

Can the halo model describe 2nd- and 3rd-order correlation functions of gravitational lensing consistently?

DISSERTATION zur Erlangung des Doktorgrades (Dr. rer. nat.) der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

von **Sandra Maria Martin** aus Wichita Falls, Texas

Bonn, Mai 2019

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf dem Hochschulschriftenserver der ULB Bonn <http://nbn-resolving.de/urn:nbn:de:hbz:5n-56282> elektronisch publiziert.

1. Gutachter:	Prof. Dr. Peter Schneider
2. Gutachter:	Prof. Dr. Cristiano Porciani
Tag der Promotion:	9. Oktober 2019
Erscheinungsjahr:	2019

Für meine Münchner Großeltern.

Abstract

To understand and model the distribution of and the relation between baryonic and dark matter in the Universe is one of the key challenges in contemporary astrophysics. A well-established theoretical description is given by the semi-analytical *halo model*, which combines the *dark matter halo model* and the *halo occupation distribution* (HOD). Whereas the former reduces the complex distribution of dark matter to the clustering of dark matter halos on large scales and the radial distribution of dark matter within these halos on small scales, the latter incorporates galaxies based on the assumption that galaxies can only form and live within dark matter halos. The validity of the halo model is determined by how well its predictions match ever-newer observations.

A unique tool to map the matter distribution in the Universe is the *gravitational lensing effect*, the phenomenon that light rays emitted from distant objects get differentially deflected by the gravitational potential of the intervening matter distribution, visible or dark. As statistical applications of the weak gravitational lensing effect, *galaxy-galaxy lensing* (G2L) and *galaxy-galaxy-galaxy lensing* (G3L) probe the average matter density profile about galaxies and pairs of galaxies, respectively, thereby revealing the relation between galaxies and their dark host halos. The halo model is known to provide a good description of second-order statistics as G2L, but so far neither a quantitative comparison of halo model predictions for G3L to observations nor direct model fits to observations of G3L are available.

The main goal of this doctoral thesis is to test whether the halo model can describe measurements of G2L and G3L consistently. To this end halo model fits are performed to the G2L signal measured from the Canada-France Hawaii Telescope Lensing Survey (CFHTLenS) for 29 galaxy samples of stellar mass ($5 \times 10^9 M_\odot \leq M_* \leq 2 \times 10^{11} M_\odot$), luminosity ($-23 \leq M_r \leq -18$) and galaxy-type, further differentiating between low ($0.2 \leq z_{\text{ph}} < 0.44$) and high redshift ($0.44 \leq z_{\text{ph}} < 0.6$) samples. Based on the best-fit models, predictions of G3L in terms of the *aperture statistics* $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are generated, which are confronted with their observational counterparts from CFHTLenS. The comparison shows that the halo model can successfully describe G3L at a level of accuracy that is on par with that of dark matter simulations into which baryonic physics is incorporated using semi-analytical models (SAMs).

Moreover, first-time halo model predictions of the more intuitive representation of G3L as *excess mass maps* are presented. Trends of excess mass with lens-lens separation, galaxy properties, and redshift are studied and are discussed together with the respective predictions for the aperture statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$. The results suggest that excess mass

increases with stellar mass and luminosity, and decreases with redshift. The results confirm the observation of excess mass to increase with decreasing lens-lens separation, and to be more than one order of magnitude higher around pairs of early-type compared to late-type galaxies. Additionally, the dependence of excess mass on halo model properties is explored; i.e. the contributions of the one-, two-, and three-halo terms are quantified. For a projected lens-lens separation of 1 arcmin the one-halo term is found to be suppressed for late-type galaxies as a consequence of them being typically field galaxies. The results for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ show that for all other samples the one-halo term clearly dominates up to aperture scales of 10 arcmin. A sensitivity analysis regarding the dependence of **G3L** on individual **HOD** parameters shows that, first, changes are maximal in the range probed by **CFHTLenS** (1 – 10 arcmin). Second, changes in **G3L** exceed 20% for four out of five parameters when varied individually by $\pm 20\%$ around their best-fit values, indicating that simultaneous model fits to **G2L** and **G3L** will help to constrain the **HOD**. Finally, it is tested whether the halo model can describe map features observed with **CFHTLenS**, in particular a vertical bulge-like feature that is absent in predictions of **SAMs**. Although the halo model cannot reproduce the feature, the predictions match the observations regarding the amplitude of the signal around the lenses and the rate of drop-off of the signal towards the outer regions of the map.

Contents

	Page
Preface	1
1 The standard model of Cosmology – A brief review	5
1.1 From General Relativity to a homogeneous & isotropic world model	5
1.1.1 Theory of General Relativity	5
1.1.2 The pillars of the Λ CDM model	6
1.1.3 Friedmann-Robertson-Walker models	7
1.1.4 The content of our Universe	7
1.1.5 Distance measures	10
1.2 The formation of structures	10
1.2.1 The origin of structure	11
1.2.2 Evolution of the density fluctuations	12
1.3 The Universe – a random field	16
1.3.1 Correlation functions	17
1.3.2 Spectra	18
1.3.3 Dark matter spectra from perturbation theory	19
2 The halo model	21
2.1 The dark matter halo model	23
2.1.1 Halo mass function	23
2.1.2 Halo density profile	24
2.1.3 Dark matter halo bias	25
2.2 The halo occupation distribution	25
2.2.1 Modelling the relation between dark matter halos and galaxies	26
2.2.2 Derived parameters	27
2.2.3 Conversion to the ratio of stellar to halo mass	27
2.3 Halo model galaxy-matter power- and bispectrum	28
2.3.1 Halo model galaxy-dark matter power spectrum	28
2.3.2 Halo model galaxy-dark matter bispectrum	29
2.3.3 Projecting the spectra from 3D to 2D - Limber's Equation	30
2.3.4 Reduced bispectrum	31

3	Galaxy-galaxy(-galaxy) lensing	33
3.1	Definitions	34
3.1.1	Galaxy-galaxy lensing	35
3.1.2	Galaxy-galaxy-galaxy lensing	36
3.2	Halo model interpretation	38
3.2.1	Modelling galaxy-galaxy lensing	39
3.2.2	Modelling galaxy-galaxy-galaxy lensing	39
3.2.3	Galaxy-galaxy-galaxy lensing for a Gaussian bispectrum	41
4	Halo model fits to the Millennium Simulation & the CFHTLenS G2L	43
4.1	Data	44
4.1.1	CFHTLenS	44
4.1.2	Mock data – Millennium Simulation & Semi-Analytical Model	45
4.2	Method	45
4.2.1	Fitting procedure	46
4.2.2	Error estimates on best-fit parameters and model predictions	47
4.3	Limitations of the model and the fitting procedure	49
4.3.1	Model fits to G2L mock data	49
4.3.2	Accuracy of predicting the mean halo mass and fraction of centrals	51
4.3.3	Accuracy of the HOD model prediction	54
4.3.4	Conclusions on the limitations of the fitting routine and the model	56
4.4	Constraining HODs with CFHTLenS data	56
4.4.1	Model fits to G2L CFHTLenS data	57
4.4.2	Mean halo mass and fraction of centrals	59
4.4.3	Ratio of stellar to halo mass	64
4.4.4	Conclusions on the model fits to G2L from CFHTLenS	65
5	First-time prediction of excess mass maps using the halo model – dependence of G3L on galaxy & model properties	67
5.1	Method	68
5.1.1	Predicting G3L	68
5.1.2	Uncertainties in predicting G3L	68
5.2	Dependence of excess mass on galaxy properties	70
5.3	One-halo, two-halo, and three-halo term contributions	75
5.4	Sensitivity of G3L with respect to HOD parameters	78
5.5	Excess mass maps – model predictions versus observations	84
5.6	Conclusions	86
6	Testing G3L halo model predictions against CFHTLenS observations	89
6.1	Data	89
6.2	Method	90
6.2.1	Predicting the G3L aperture statistics	90
6.2.2	Uncertainties of the model predictions	90
6.3	G3L aperture statistics – model predictions versus observations	91
6.4	Conclusions	96

