophysics* **536** (Dec, 2011) A53.

[272] M. Tewes, F. Courbin, G. Meylan, C. S. Kochanek, E. Eulaers, N. Cantale et al., *Cosmograil: Measuring time delays of gravitationally lensed quasars to constrain cosmology, The Messenger* **150** (Dec, 2012) 49.

Bibliography

[273] S. R. Kumar, M. Tewes, C. S. Stalin, F. Courbin, I. Asfandiyarov, G. Meylan et al., *Cosmograil: the cosmological monitoring of gravitational lenses. xiv. time delay of the doubly lensed quasar sdss j1001+5027, Astronomy & Astrophysics* **557** (Sep, 2013) A44.

[274] M. Tewes, F. Courbin and G. Meylan, *Cosmograil: the cosmological monitoring of gravitational lenses. xi. techniques for time delay measurement in presence of microlensing, Astronomy & Astrophysics* **553** (May, 2013) 120.

[275] L. V. E. Koopmans, T. Treu, C. D. Fassnacht, R. D. Blandford and G. Surpi, *The hubble constant from the gravitational lens b1608+656, The Astrophysical Journal* **599** (Dec, 2003) 70.

[276] O. Wucknitz, A. D. Biggs and I. W. A. Browne, *Models for the lens and source of b0218+357: a lensclean approach to determine h0, Monthly Notices of the Royal Astronomical Society* **349** (Mar, 2004) 14.

[277] T. York, N. Jackson, I. W. A. Browne, O. Wucknitz and J. E. Skelton, *The hubble constant from the gravitational lens class b0218+357 using the advanced camera for surveys, Monthly Notices of the Royal Astronomical Society* **357** (Feb, 2005) 124.

[278] P. Jakobsson, J. Hjorth, I. Burud, G. Letawe, C. Lidman and F. Courbin, *An optical time delay for the double gravitational lens system fbq 0951+2635, Astronomy and Astrophysics* **431** (Feb, 2005) 103.

[279] C. Vuissoz, F. Courbin, D. Sluse, G. Meylan, M. Ibrahimov, I. Asfandiyarov et al., *Cosmograil: the cosmological monitoring of gravitational lenses. v. the time delay in sdss j1650+4251, Astronomy and Astrophysics* **464** (Mar, 2007) 845.

[280] D. Paraficz, J. Hjorth and Á. Elíasdóttir, *Results of optical monitoring of 5 sdss double qsos with the nordic optical telescope, Astronomy and Astrophysics* **499** (May, 2009) 395.

[281] R. Fadely, C. R. Keeton, R. Nakajima and G. M. Bernstein, *Improved constraints on the gravitational lens q0957+561. ii. strong lensing, The Astrophysical Journal* **711** (Mar, 2010) 246.

[282] S. H. Suyu, T. Treu, S. Hilbert, A. Sonnenfeld, M. W. Auger, R. D. Blandford et al., *Cosmology from gravitational lens time delays and planck data, The Astrophysical Journal Letters* **788** (Jun, 2014) L35.

[283] J. Coles, *A new estimate of the hubble time with improved modeling of gravitational lenses, The Astrophysical Journal* **679** (May, 2008) 17.

[284] J. Liesenborgs and S. D. Rijcke, *Lensing degeneracies and mass substructure, Monthly Notices of the Royal Astronomical Society* **425** (Sep, 2012) 1772.

[285] P. Schneider and D. Sluse, *Source-position transformation: an approximate invariance in strong gravitational lensing, Astronomy & Astrophysics* **564** (Apr, 2014) A103.

[286] P. Schneider and D. Sluse, *Mass-sheet degeneracy, power-law models and external convergence: Impact on the determination of the hubble constant from gravitational lensing, Astronomy & Astrophysics* **559** (Nov, 2013) A37.

[287] G. C. F. Chen, S. H. Suyu, K. C. Wong, C. D. Fassnacht, T. Chiueh, A. Halkola et al., *Sharp - iii: First use of adaptive optics imaging to constrain cosmology with gravitational lens time delays, arXiv* **astro-ph.CO** (Jan, 2016) , [1601.01321v1].

[288] M. Tewes, F. Courbin, G. Meylan, C. S. Kochanek, E. Eulaers, N. Cantale et al., *Cosmograil: the cosmological monitoring of gravitational lenses. xiii. time delays and 9-yr optical monitoring of the lensed quasar rx j1131-1231, Astronomy & Astrophysics* **556** (Aug, 2013) 22.