Conclusions & Outlook	99
A Natural constants & model parameters	105
B Eisenstein & Hu transfer function	107
C Normalisation of the linear power spectrum	111
D Derivative in the halo mass function	113
E Numerical interfaces to the bispectrum	115
E.1 Excess mass maps	115
E.2 Aperture statistics	116
F Model fits to the mock & CFHTLenS G2L	119
G First-time halo model prediction of excess mass	129
H Testing G3L halo model predictions	137
Bibliography	143
List of Figures	148
List of Tables	151
List of Acronyms	155
Acknowledgements	159

Preface

The overwhelming wealth of galaxies observed on the sky (Fig. 1) motivates the science of the Universe as a whole, cosmology. The aim is to explain the origin, the evolution, the variety and the large-scale distribution of galaxies. The earliest information available regarding the distribution of matter in the Universe comes from a snapshot of the Universe when it was only 380 000 years old. At this time the initially infinitely dense and hot Universe had sufficiently expanded and its temperature had sufficiently dropped for neutral atoms to form and for photons to stream freely. This relic Cosmic Microwave Background (CMB) radiation carries information about the spatial temperature distribution, and thereby on the distribution of matter at that time. Latest CMB measurements confirmed that this distribution is highly isotropic and homogeneous with only tiny fluctuations at a level of 10^{-5} (Planck Collaboration et al. 2016b). These are the seeds of today's structures. The question is how these tiny initial perturbations evolved to the massive structures observed today in form of planets, stars, galaxies, galaxy clusters, and the large-scale structure.



Figure 1: Hubble Space Telescope image of the Pegasus constellation. Aside from a couple of foreground stars, each visible object in this photo is a galaxy. [Source: <http://hubblesite.org/images/gallery>]

According to the standard model of cosmology, the Lambda Cold Dark Matter (Λ CDM) model (**Chapter 1**), dark matter plays a crucial role in the formation of structures by forming the first overdensities in the Universe. Freed from the constant photon pressure after the release of the CMB, the baryons fell into these dark matter overdensities, gravitationally collapsed and hierarchically evolved to the structures observed today. Due to this coupled - however, not identical - evolution, a relation between today's galaxies and their dark matter environment is expected. To infer this relation is, however, challenging. First, dark matter is by its very nature dark and can only be detected indirectly through its gravitational interactions (e.g. Bertone et al. 2005). Second, the complex hydrodynamical interactions baryons undergo from the beginning of time until today can only be described with limitations analytically and numerically, although recently great improvements have been achieved (Vogelsberger et al. 2014, Schaye et al. 2015).

By combining the findings from perturbation theory, numerical dark matter N -body simulations, and observations, the *halo model* provides a well established description of the distributions of dark matter and galaxies, and of their relation (Cooray & Sheth 2002). The model consists of the *dark matter halo model* and the *halo occupation distribution* (HOD) (**Chapter 2**). Whereas the former replaces the complex spatial distribution of dark matter by the distribution of distinct *dark matter halos*, the latter includes galaxies into this framework based on the physical assumption that galaxies can only form and live in dark matter halos, because baryonic matter needs the gravitational potential of the dark matter overdensities to cool and form stars.

The halo model allows for an analytical description of the statistical properties of the matter and galaxy field. Emphasis being on statistical properties, because no theory is or ever will be able to predict the exact realisation of the Universe down to individual galaxies, i.e. no theory can predict the Andromeda galaxy to be a neighbouring galaxy of the Milky Way at a distance of 2.5 million light years. However, one can predict the probability to find a galaxy with the same properties as the Andromeda galaxy at a distance of 2.5 million light years from a galaxy with the same properties as the Milky Way by using the concept of n -point correlation functions, or their Fourier space analogues: the spectra. The relation between galaxies and dark matter would be fully determined, if all n -point correlation functions are known. However, state of the art are the measurement of the two- and the three-point correlation function. Halo model predictions successfully describe the measured two-point cross-correlation of galaxies and dark matter (Leauthaud et al. 2012, Velander et al. 2014, Coupon et al. 2015, Clampitt et al. 2017). However, the validity of the halo model predictions with regard to third-order cross-correlations has not yet been tested.

Ideal measurements to confront the halo model are those exploiting the *gravitational lensing effect* (Schneider 2006), i.e. the phenomenon that light rays emitted by distant 'sources' get on their path to earth differentially deflected by the gravitational potential of the intervening matter distribution. The gravitational lensing effect distorts the intrinsic shapes of the source images, thereby imprinting information about the (dark) matter environment of the lensing objects into the sheared images. While the lensing effect can lead to strong distortions resulting in arc-like shaped images, this work is concerned with subtle distortions attributed to the *weak lensing effect* (**Chapter 3**), particularly with one

of its statistical applications: *galaxy-galaxy lensing* (G2L). G2L correlates the sheared image of a background galaxy with the position of a lens galaxy, and probes by that the average (dark) matter environment around a typical lens. Thus, G2L provides not only a possibility to map dark matter, but also to infer the statistical properties of the galaxy-halo connection (e.g. Zheng et al. 2007, Zehavi et al. 2011, Leauthaud et al. 2011, Coupon et al. 2012, Velander et al. 2014).

The newest tool in the field of galaxy-galaxy lensing is *galaxy-galaxy-galaxy lensing* (G3L), which extends the concept of G2L to third-order (Schneider & Watts 2005). By considering the cross-correlation between the sheared image of a source galaxy with the positions of two lens galaxies, G3L probes the average (dark) matter environment around pairs of galaxies. Thus, G3L is a promising tool to study the dark matter environment of small gravitationally bound systems (Simon et al. 2008, Simon et al. 2013, Simon et al. 2019), and contains valuable information about the relationship between galaxies and their dark host that G2L cannot provide. An intuitive visualisation of G3L is provided in terms of *excess mass maps* (Simon et al. 2008), which map the projected average mass around pairs of lenses in excess of the mass measured around the individual lenses with G2L. This means that there is more (dark) matter associated with two galaxies that are physically close compared to two galaxies that are isolated.

The aim of this doctoral thesis is to test the ability of the halo model to consistently describe the second- and third-order cross-correlation of galaxies and (dark) matter, as probed by G2L and G3L. To this end, the halo model introduced in **Chapter 2** is in a first step fitted to G2L measured from the Canada-France Hawaii Telescope Lensing Survey (CFHTLenS) (**Chapter 4**). The best-fit models are in a second step used to generate halo model predictions for G3L, which are confronted with their measured counterparts from CFHTLenS (**Chapter 6**). This confrontation constitutes the next level test for the halo model and the assumptions it is based on.