[289] J. L. Sersic, *Atlas de galaxias australes, Cordoba* (Jan, 1968) .

[290] L. Hernquist, *An analytical model for spherical galaxies and bulges, Astrophysical Journal* **356** (Jun, 1990) 359.

[291] P. Saha and L. L. R. Williams, *Gravitational lensing model degeneracies: Is steepness all-important?, The Astrophysical Journal* **653** (Dec, 2006) 936.

[292] S. M. Faber, C. N. A. Willmer, C. Wolf, D. C. Koo, B. J. Weiner, J. A. Newman et al., *Galaxy luminosity functions to z 1 from deep2 and combo-17: Implications for red galaxy formation, The Astrophysical Journal* **665** (Aug, 2007) 265.

[293] L. V. E. Koopmans, A. Bolton, T. Treu, O. Czoske, M. W. Auger, M. Barnabè et al., *The structure and dynamics of massive early-type galaxies: On homology, isothermality, and isotropy inside one effective radius, The Astrophysical Journal Letters* **703** (Sep, 2009) L51.

[294] M. Barnabè, O. Czoske, L. V. E. Koopmans, T. Treu and A. S. Bolton, *Two-dimensional kinematics of slacs lenses - iii. mass structure and dynamics of early-type lens galaxies beyond z 0.1, Monthly Notices of the Royal Astronomical Society* **415** (Aug, 2011) 2215.

[295] M. Barnabè, A. A. Dutton, P. J. Marshall, M. W. Auger, B. J. Brewer, T. Treu et al., *The swells survey - iv. precision measurements of the stellar and dark matter distributions in a spiral lens galaxy, Monthly Notices of the Royal Astronomical Society* **423** (Jun, 2012) 1073.

[296] A. Lewis and S. Bridle, *Cosmological parameters from cmb and other data: A monte carlo approach, Physical Review D* **66** (Nov, 2002) 103511.

[297] I. Jee, E. Komatsu and S. H. Suyu, *Measuring angular diameter distances of strong gravitational lenses, Journal of Cosmology and Astroparticle Physics* **11** (Nov, 2015) 033.

[298] I. Jee, E. Komatsu, S. H. Suyu and D. Huterer, *Time-delay cosmography: Increased leverage with angular diameter distances, eprint arXiv* **1509** (Sep, 2015) 3310.

[299] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley et al., *Nine-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological parameter results, The Astrophysical Journal Supplement* **208** (Oct, 2013) 19.

[300] J. Bergé, L. Gamper, A. Réfrégier and A. Amara, *An ultra fast image generator (ufig) for wide-field astronomy, Astronomy and Computing* **1** (Feb, 2013) 23.

[301] C. S. Kochanek, *What do gravitational lens time delays measure?, The Astrophysical Journal* **578** (Oct, 2002) 25.

[302] Y. D. Hezaveh, N. Dalal, D. P. Marrone, Y.-Y. Mao, W. Morningstar, D. Wen et al., *Detection of lensing substructure using alma observations of the dusty galaxy sdp.81, The Astrophysical Journal* **823** (May, 2016) 37.

[303] T. Kacprzak, J. Zuntz, B. Rowe, S. Bridle, A. Refregier, A. Amara et al., *Measurement and calibration of noise bias in weak lensing galaxy shape estimation, Monthly Notices of the Royal Astronomical Society* **427** (Dec, 2012) 2711.

[304] A. Mesinger, S. Furlanetto and R. Cen, *21cmfast: a fast, seminumerical simulation of the high-redshift 21-cm signal, Monthly Notices of the Royal Astronomical Society* **411** (Feb, 2011) 955.

[305] M. Reinecke, K. Dolag, R. Hell, M. Bartelmann and T. A. Enßlin, *A simulation pipeline for the planck mission, Astronomy and Astrophysics* **445** (Jan, 2006) 373.

[306] J. R. Peterson, J. G. Jernigan, S. M. Kahn, A. P. Rasmussen, E. Peng, Z. Ahmad et al., *Simulation of astronomical images from optical survey telescopes using a comprehensive photon monte carlo approach, The Astrophysical Journal Supplement Series* **218** (May, 2015) 14.

[307] J. B. Juin, D. Yvon, A. Réfrégier and C. Yèche, *Cosmology with wide-field sz cluster surveys: selection and systematic effects, Astronomy and Astrophysics* **465** (Apr, 2007) 57.