Moreover, first-time halo model predictions of G3L in terms of excess mass maps are presented in **Chapter 5**. The dependence of excess mass on galaxy and halo model properties is studied in order to build up a knowledge base on which to interpret present and future measurements of G3L. Additionally, it is investigated whether the halo model can reproduce the map features found in recent measurements from CFHTLenS and from the Millennium Simulation (Simon et al. 2019).

Chapter 1

The standard model of Cosmology – A brief review

Cosmology deals with the Universe as a whole and its evolution over the course of time. The aim is a thorough, mathematical description of all physical processes taking place in it with regard to its constituents and the acting forces. This includes a theory for the distribution and evolution of structures in the Universe. The halo model formalism is one branch of this theory and is described in detail in Chapter 2. In this chapter the aspects of the standard model of cosmology, the Lambda Cold Dark Matter ([\$\Lambda\$ CDM](#)) model, which are essential for the comprehension of the halo model are briefly reviewed. This review is large based on cosmological standard literature, e.g. [Peacock \(1999\)](#), [Dodelson \(2003\)](#) and [Schneider \(2014\)](#). If topics or derivations are not covered by this literature, specific references are given.

1.1 From General Relativity to a homogeneous & isotropic world model

1.1.1 Theory of General Relativity

The mathematical framework for the [\$\Lambda\$ CDM](#) model was laid when Einstein published the theory of General Relativity ([GR](#)) in 1915 ([Einstein 1915](#)). [GR](#) unifies Special Relativity ([Einstein 1905](#)) and Newton's law of gravitation ([Newton 1687](#)) and extends them for the case of large-scales, making [GR](#) the description of gravitation in modern physics.

In this theory gravity is interpreted as a geometric property of spacetime: Not only does mass act as a source of gravity, but mass must be seen as a part of a more general quantity of energy and momentum. The energy-momentum tensor $T_{\mu\nu}$ is then the source of the gravitational field, which in turn determines the geometry of spacetime described by the Einstein tensor $G_{\mu\nu}$. This concept is comprised in Einstein's field equations:

$$G_{\mu\nu} = -\frac{8\pi G_{\text{N}}}{c^4} T_{\mu\nu} - \Lambda g_{\mu\nu}, \quad (1.1)$$

where the factor $\frac{8\pi G_{\text{N}}}{c^4}$ is called Einstein's gravitational constant with G_{N} being Newton's gravitational constant and c being the speed of light (see Table [A.1](#) in Appendix [A](#) for

the values of natural and cosmological constants). The field equations are generalised by adding the constant term $\Lambda g_{\mu\nu}$, either to include the possibility of a static universe or to quantify the accelerated expansion of the Universe (Riess et al. 1998, Perlmutter et al. 1999) in terms of the *cosmological constant* Λ . The latter is associated with *dark energy* and can be interpreted as a vacuum energy density (Frieman et al. 2008). The metric $g_{\mu\nu}$ defines on one hand spatial and temporal distances in spacetime, and on the other hand the geodesics on which free particles and light rays travel. Accordingly, the metric tensor determines the gravitational potential and builds a bridge between the geometry of spacetime and gravity.

1.1.2 The pillars of the Λ CDM model

The large-scale distribution of galaxies (Colless 1999) and the temperature distribution of the Cosmic Microwave Background (CMB) radiation (Planck Collaboration et al. 2016b) indicate that the Universe is nearly isotropic. If the position of Earth in the Universe is not particular, isotropy results directly in the homogeneity of the Universe when averaging over scales > 200 Mpc. The assumption of isotropy and homogeneity is formulated as the *Cosmological Principle*, which is one pillar of the Λ CDM model.

The second pillar is the observation, made initially by Hubble in 1928 (Hubble 1929), that most galaxies move away from Earth with a radial velocity v proportional to their distance D

$$v = H_0 D, \quad (1.2)$$

with the *Hubble constant*

$$H_0 := 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}. \quad (1.3)$$

The actual value of H_0 is parametrised by the dimensionless constant h , which is of order unity (Table A.1). The *Hubble law* can be interpreted as the expansion of space itself, rather than the movement of galaxies away from Earth. This expansion is parametrised by the scale factor $a(t)$, which relates the actual position $\mathbf{r}(t)$ of an object on an expanding sphere with the position \mathbf{x} of the *comoving observer*

$$\mathbf{r}(t) = a(t)\mathbf{x}. \quad (1.4)$$

Conventionally, the scale factor is normalised to unity today, $a(t_0) = 1$. One can then define the *expansion rate* by

$$H(t) = \frac{\dot{a}(t)}{a(t)}, \quad (1.5)$$

which reduces to the Hubble constant for $t = t_0$.

As a consequence of the expansion, a light ray emitted from a distant object travelling through the expanding Universe towards an observer experiences a decrease in the energy of its photons; the photons get redshifted. If peculiar velocities can be neglected, the ratio

$$z := \frac{\lambda - \lambda_0}{\lambda_0}, \quad (1.6)$$

with λ being the observed wavelength of a spectral line and λ_0 the corresponding one in restframe, called *redshift*, can be used as a proxy for the radial velocity and the distance of the object. The redshift and the scale factor correspond to each other via

$$a = \frac{1}{1+z}. \quad (1.7)$$

1.1.3 Friedmann-Robertson-Walker models

To wed the mathematical framework for gravity represented by **GR** with the observation of a homogeneous, isotropic, expanding and infinite Universe, was the challenge for cosmologists in the 20th century.

In 1936 Robertson and Walker ([Robertson 1935](#), [1936a,b](#); [Walker 1937](#)) showed that a homogeneous, isotropic and expanding (or contracting) universe can be realised in the framework of **GR** by assuming a metric of the form

$$ds^2 = c^2 dt^2 - a^2(t)[d\chi^2 + f_K^2(\chi)(d\theta^2 + \sin^2(\theta)d\varphi^2)], \quad (1.8)$$

where s is the separation of two neighbouring events in spacetime, t the cosmic time, $a(t)$ the cosmic scale factor, θ and φ the angular coordinates, χ the comoving radial distance and $f_K(\chi)$ the comoving angular diameter distance. The latter depends on the curvature of space K in the following way:

$$f_K(\chi) = \begin{cases} K^{-1/2} \sin(K^{1/2} \chi) & \text{for } K > 0, \\ \chi & \text{for } K = 0, \\ (-K)^{-1/2} \sinh [(-K)^{1/2} \chi] & \text{for } K < 0, \end{cases} \quad (1.9)$$

which in turn depends on the content of the Universe (see Sect. 1.1.4).

The Robertson-Walker metric is an exact solution to Einstein's field equations for which the latter reduce to two independent dynamical equations describing the behaviour of the scale factor, the *Friedmann equations*,

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G_N}{3} \rho - \frac{Kc^2}{a^2}, \quad (1.10)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_N}{3} \left(\rho + \frac{3p}{c^2}\right), \quad (1.11)$$

with density $\rho(t)$ and pressure $p(t)$.

1.1.4 The content of our Universe

According to the Friedmann equations, the dynamics of the Universe depend on the density and pressure of its content. Generally, every constituent needs to obey the first law of thermodynamics, which states that the inner energy U of a system depends on the interplay of its pressure P and its volume V . In comoving coordinates the first law of thermodynamics reads

$$d(\rho c^2 a^3) = -p d(a^3), \quad (1.12)$$

where according to Special Relativity ρc^2 is the energy density of ordinary matter. In cosmology one distinguishes between three types of matter: pressureless matter, radiation, and vacuum energy. The constituents of *pressureless matter* have a random velocity much smaller than the speed of light, $p_m \ll \rho_m c^2$, so that pressure can be neglected, $p_m = 0$. Applying the first law of thermodynamics leads to the following expression for the matter density

$$\rho_m(t) = \rho_{m,0} a^{-3}(t), \quad (1.13)$$

where $\rho_{m,0}$ is today's density. In the case of *radiation* the random velocity of the constituents is comparable to the speed of light and pressure can be described by $p_r = \rho_r c^2/3$. Then the radiation density is given by

$$\rho_r(t) = \rho_{r,0} a^{-4}(t), \quad (1.14)$$

with $\rho_{r,0}$ being the current radiation density. Due to cosmic expansion the particles get redshifted, which results in the additional factor of a^{-1} . It is assumed that in the last case of *vacuum energy* the density does not vary with time and is characterised by the cosmological constant Λ ,

$$\rho_\Lambda = \frac{\Lambda}{8\pi G_N}. \quad (1.15)$$

As a consequence the pressure is negative, $p_\Lambda = -\rho_\Lambda c^2$. In summary, the total density and pressure of the Universe are given by

$$\begin{aligned} \rho(a) &= \rho_m + \rho_r + \rho_\Lambda = \rho_{m,0} a^{-3} + \rho_{r,0} a^{-4} + \frac{\Lambda}{8\pi G_N}, \\ p(a) &= \frac{\rho_r c^2}{3} - \rho_\Lambda c^2 = \frac{\rho_{r,0} c^2}{3a^4} - \rho_\Lambda c^2. \end{aligned} \quad (1.16)$$

Going back to the Friedmann equations, one can now focus on the limiting case where the curvature of spacetime K is zero, which corresponds to a flat universe, and define the *critical density* as

$$\rho_{\text{crit}}(a) := \frac{3H^2(a)}{8\pi G_N}, \quad (1.17)$$

which is the density of the Universe required to maintain a flat cosmology. This characteristic density is used to normalise the densities and define the *density parameters* by

$$\Omega_i(a) := \frac{\rho_i(a)}{\rho_{\text{crit}}(a)}, \quad (1.18)$$

which results in

$$\Omega_m(a) = \frac{\rho_m(a)}{\rho_{\text{crit}}(a)} = \frac{8\pi G_N}{3H^2(a)} \rho_{m,0} a^{-3} = \frac{\Omega_m}{a + \Omega_m(1-a) + \Omega_\Lambda(a^3 - a)}, \quad (1.19)$$

for the matter density (neglecting contributions from Ω_r). Analogously, the time dependent density parameter for dark energy can be derived,

$$\Omega_\Lambda(a) = \frac{\Lambda}{3H^2(a)} = \frac{\Omega_\Lambda a^3}{a + \Omega_m(1-a) + \Omega_\Lambda(a^3 - a)}. \quad (1.20)$$

In this work Ω_m and Ω_Λ represent the current values, obeying the general definition for the current density parameters

$$\Omega_i = \frac{\rho_{i,0}}{\rho_{\text{crit},0}}. \quad (1.21)$$

The critical density for today is given by

$$\rho_{\text{crit},0} = \frac{3H_0^2}{8\pi G_N} \approx 1.88 \times 10^{-29} h^2 \text{ g cm}^{-3} \approx 2.77485 \times 10^{11} h^2 \frac{\text{M}_\odot}{\text{Mpc}^3}, \quad (1.22)$$

corresponding to six hydrogen atoms per cubic meter for $h = 0.73$.

The first Friedmann Eq. (1.10), also known as the expansion equation, can now be rewritten in terms of the density parameters

$$H^2(a) = H_0^2 \left(\Omega_r a^{-4} + \Omega_m a^{-3} - \frac{Kc^2}{a^2 H_0^2} + \Omega_\Lambda \right), \quad (1.23)$$

where the curvature of spacetime depends on the total density parameter $\Omega_0 = \Omega_r + \Omega_m + \Omega_\Lambda$,

$$K = \frac{H_0^2}{c^2} (\Omega_0 - 1). \quad (1.24)$$

The physical interpretation of the expansion equation, written in this form, is that at the very beginning ($a \ll 1$) the Universe was dominated by radiation. For a flat Universe, the solution to the Friedmann equation is

$$a(t) = \left(2H_0 \sqrt{\Omega_r} t \right)^{1/2}. \quad (1.25)$$

With growing scale factor a , matter takes over at *radiation-matter equality*,

$$a_{\text{eq}} = \frac{\Omega_r}{\Omega_m} = 4.2 \times 10^{-5} (\Omega_m h^2)^{-1}, \quad (1.26)$$

corresponding to $z \approx 3000$. This transition is particularly for structure formation an important milestone, as in the matter-dominated era the solution to the expansion equation is given by

$$a(t) = \left(\frac{3}{2} H_0 \sqrt{\Omega_m} t \right)^{2/3}, \quad (1.27)$$

which implies a slow down in the expansion of the Universe compared to the radiation-dominated era. Today, the Universe is dominated by the cosmological constant, or dark energy, with

$$a(t) = e^{H_0 \sqrt{\Omega_\Lambda} t}, \quad (1.28)$$

corresponding to an accelerated expansion.

The set of density parameters measured by e.g. [Hinshaw et al. \(2013\)](#) for our Universe (Table A.1) indicates that it belongs to the class of nearly flat universes which expand forever (Fig. 1.1). Consequently, the size of the Universe must - at least formally - have been zero in the past, leading to an infinitely dense and hot state called *Big Bang*, which is defined as the origin of time, $t = 0$.

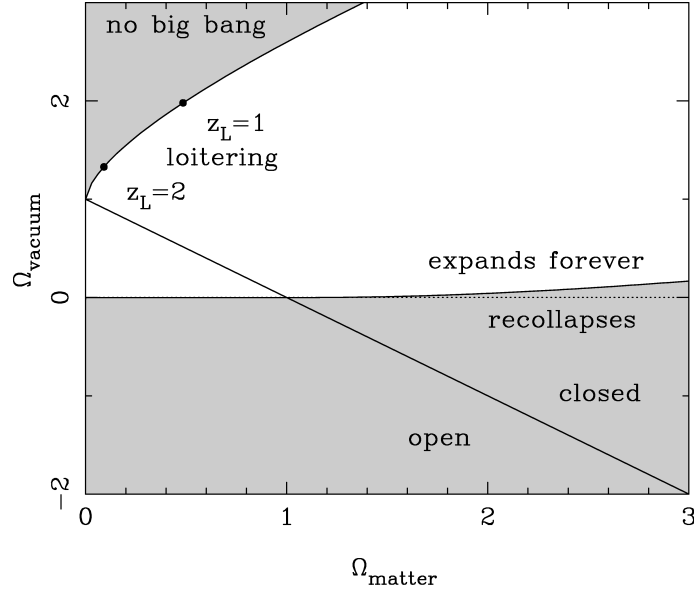


Figure 1.1: Classification of cosmological models with different combinations of Ω_m and Ω_Λ . The density parameters for our Universe are given by $\Omega_m = 0.279$ and $\Omega_\Lambda = 0.721$ [Source: Peacock (1999)].