[308] S. Pires, J.-L. Starck, A. Amara, R. Teyssier, A. Réfrégier and J. Fadili, *Fast statistics for weak lensing (fastlens): fast method for weak lensing statistics and map making, Monthly Notices of the Royal Astronomical Society* **395** (May, 2009) 1265.

[309] J. P. Dietrich and J. Hartlap, *Cosmology with the shear-peak statistics, Monthly Notices of the Royal Astronomical Society* **402** (Feb, 2010) 1049.

[310] L. Marian, R. E. Smith, S. Hilbert and P. Schneider, *Optimized detection of shear peaks in weak lensing maps, Monthly Notices of the Royal Astronomical Society* **423** (Jun, 2012) 1711.

[311] C. Heymans, L. V. Waerbeke, D. Bacon, J. Berge, G. Bernstein, E. Bertin et al., *The shear testing programme - i. weak lensing analysis of simulated ground-based observations, Monthly Notices of the Royal Astronomical Society* **368** (May, 2006) 1323.

[312] R. Massey, C. Heymans, J. Bergé, G. Bernstein, S. Bridle, D. Clowe et al., *The shear testing programme 2: Factors affecting high-precision weak-lensing analyses, Monthly Notices of the Royal Astronomical Society* **376** (Mar, 2007) 13.

[313] S. Bridle, J. Shawe-Taylor, A. Amara, D. Applegate, S. T. Balan, G. Bernstein et al., *Handbook for the great08 challenge: An image analysis competition for cosmological lensing, Annals of Applied Statistics* **3** (Jan, 2009) 6.

[314] A. Refregier and A. Amara, *A way forward for cosmic shear: Monte-carlo control loops, Physics of the Dark Universe* **3** (Apr, 2014) 1.

[315] C. Bruderer, C. Chang, A. Refregier, A. Amara, J. Bergé and L. Gamper, *Calibrated ultra fast image simulations for the dark energy survey, The Astrophysical Journal* **817** (Jan, 2016) 25.

[316] A. J. Benson, A. Farahi, S. Cole, L. A. Moustakas, A. Jenkins, M. Lovell et al., *Dark matter halo merger histories beyond cold dark matter - i. methods and application to warm dark matter, Monthly Notices of the Royal Astronomical Society* **428** (Jan, 2013) 1774.

[317] R. Barkana, Z. Haiman and J. P. Ostriker, *Constraints on warm dark matter from cosmological reionization, The Astrophysical Journal* **558** (Sep, 2001) 482.

[318] R. H. Wechsler, J. S. Bullock, J. R. Primack, A. V. Kravtsov and A. Dekel, *Concentrations of dark halos from their assembly histories, The Astrophysical Journal* **568** (Mar, 2002) 52.

[319] A. D. Ludlow, J. F. Navarro, M. Boylan-Kolchin, P. E. Bett, R. E. Angulo, M. Li et al., *The mass profile and accretion history of cold dark matter haloes, Monthly Notices of the Royal Astronomical Society* **432** (Jun, 2013) 1103.

[320] A. R. Duffy, J. Schaye, S. T. Kay and C. D. Vecchia, *Dark matter halo concentrations in the wilkinson microwave anisotropy probe year 5 cosmology, Monthly Notices of the Royal Astronomical Society: Letters* **390** (Oct, 2008) L64.

[321] A. Schneider, *Structure formation with suppressed small-scale perturbations, Monthly Notices of the Royal Astronomical Society* **451** (Aug, 2015) 3117.

[322] F. Jiang and F. C. van den Bosch, *Statistics of dark matter substructure - i. model and universal fitting functions, Monthly Notices of the Royal Astronomical Society* **458** (May, 2016) 2848.

[323] F. C. van den Bosch, G. Tormen and C. Giocoli, *The mass function and average mass-loss rate of dark matter subhaloes, Monthly Notices of the Royal Astronomical Society* **359** (May, 2005) 1029.

[324] C. Giocoli, G. Tormen and F. C. van den Bosch, *The population of dark matter subhaloes: mass functions and average mass-loss rates, Monthly Notices of the Royal Astronomical Society* **386** (Jun, 2008) 2135.

[325] M. Boylan-Kolchin, C.-P. Ma and E. Quataert, *Dynamical friction and galaxy merging time-scales, Monthly Notices of the Royal Astronomical Society* **383** (Jan, 2008) 93.