1.1.5 Distance measures

In an expanding Universe with curved spacetime the definition of distance is not unique and results in a list of different measures. Three of them - the scale factor a , the redshift z and the radial comoving distance χ - have already been introduced. They are all related to each other. For example, the comoving distance can be expressed in terms of the scale factor as

$$\chi(a) = \frac{c}{H_0} \int_a^1 da' [a' \Omega_m + a'^2 (1 - \Omega_m - \Omega_\Lambda) + a'^4 \Omega_\Lambda]^{-1/2}, \quad (1.29)$$

and as a function of redshift

$$\chi(z) = \frac{c}{H_0} \int_0^z dz' [(1+z')^3 \Omega_m + (1+z')^2 (1 - \Omega_m - \Omega_\Lambda) + \Omega_\Lambda]^{-1/2}. \quad (1.30)$$

Furthermore, the *angular diameter distance* is defined as the ratio of the physical diameter $2R$ of an object and the observed angular diameter δ ,

$$D_{\text{ang}}(z) := \frac{2R}{\delta} = a(z) f_K(\chi), \quad (1.31)$$

where the second step follows from setting $d\theta = \delta$ and $ds = 2R$ in the metric.

1.2 The formation of structures

The expansion of the Universe is mathematically encapsulated in the scale factor a , which relates to temperature as $T \propto a^{-1}$. Since the scale factor increases with time from zero

to one, the temperature drops from formally infinity at the epoch of the Big Bang to the current value of 2.73 K (Fixsen 2009). The functional dependence of temperature with respect to time determines sensitively the formation history of the first structures in the Universe. The temperature determines which particle species are still in thermal equilibrium and which ones have already dropped out, being free to propagate through the Universe or *gravitationally collapse* to the structures observed today: stars, galaxies, dark matter halos, galaxy clusters and the large-scale structure.

1.2.1 The origin of structure

Heisenberg's uncertainty principle implies that the conservation of energy can be ‘violated’ for a short period of time Δt , if a virtual particle-antiparticle pair of energy ΔE is instantaneously created and destroyed again (Heisenberg 1927). These *quantum fluctuations* are believed to have existed also at the very beginning of the Universe, at the time of *Grand Unification*, when energies of $T \approx 10^{14}$ GeV¹ were governing physical processes and matter was in the state known in quantum field theory as vacuum (Sakharov 1966).

These vacuum quantum fluctuations were amplified during a postulated phase of exponential expansion of the Universe, which took place approximately 10^{-34} s after the Big Bang and was driven by the negative-pressure vacuum energy density. Although this phase of *inflation* lasted only a fraction of a second, the space expanded by a factor of 10^{40} , inflating the quantum fluctuations to macroscopic structures, which became the seeds of the large-scale structures one observes today. The theory of inflation is an extension to the standard model of cosmology introduced by Guth (1981) to explain, besides the spectrum of primordial fluctuations, the *flatness* -, the *horizon* - and the *magnetic monopole problem*. A review on inflation can be found in Martin et al. (2014).

The most promising tool to verify inflation is the measurement of the CMB radiation, which is a relic from the epoch of *recombination* and the oldest snapshot of the Universe observable today. When the Universe and thus the plasma of nucleons, electrons and photons cooled down to a temperature of $T \approx 0.3$ eV at $z \approx 1100$, the energy of the photons was no longer sufficient to reionise freshly build atoms. The nucleons and electrons could finally combine and form neutral atoms. From this time of *freeze-out* about 380 000 years after the Big Bang onwards, the CMB photons typically do not interact with baryons anymore, but propagates freely through the Universe until today.

The CMB photons, as a result, still carry information about the distribution of the primordial structures of matter at the time of freeze-out, which is invaluable information with regard to structure formation theories. For example, if a photon resides at the time of freeze-out in an overdense region, it first needs to climb out of the gravitational well of the overdensity, resulting in a redshift of its energy and thus a lower observed temperature compared to an underdense region. This phenomenon is called the Sachs-Wolf effect (Sachs & Wolfe 1967). This radiation from recombination era was first detected by Penzias & Wilson (1965) and measured with an unprecedented precision most recently by the Planck Collaboration et al. (2016a). These measurements confirmed the flatness of the Universe, the existence of fluctuations on scales larger than light could have travelled with the

¹Temperature and energy can be converted into each other with the conversion $1\text{eV} = 1.1605 \times 10^4 k_{\text{B}}\text{K}$.

speed of light until recombination sets in, the existence of adiabatic fluctuations resulting from the thermal equilibrium before recombination, and the Gaussianity of the primordial fluctuations with a slight deviation from a scale-invariant *Harrison-Zel'dovich* spectrum (Harrison 1970, Zeldovich 1972). All of these observations are in excellent agreement with the Λ CDM model and the theory of inflation. Still, one prediction of the theory of inflation has not yet been confirmed: The existence of primordial gravitational waves that should be observable in the B-mode polarisation of the CMB radiation (BICEP2/Keck and Planck Collaborations et al. 2015, Kovac 2018).

1.2.2 Evolution of the density fluctuations

The homogeneous and isotropic world model introduced in Sect. 1.1 considers so far only a Universe with mean density $\bar{\rho}(t)$, where the value of the latter determines the expansion behaviour of the Universe. The presence and evolution of density fluctuations can, however, be integrated into the Λ CDM model by treating them as a perturbation to the mean density.

Using comoving coordinates to express the deviation from the otherwise homogeneous expansion, and denoting the density at the spatial comoving coordinate \mathbf{x} at time t by $\hat{\rho}(\mathbf{x}, t)$, a density perturbation is defined by the *density contrast*,

$$\delta(\mathbf{x}, t) := \frac{\hat{\rho}(\mathbf{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \quad (1.32)$$

which is a measure for the relative amplitude of the perturbation. Ideally, one could derive an accurate analytical description of the density field in terms of the density contrast $\delta(\mathbf{x}, t)$ over the course of time, explaining the growth of the primordial structures to today's massive perturbations, e.g. in form of galaxies. This is, however, not possible due to the range of scales and the non-linearity of physics involved, so that one needs to rely on approximate methods: perturbation theory, numerical simulations and the halo model.

Although *perturbation theory* offers the preferred analytical description of the evolution of density perturbations, it is limited to the linear ($|\delta| \ll 1$), maximally mildly non-linear regime (for a review see Bernardeau et al. 2002). Considering small scales or objects like galaxies or galaxy clusters, one enters the non-linear regime ($|\delta| \gtrsim 1$) and *numerical simulations* need to be performed to describe the growth of structures. Simulations, however, have the drawback of being expensive with respect to computing time and limited in resolution. Further, although dark matter simulations are a standard tool today (Springel et al. 2005), comparable hydrodynamical simulations are just starting to become available (Vogelsberger et al. 2014, Schaye et al. 2015).

Built on the findings of perturbation theory, dark matter simulations and observations, the *halo model* provides not only an analytical description, but as well a physical interpretation of the distribution and evolution of dark and luminous structures in the Universe on all scales. This work explores the limitations of the halo model, which will be introduced in Chapter 2, therefore a short introduction over equations from perturbation theory entering the halo model is given in the next paragraphs.