[326] J. Peñarrubia, J. F. Navarro and A. W. McConnachie, *The tidal evolution of local group dwarf spheroidals, The Astrophysical Journal* **673** (Jan, 2008) 226–240.

[327] J. Peñarrubia, A. J. Benson, M. G. Walker, G. Gilmore, A. W. McConnachie and L. Mayer, *The impact of dark matter cusps and cores on the satellite galaxy population around spiral galaxies, Monthly Notices of the Royal Astronomical Society* **406** (Aug, 2010) 1290.

[328] J. D. Emberson, T. Kobayashi and M. A. Alvarez, *Evolution of low mass galactic subhalos and dependence on concentration, The Astrophysical Journal* **812** (Oct, 2015) 9.

[329] B. Turner and T. V. Zandt, *A tutorial on approximate bayesian computation, Journal of Mathematical Psychology* (Jan, 2012) .

[330] J. Liepe, P. Kirk, S. Filippi, T. Toni and C. Barnes. . ., *A framework for parameter estimation and model selection from experimental data in systems biology using approximate bayesian computation, Nature protocols* (Jan, 2014) .

[331] S. Dye, C. Furlanetto, A. M. Swinbank, C. Vlahakis, J. W. Nightingale, L. Dunne et al., *Revealing the complex nature of the strong gravitationally lensed system h-atlas j090311.6+003906 using alma, arXiv* **astro-ph.GA** (Mar, 2015) , [1503.08720v2].

[332] M. Rybak, J. P. McKean, S. Vegetti, P. Andreani and S. D. M. White, *Alma imaging of sdp.81 - i. a pixelated reconstruction of the far-infrared continuum emission, Monthly Notices of the Royal Astronomical Society: Letters* **451** (Jul, 2015) L40.

Bibliography

[333] J. Hunter, *Matplotlib: A 2d graphics environment, Computing in science and engineering* (Jan, 2007) .

[334] J. Akeret, S. Seehars, A. Amara, A. Refregier and A. Csillaghy, *Cosmohammer: Cosmological parameter estimation with the mcmc hammer, Astrophysics Source Code Library* (Mar, 2013) 1303.003.

List of Symbols

a Scale factor.

$\hat{\alpha}$ Physical deflection angle.

A_s Amplitude of the primordial power spectrum.

$\vec{\beta}$ Angular position in the source plane.

χ Radial co-moving distance.

d_A Angular diameter distance.

D_d Angular diameter distance to the lens.

D_ds Angular diameter distance from the lens to the source.

d_L Luminosity distance.

D_s Angular diameter distance to the source.

H_0 Hubble constant, i.e. $H(a)$ evaluated at $a = 1$.

ψ Lensing potential.

n_s Spectral index of the primordial power spectrum.

Ω	Density parameter ρ/ρ_c.
Ω_X	Density parameter of species X today.
ρ	Energy density.
r_s	Physical deflection angle.
sSFR	specific star formation rate = star formation rate/stellar mass.
SFRD	Star-formation rate density.
T	Transverse co-moving distance.
T_d	Transverse co-moving distance to the lens.
T_{ds}	Transverse co-moving distance from the lens to the source.
$\vec{\theta}$	Angular position in the observer plane.
$T(k)$	Transfer function.
T_s	Transverse co-moving distance to the source.
w	Equation of state parameter.
$Z(m, \text{SFR})$	Angular position in the observer plane.
z	Redshift.

Index

Curriculum Vitae

Simon Birrer

Born on 3rd of February 1988 in Sursee, Lucerne, Switzerland
from Luthern LU and Ebikon LU.

Address: Lindenweg 29, CH-6033 Buchrain, Switzerland
Phone: +41 77 433 38 77
E-Mail: simon.birrer@phys.ethz.ch

Education

Since May 2013	PhD candidate at Institute for Astronomy, ETH Zurich, Switzerland
Mars 2013	Master of Science in Physics, ETH Zurich, Switzerland
September 2011	Bachelor of Science in Physics, ETH Zurich, Switzerland
Fall 2010	Undergraduate exchange program at Hong Kong University of Science and Technology (HKUST), Hong Kong
Mai 2007	Federal Matura, Kantonsschule Alpenquai, Lucerne, Switzerland

Teaching Assistance

Fall 2014 - 2016	Advanced physics lab, student telescope
Fall 2013	Astrophysics I

Fall 2012	Physics I for D-Math/D-Phys
Spring 2010, 2011	Analysis II for D-Math/D-Phys
2008 - 2016	Coaching high school students for the Swiss Physics Olympiads

Mentoring

Fall 2016   Co-supervisor of Felix Kuhn (master thesis)
Strong and weak lensing in the COSMOS field

Fall 2016   Co-supervisor of Felix Kuhn (semester thesis)
Lens modeling with Dark Energy Survey data

Spring 2016   Co-supervisor of Florian Lienhard (semester thesis)
Dark matter rendering on the Dark Energy Survey

Spring 2016   Co-supervisor of Felix Mayor (semester thesis)
Line-of-sight reconstruction in the strong lens RXJ1131-1231

Fall 2015   Co-supervisor of Cyril Welschen (master thesis)
Lens environment in the COSMOS field

Spring 2015   Co-Supervisor of Kevin Fusshoeller (semester thesis)
Automatized analysis of strong lensing systems

Spring 2014   Co-supervisor of Janik Andrejkovic (semester thesis)
Shapelet decomposition of dark matter substructure

Talks and Seminars

October 2016   Research Seminar, IfA, ETH Zurich, Switzerland. Seminar talk.

Mai 2016   Dark Energy Survey Collaboration meeting, woking group presentation, SLAC, USA.

Mai 2016   STRIDES workshop, UCLA, USA. Contributed talk.

April 2016	Astronomische Vereinigung Aarau, outreach talk, Aarau, Switzerland
Mars 2016	Urania Sternwarte Zürich, outreach talk, Zurich, Switzerland
February 2016	Dark Matter@ETH workshop, review talk, Zurich, Switzerland.
October 2015	Dark Energy Survey Collaboration meeting, woking group presentation, Madrid, Spain.
July 2015	Tata Institute for fundamental research, Mumbai, India, seminar talk.
Mai 2015	Dark Energy Survey Collaboration meeting, plenary talk, Ann Arbor, Michigan, USA.
April 2015	STRIDES workshop, EPFL, Lausanne, Switzerland. Contributed talk.
Mars 2015	Research Seminar, IfA, ETH Zurich, Switzerland. Seminar talk.
Mars 2015	Treffpunkt Science City, main outreach program of ETHZ, outreach talk, Zurich, Switzerland
February 2015	Swiss Cosmology Days (conference), Geneva, Switzerland. Contributed talk.
September 2014	PhD day of the Swiss Study Fundation, Fribourg, Switzerland. Contributed talk.
October 2014	STRIDES workshop, Cambridge, UK. Contributed talk.
February 2014	Saas-Fee meeting advanced course, Engelberg, Switzerland. Winter school.
August 2014	Galaxy evolutin workshop, Santa Cruz, US. Contributed talk.
June 2014	EWASS, Geneva, Switzerland. Contributed talk.
2014 - 2016	Guide for public tours to the students telescope at ETH Zurich
September 2013	Research Seminar, IfA, ETH Zurich, Switzerland. Seminar talk.

February 2013	Swiss Cosmology Day (conference), Bern, Switzerland. Contributed talk.
2008 - 2016	Seminars of the Swiss Study Fundation, including "effective presentation", "negotiation skills", "Law & Politics", "Art of debating", "Ethics", "Life Science".
2008 - 2015	Five Summer schools of the Swiss and German Study Fundation, "Physics of the climate", "Society transformation process in eastern Europe", "Understanding Quantum physics", "Economy of Luck", "Quantum gravity".

Other Work Experiences

2014 - 2016	International Physics Olympiad (IPhO) 2016, Executive Chairman *(Co-chaired with the President of University of Zurich). Leading of the steering committee, supervision of the academic and administrative project management. Budget of 3 Mio. CHF, about 1000 people and many national institutions involved in the event.*
2012 - 2016	Municipality planing panel (political mandate), member *Strategic planing and supervision of the municipality in terms of traffic and urban development. Elected by the municipality board.*
2010 - 2015	Swiss Physics Olympiad, national coordinator *Responsible for the national selection process and training for the International Physics Olympiad.*
2008 - 2016	Swiss Army mountain specialist battalion, member *Milicia service. Instructor and guide for other army formations,alpine rescue missions.*
11/2007 - 4/2008	Military service, Andermatt, Switzerland *Swiss Army mountain specialist school.*
2007, 2008	Post Finance, Luzern, Switzerland

Backoffice employee of the Swiss postal bank.