Linear perturbation theory

In this subsection the equations governing the evolution of density fluctuations in the linear regime are introduced. To this end, several assumptions are made. Firstly, although in general structure formation needs to be treated in the framework of GR, the restriction of the analysis to scales smaller than the *comoving horizon size*,

$$d_{\text{H}}(a) = \int_0^a \frac{c da'}{a'^2 H(a')}, \quad (1.33)$$

allows the application of Newtonian physics instead. Secondly, for now the era of matter-domination is considered where pressure is negligible and dark matter can be treated as a perfect fluid² with density $\hat{\rho}(\mathbf{x}, t)$ and peculiar velocity $\mathbf{v}(\mathbf{x}, t)$.

The density field can then be described by a triplet of equations: the *continuity equation*, which expresses the conservation of mass, the *Euler equation*, which accounts for the conservation of momentum, and the *Poisson equation*, which relates the density contrast to the gravitational potential $\Phi(\mathbf{x}, t)$. This triplet of non-linear, coupled partial differential equations can in general not be solved analytically, but only numerically. An analytical solution can be found if one assumes the perturbations in δ and $|\mathbf{v}|$ to be small ($\delta \ll 1$). Then the fluid equations can be linearised in these quantities resulting in the linearised continuity, Euler and Poisson equation:

$$\begin{aligned} \frac{\partial \delta(\mathbf{x}, t)}{\partial t} + \frac{1}{a} \nabla_{\mathbf{x}} \cdot \mathbf{v}(\mathbf{x}, t) &= 0, \\ \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} + \frac{\dot{a}}{a} \mathbf{v}(\mathbf{x}, t) &= -\frac{1}{a} \nabla_{\mathbf{x}} \Phi(\mathbf{x}, t), \\ \nabla_{\mathbf{x}}^2 \Phi(\mathbf{x}, t) &= \frac{3H_0^2 \Omega_{\text{m}}}{2a} \delta(\mathbf{x}, t). \end{aligned} \quad (1.34)$$

The three equations can be combined into a single second-order differential equation for the density contrast,

$$\frac{\partial^2 \delta(\mathbf{x}, t)}{\partial t^2} + \frac{2\dot{a}}{a} \frac{\partial \delta(\mathbf{x}, t)}{\partial t} - \frac{3H_0^2 \Omega_{\text{m}}}{2a^3} \delta(\mathbf{x}, t) = 0, \quad (1.35)$$

which no longer explicitly contains the spatial coordinate \mathbf{x} or derivatives of \mathbf{x} . Consequently, the shape of a given perturbation is preserved in the linear regime and only the density contrast changes. Therefore, the solution can be split into a temporal part $D(t)$ and a spatial part $\Delta(\mathbf{x})$,

$$\delta(\mathbf{x}, t) = D_-(t)\Delta_-(\mathbf{x}) + D_+(t)\Delta_+(\mathbf{x}), \quad (1.36)$$

with one decaying (−) and one growing (+) solution. For an Einstein-de Sitter model ($\Omega_{\text{m}} = 1, \Omega_{\Lambda} = 0$) it turns out that $D_-(t) \propto a^{-3/2}(t)$ is monotonically decreasing with time and is thus less relevant for the formation of structures. However, $D_+(t)$ is in this case

²Dark matter is collisionless so that multi-streams can occur. Although important on small scales, on large-scale this effect can be neglected and the assumption of a perfect fluid is justified.

Table 1.1: Growth behaviour of density perturbations within linear perturbation theory before matter-radiation equality, between matter-radiation equality and recombination, and after recombination. See text for details.

Epoch	Perturbation size	Dark matter	Radiation	Baryons
$a < a_{\text{eq}}$	$\lambda > d_{\text{H}}$	$\delta_{\text{dm}} \propto a^2$	$\delta_{\text{r}} \propto a^2$	$\delta_{\text{b}} \propto a^2$
	$\lambda < d_{\text{H}}$	constant	oscillating	oscillating
$a_{\text{eq}} < a < a_{\text{rec}}$	all λ	$\delta_{\text{dm}} \propto a$	oscillating	oscillating
$a_{\text{rec}} < a$	all λ	$\delta_{\text{dm}} \propto a$	oscillating	$\delta_{\text{b}} \propto a$

proportional to the scale factor $a(t)$ and describes the evolution of growing perturbations as a function of time, resulting in the name *linear growth factor*.

For a general cosmology the functional dependencies on the scale factor are more complex but the general behaviour is preserved. It can be shown that the linear growth factor is given by

$$D_+(a) \propto \mathcal{D}_+(a) = \frac{5\Omega_{\text{m}}}{2} \frac{H(a)}{H_0} \int_0^a \frac{da'}{[\Omega_{\text{m}}/a' + \Omega_{\Lambda}a'^2 - (\Omega_{\text{m}} + \Omega_{\Lambda} - 1)]^{3/2}}. \quad (1.37)$$

The constant of proportionality is set by normalising $\mathcal{D}_+(a)$ to unity for today, such that

$$D_+(a) = \frac{\mathcal{D}_+(a)}{\mathcal{D}_+(a=1)}. \quad (1.38)$$

For a general Λ CDM cosmology the integral can not be solved analytically. However, in this work the fitting function by Carroll et al. (1992),

$$D_+(a) = \frac{5}{2} a \Omega_{\text{m}}(a) \left[\Omega_{\text{m}}^{4/7}(a) - \Omega_{\Lambda}(a) + \left(1 + \frac{\Omega_{\text{m}}(a)}{2} \right) \left(1 + \frac{\Omega_{\Lambda}(a)}{70} \right) \right]^{-1}, \quad (1.39)$$

is used. Note that the density parameters $\Omega_{\text{m}}(a)$ and $\Omega_{\Lambda}(a)$ are explicitly scale-factor dependent (Eqs. 1.19 and 1.20) and that the growth factor $D_+(a)$ is dimensionless.

The growth of other density perturbations in the linear regime, as summarised in Table 1.1, is now discussed in a qualitative way. For the *radiation-dominated era* ($a < a_{\text{eq}}$) GR can not be neglected and the continuity, Euler and Poisson equation need to be written in their relativistic form. Radiation rules the expansion behaviour of the Universe according to Eq. (1.25) and has to be accounted for in the Poisson equation. Perturbations can be either larger than the horizon or smaller. In the former case physical interactions are not possible and, thus, pressure does not play a significant role compared to gravity. From the fluid equations it follows that all three species of perturbations (radiation, baryonic matter and dark matter) behave the same and grow like $\delta \propto a^2$. However, if a perturbation is smaller than the horizon, or enters the horizon during the radiation-dominated phase, the competition between gravity and pressure gives rise to oscillations in the coupled photon-baryon fluid. In the case of sub-horizon dark matter perturbations, the dark matter overdensities are independent of the particle pressure, but their growth is nevertheless

suppressed due to the strong expansion behaviour of the Universe during the radiation-dominated era, which counteracts the effect of gravity. The dark matter structure growth is suppressed until matter-radiation equality is reached at a_{eq} , corresponding to a horizon size of $d_{\text{H}}(a_{\text{eq}}) \approx 16(\Omega_{\text{m}}h)^{-2} \text{Mpc}$. All perturbations smaller than this threshold size are suppressed.