Awards and Support

2016   "Kugelpyramide" of the Association of Swiss Scientific Olymiads

Every year, the price gets awarded to a personality who made significant contributions to Youth and Science

2016   Experimental Innovation Award, Advanced Student Lab

Every year the department of physics of ETH Zurich awards the prize for the best innovation in the Advanced Student Lab.

2013 - 2016   Promoted by the Swiss Study Fundation, PhD program

The Swiss Study Fundation promotes prosperous PhD students in acquiring interdisciplinary and non-academic skills in summer schools, workshops and meetings and provides access to funding and a networking platform.

2009-2013   Promoted by the Swiss Study Fundation, student program

The Swiss Study Fundation promotes prosperous jung students in acquiring interdisciplinary and non-academic skills in summer schools, workshops and meetings and provides access to funding and a networking platform.

2007   Swiss Physics Olympiad

Highschool level individual physics competition.

1. place (winner) preliminary round

1. place (winner) national round

Honorable mention, International Physics Olympiad, Isfahan, Iran

Peer-Reviewed Publications

4. **S. Birrer**, A. Amara, and A. Refregier. *The mass-sheet degeneracy and time-delay cosmography: analysis of the strong lens rxj1131-1231. Journal of Cosmology and Astroparticle*

Physics **08** (Aug, 2016) 020

3. B. Nord, E. Buckley-Geer, H. Lin, H. T. Diehl, J. Helsby, N. Kuropatkin, A. Amara, T. Collett, S. Allam, G. B. Caminha, C. D. Bom, S. Desai, H. Dumet-Montoya, M. E. da S Pereira, D. A. Finley, B. Flaugher, C. Furlanetto, H. Gaitsch, M. Gill, K. W. Merritt, A. More, D. Tucker, A. Saro, E. S. Rykoff, E. Rozo, **S. Birrer**, F. B. Abdalla, A. Agnello, M. Auger, R. J. Brunner, M. C. Kind, F. J. Castander, C. E. Cunha, L. N. da Costa, R. J. Foley, D. W. Gerdes, K. Glazebrook, J. Gschwend, W. Hartley, R. Kessler, D. Lagattuta, G. Lewis, M. A. G. Maia, M. Makler, F. Menanteau, A. Niernberg, D. Scolnic, J. D. Vieira, R. Gramillano, T. M. C. Abbott, M. Banerji, A. Benoit- Levy, D. Brooks, D. L. Burke, D. Capozzi, A. C. Rosell, J. Carretero, C. B. D'Andrea, J. P. Dietrich, P. Doel, A. E. Evrard, J. Frieman, E. Gaztanaga, D. Gruen, K. Honscheid, D. J. James, K. Kuehn, T. S. Li, M. Lima, J. L. Marshall, P. Martini, P. Melchior, R. Miquel, E. Neilsen, R. C. Nichol, R. Ogando, A. A. Plazas, A. K. Romer, M. Sako, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, J. Thaler, A. R. Walker, W. Wester, Y. Zhang, and T. D. Collaboration. *Observation and confirmation of six strong-lensing systems in the dark energy survey science verification data. The Astrophysical Journal* **827** (Aug, 2016) 51

2. **S. Birrer**, A. Amara, and A. Refregier. *Gravitational lens modeling with basis sets. The Astrophysical Journal* **813** (Nov, 2015) 102

1. **S. Birrer**, S. Lilly, A. Amara, A. Paranjape, and A. Refregier. *A simple model linking galaxy and dark matter evolution. The Astrophysical Journal* **793** (Sep, 2014) 12

Submitted Publications

2. N. Caplar, S. Tacchella and **S. Birrer**. *Quantitative evaluation of gender bias in astronomical publications from citation counts. Submitted to Nature Astronomy* (Oct, 2016)

1. **S. Birrer**, C. Welschen, A. Amara, and A. Refregier. *Line-of-sight effects in strong lensing: Putting theory into practice. Submitted to JCAP* (Oct, 2016)

Acknowledgements

I started my undergraduate studies at ETH in Fall 2008. And now after more than eight years ago, it's time to move on - and to say thank you to many people that made this time joyful and with whomes support I overcame many challenges.

First, I would like to express my gratitude to my supervisors **Adam Amara** and **Alexandre Refregier**. It was a pleasure working with Adam and Alex and a privilege of having two supervisors. They gave me support and supervision where I needed it and left me with a lot of freedom in proceeding my work, what I very much appreciate. Their demand on quality in the science and its presentation pushed me and I am very grateful for their support throughout.

I enjoyed being part of ETH Cosmology Research Group. It was a stimulating environment with supportive colleagues and friends, more than just a work. Thank you Cosmology Group PhD friends, **Claudio Bruderer**[1], **Andrina Nicola**[2], **Sebastian Seehars**[3] and **Jörg Herbel** for the support you gave me, and definitely also for the fun we had. I enjoyed tremendous support from **Joel Akeret** in any aspects of software development. He taught me many useful tools, advised me in design and software solutions, provided group internal packages and his help often saved me a lot of time to be spent on science rather than IT problems. I also benefited of **Aseem Paranjape** and **Aurel Schneider** from their expertise and experience in research. Throughout my employment in the Cosmology Group, I enjoyed the great support of **Annie Aguis-Luethi**. I am probably very spoiled now in administrative support in every aspect.

---

[1] who would make a silly comment, if not listed first ;-)
[2] who owns also to be listed first, see footnote of Claudio
[3] He taught me advanced statistical methods and I got the thesis template from him

241

Acknowledgements

During my masters thesis and at the start into my PhD, I enjoyed a lot of discussions with **Simon J. Lilly** on galaxy evolution. This broadened my view and I can still benefit from this now. Of great use were also discussions with **Neven Caplar** and **Sandro Tacchella** about science related to galaxy evolution and for letting me joint their data science project on citation counts. I thank **Sherry Suyu** for being my co-examiner and for many fruitful discussions and the detailed and very useful feedback on my work.

The time at ETH was not only about science, it was a life experience. From the first semester at ETH I met many people and we are still closed friends. This thank goes in particular to **Jens**, **Aline**, **Peggy**, **Gian**, **Anja**, **Samuel**, **Christa**, **Andrea**, **Anna** and **Claudio**. I would like to thank all my PhD colleagues at the Institute for Astronomy for the many moments of joy[4]. It was a pleasure having you around.

All the way through to this point, I got supported by my family. They supported always the path I wanted to take.

During my PhD, I enjoyed intense interactions with the Dark Energy Survey collaboration, in particular the strong lensing working group, the STRIDES collaboration and the H0LiCOW collaboration. I would like to thank all members for the stimulating discussions and their feedback on my work. I also had the chance to work with several masters and bachelor project students. From their work I could benefit and have a direct or indirect impact on the work presented in this thesis. This thank goes to **Janik Andrejkovic**, **Kevin Fusshoeller**, **Cyril Welschen**, **Felix Mayor**, **Florian Lienhard** and **Felix Kuhn**. Three publication of which I was the corresponding author went through a referee process. Comments and suggestions made by the anonymous referees did improve the quality and presentation of those publications and so this thesis.

Besides of my scientific studies, I was the teaching assistant for the student telescope at ETH Zurich, where undergraduates get the chance to proceed their first astronomical observations. The work spent by **Andreas Bazzon** and **Hans-Martin Schmid** to commission and maintaining the telescope helped me a lot in sharing the passion for astronomy with the students I was supervising. Since my undergraduate studies, I got promoted by the Swiss Study Foundation

---

[4]This includes many cakes. Thank you Anna, Lia & co!

through seminars and workshops, not directly related to the science I presented in this thesis but of great use during my PhD and beyond.

This thesis had not been possible without the extensive use of open source software and computing power. For heavy computing tasks, I used the computer cluster Monch at the Swiss National Supercomputing Center[5]. I would like to thank the open access software community for providing a wealth of useful packages that made my live much easier as a PhD student. Especially, I would like to thank the Python community and the developer of the packages numpy [6], scipy [7], astropy [8], matplotlib (333)[9], FASTELL (199), corner [10] and CosmoHammer (334).

I would also like to acknowledge the SAO/NASA Astropysical Data System[11], the arXiv preprint server[12], the Milkulsky Archive for Space Telescopes (MAST)[13] and the wealth of publicly available data and data products from the NASA and ESA missions and telescopes. I thank the public of Switzerland for the generous support of science through funding the ETH domain and the Swiss National Science Foundation.

---

[5]www.cscs.ch
[6]www.numpy.org
[7]www.scipy.org
[8]www.astropy.org
[9]www.matplotlib.org
[10]https://github.com/dfm/triangle.py
[11]www.adsabs.harvard.edu/
[12]www.arxiv.org
[13]http://archive.stsci.edu