After *matter-radiation equality* the expansion behaviour of the Universe changes to that of Eq. (1.27). Although oscillations continue in the still relativistic photon-baryon fluid, dark matter perturbations feel the decrease in the expansion rate of the Universe and resume growing with $\delta \propto a$. Only at *recombination*, baryons are released from the photon pressure and can grow in the same way as dark matter with $\delta \propto a$, marking the start of a correlated evolution.

It is convenient to formulate the scale-dependent effects that density perturbations undergo during horizon crossing and matter-radiation transition in Fourier-space and summarise them in the *transfer function* $T(\mathbf{k})$. With the Fourier-space analogue of the length λ being given by the length $k = 2\pi/\lambda$ of the *comoving wave vector* \mathbf{k} , the *Fourier transform* of the real space density $\delta(\mathbf{x}, t)$ is defined as

$$\tilde{\delta}(\mathbf{k}, t) = \int_{\mathbb{R}^3} d^3x \delta(\mathbf{x}, t) e^{-i\mathbf{x}\cdot\mathbf{k}}. \quad (1.40)$$

Let a_i denote the scale factor at a time t_i in the radiation-dominated phase. Further, let $\tilde{\delta}(\mathbf{k}_{\text{s}}, a_i)$ be a small-scale perturbation that enters the horizon at a_i , and $\tilde{\delta}(\mathbf{k}_{\text{l}}, a_i)$ a large-scale perturbation that enters the horizon only later in the matter-dominated era. Under the assumption that both perturbations evolved linearly until today to $\tilde{\delta}(\mathbf{k}_{\text{s}}, a_0)$ and $\tilde{\delta}(\mathbf{k}_{\text{l}}, a_0)$, respectively, the suppression of the former compared to the latter is captured by the transfer function $T(\mathbf{k})$, which is defined as

$$\frac{\tilde{\delta}(\mathbf{k}_{\text{s}}, a_0)}{\tilde{\delta}(\mathbf{k}_{\text{l}}, a_0)} \equiv T(\mathbf{k}) \frac{\tilde{\delta}(\mathbf{k}_{\text{s}}, a_i)}{\tilde{\delta}(\mathbf{k}_{\text{l}}, a_i)}. \quad (1.41)$$

By construction $T(k) \rightarrow 1$ for $k \rightarrow 0$, resembling the observed qualitative behaviour. Fitting functions have been derived by, e.g. [Bardeen et al. \(1986\)](#) and [Eisenstein & Hu \(1998\)](#). In this work the one of [Eisenstein & Hu \(1998\)](#) is adopted, which describes the growth of cold dark matter perturbations in the presence of baryons and vice versa,

$$T(k) = \frac{\Omega_{\text{c}}}{\Omega_{\text{m}}} T_{\text{c}}(k) + \frac{\Omega_{\text{b}}}{\Omega_{\text{m}}} T_{\text{b}}(k). \quad (1.42)$$

Here Ω_{m} is the sum of the baryonic matter density Ω_{b} and the cold dark matter density Ω_{c} . A summary of all the equations necessary for the numerical implementation of the transfer function by [Eisenstein & Hu \(1998\)](#) is given in [Appendix B](#).

Dark matter

In the preceding discussion of linear structure growth the existence of *dark matter* was simply assumed. In this subsection the evidence for its existence as well as its influence on structure formation is briefly reviewed.

The evidence for the existence of dark matter is manifold: the flat rotation curves of spiral galaxies (Sofue & Rubin 2001), the velocity dispersions of elliptical galaxies (Saglia et al. 1993), the kinematics of galaxies in galaxy clusters (Zwicky 1933), the hot X-ray gas in galaxy clusters (Allen et al. 2011), the Baryonic Acoustic Oscillations (BAO) (Eisenstein et al. 2005), and the gravitational lensing effect (Schneider 2006). However, its nature remains unknown (Roszkowski et al. 2018), although a purely baryonic explanation in the form of MASSive Compact Halo Objects (MACHOs) has already been excluded (Tisserand et al. 2007). Assuming GR is the correct theory of gravity, one is left with elementary particles as candidates, where the options are diverse but none have so far been detected (for a review see Bertone et al. 2005). A scenario based on exclusively Hot Dark Matter (HDM) particles, which were still relativistic at the time of their decoupling, can be ruled out, because it fails to reproduce the assumed bottom-up formation of structures. Cold Dark Matter (CDM) particles on the other hand do support hierarchical structure formation. They have a low number density, but large particle mass - the last characteristic captured in the term for the most probable dark matter candidate, the Weakly Interacting Massive Particle (WIMP).

Cold dark matter could also explain why the observed temperature fluctuations in the CMB are of the order of 10^{-5} (Planck Collaboration et al. 2016a), although according to linear perturbation theory the fluctuations at the time of recombination must have been of the order of 10^{-3} to obtain the structures we observe today (Wilson & Silk 1981, Wilson 1983, Bond & Efstathiou 1984). Interacting only gravitationally and weakly, cold dark matter decoupled early from thermal equilibrium and gravitationally collapsed before the release of the CMB radiation, which allowed dark matter overdensities to grow to an amplitude of the order of 10^{-3} until recombination. Although the baryons get attracted by this gravitational potential, the radiation pressure of the photons prevents them from falling in. This counterplay of gravity and pressure is today still visible as the baryonic acoustic oscillation features in the CMB power spectrum and the large-scale galaxy distribution. Only after recombination is the photon pressure removed and the baryons fall into the potential wells of the dark matter overdensities. The perturbations in the CMB with an amplitude of 10^{-5} must therefore be evaluated as the density contrast of baryons at recombination, which is the last information the photons received and carried to us.

The importance of dark matter in our Universe regarding structure formation extends beyond the possibility to explain the amplitude of the CMB temperature fluctuations. Dark matter is assumed to also play an important role in the formation of stars and galaxies at later times. The halo model, e.g. is based on the assumption that galaxies can only form and live inside a dark matter environment, as will be discussed in Chapter 2.

1.3 The Universe – a random field

Perturbation theory can describe the evolution of linear density perturbations in an expanding Universe and predict the density field at a certain time, if knowledge on the initial conditions is given. However, the specific conditions of the initial density field right after inflation can neither be measured nor theoretically predicted. Inflationary models can only predict the statistical properties of the initial density field based on the assumption that

the seeds of the perturbations in the primordial density field, the quantum fluctuations, are generated in a random process. Then the density field in our Universe can be treated as one possible realisation $g(\mathbf{x})$ of a *random field*. The statistical properties of a random field are characterised by a formally infinite series of moments, the *n-point correlation functions* or their Fourier space analoga, the *spectra*.

Since this work deals with two- and three-point correlation functions of gravitational lensing, relevant definitions regarding correlation functions and spectra are given in Sects. 1.3.1 and 1.3.2, respectively. This summary is based on the review on statistical measures related to random fields by [Bernardeau et al. \(2002\)](#).

1.3.1 Correlation functions

The infinitesimal joint probability of one particular realisation of a random field to occur, i.e. that the random field takes at the n considered positions in the field the values $g(\mathbf{x}_n)$ within the intervals $dg(\mathbf{x}_n)$, is given by

$$dP = P[g(\mathbf{x}_1), \dots, g(\mathbf{x}_n)] dg(\mathbf{x}_1) \dots dg(\mathbf{x}_n). \quad (1.43)$$

The statistical properties of the underlying random field are then fully specified by all k moments ($k \rightarrow \infty$) of the random field, which are defined by the average over the ensemble:

$$\langle g(\mathbf{x}_1) \dots g(\mathbf{x}_k) \rangle = \int dg(\mathbf{x}_1) \dots dg(\mathbf{x}_k) g(\mathbf{x}_1) g(\mathbf{x}_2) \dots g(\mathbf{x}_k) P[g(\mathbf{x}_1), \dots, g(\mathbf{x}_k)]. \quad (1.44)$$

In practice, neither the whole ensemble of stochastic realisations can be accessed, since there is only our Universe to observe. Nor can the formally infinite number of moments k be computed, but currently because of computational limitations more realistically $k \leq 4$. For the former problem the principle of *ergodic fields* can be applied, where the ensemble average is replaced by a volume average over different fields on the sky. The latter limitation is less dramatic for early times, since the *central limit theorem* states that a large number of random processes results in a *Gaussian field*. The latest measurement of the [CMB](#) radiation confirmed that the temperature distribution at the time of recombination is nearly Gaussian ([Planck Collaboration et al. 2016c](#)). Then the probability distribution can be described by a multi-variate Gaussian:

$$P[g(\mathbf{x}_1), \dots, g(\mathbf{x}_k)] = \frac{1}{\sqrt{(2\pi)^n \det(\mathcal{C})}} \exp \left(-\frac{1}{2} \sum_{i,j=1}^n g(\mathbf{x}_i) \mathcal{C}_{ij}^{-1} g(\mathbf{x}_j) \right), \quad (1.45)$$

with $\mathcal{C}_{ij} = \langle g(\mathbf{x}_i) g(\mathbf{x}_j) \rangle$ being the covariance matrix. An advantage of a Gaussian random field is, that it is completely determined by the second moment, the *two-point correlation function*:

$$\xi(\mathbf{x}_1, \mathbf{x}_2) = \langle g(\mathbf{x}_1) g^*(\mathbf{x}_2) \rangle. \quad (1.46)$$

Deviations from Gaussianity can only be detected by higher-order correlation functions. The lowest-order probe of non-Gaussianities is the *three-point correlation function*,

$$\zeta(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \langle g(\mathbf{x}_1) g(\mathbf{x}_2) g(\mathbf{x}_3) \rangle. \quad (1.47)$$

However, the computation of the three-point correlation function from observations or simulations is computationally expensive. Therefore, the four-point correlation function:

$$\eta(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \langle g(\mathbf{x}_1) g(\mathbf{x}_2) g(\mathbf{x}_3) g(\mathbf{x}_4) \rangle, \quad (1.48)$$

is just becoming state of the art in cosmology. For homogeneous random fields the two-point correlation function depends only on the separation $|\mathbf{x}_1 - \mathbf{x}_2|$, the three-point correlation function on two vectors representing two sides of the spanned triangle $\mathbf{x}_{12} = \mathbf{x}_1 - \mathbf{x}_2$ and $\mathbf{x}_{13} = \mathbf{x}_1 - \mathbf{x}_3$, and the four-point correlation function on three vectors $\mathbf{x}_{12}, \mathbf{x}_{23}, \mathbf{x}_{34}$ forming the quadrangle.

1.3.2 Spectra

For many cosmological analyses it is convenient to work in Fourier space. In Chapter 3 it will be demonstrated how an analytical expression for galaxy-galaxy(-galaxy) lensing can be obtained from the Fourier transform of the second- and third-order correlation function: the power spectrum and bispectrum, respectively, which shall be introduced in this subsection.

With the Fourier transform of a random realisation $g(\mathbf{x})$ being defined by

$$g(\mathbf{k}) = \int d^n x e^{-i\mathbf{x}\cdot\mathbf{k}} g(\mathbf{x}), \quad (1.49)$$

the transformation of the two-point correlation function (Eq. 1.46) to Fourier space gives

$$\begin{aligned} \langle g(\mathbf{k}) g^*(\mathbf{k}') \rangle &= \int d^n x e^{-i\mathbf{x}\cdot\mathbf{k}} \int d^n x' e^{i\mathbf{x}'\cdot\mathbf{k}'} \langle g(\mathbf{x}) g^*(\mathbf{x}') \rangle \\ &= (2\pi)^n \delta_{\text{D}}(\mathbf{k} - \mathbf{k}') \int d^n y e^{i\mathbf{y}\cdot\mathbf{k}} \xi(|\mathbf{y}|) \\ &= (2\pi)^n \delta_{\text{D}}(\mathbf{k} - \mathbf{k}') P(|\mathbf{k}|), \end{aligned} \quad (1.50)$$

where in the second step the substitution $\mathbf{x}' = \mathbf{x} + \mathbf{y}$ is made. The quantity $P(|\mathbf{k}|)$ is the Fourier transform of the two-point correlation function, the *power spectrum*

$$P(|\mathbf{k}|) = \int d^n y e^{-i\mathbf{y}\cdot\mathbf{k}} \xi(|\mathbf{y}|), \quad (1.51)$$

which is a statistical measure for the power on the length scale $2\pi/k$.

Analogously, the Fourier transform of the three-point correlation function (Eq. 1.47) is defined by the relation

$$\langle g(\mathbf{k}_1) g(\mathbf{k}_2) g(\mathbf{k}_3) \rangle = (2\pi)^n \delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \quad (1.52)$$

with $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ defining a triangle. The *bispectrum* is then given by

$$B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \int d^n y \int d^n z e^{-i\mathbf{y}\cdot\mathbf{k}_1} e^{-i\mathbf{z}\cdot\mathbf{k}_2} \zeta(\mathbf{y}, \mathbf{z}). \quad (1.53)$$

1.3.3 Dark matter spectra from perturbation theory

For dark matter, the linear power spectrum and bispectrum can be derived analytically using linear perturbation theory. The two are ingredients of the halo model (Chapter 2), which allows for a derivation of analytical expressions for the power spectra and bispectra in the non-linear regime, exceeding the scope of perturbation theory covered in this chapter.

The power spectrum for a matter perturbation at scale factor a that linearly evolved through time can be obtained from the *primordial power spectrum* $P_{\text{prim}}(k)$. To that end, the growth factor is applied to account for the expansion of the Universe, and the transfer function to incorporate the scale-dependent effects that density perturbations undergo during horizon crossing and matter-radiation transition:

$$P_{\text{lin}}(k, a) = A D_+^2(a) T_k^2 P_{\text{prim}}(k), \quad (1.54)$$

where A is a constant of normalisation. The derivation of an analytical expression for the normalisation A can be found in Appendix C. The primordial power spectrum is set by inflation. It is assumed that right after inflation all perturbations have been larger than the horizon, meaning that no interactions occur. Thus, there is no preferred length scale in the Universe at that time and the primordial spectrum must be of the form of a power-law $P_{\text{prim}} \propto k^{n_s}$, where the constant $n_s \lesssim 1$ is the spectral index (Table A.1). For $n_s = 1$ one speaks of a *Harrison-Zel'dovich spectrum* (Harrison 1970, Zeldovich 1972).

For a derivation of the lowest order non-vanishing bispectrum term from perturbation theory see, e.g. Cooray & Sheth (2002). Here only the result is given:

$$B_{\text{lin}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P_1P_2 + 2F_2(\mathbf{k}_1, \mathbf{k}_3)P_1P_3 + 2F_2(\mathbf{k}_2, \mathbf{k}_3)P_2P_3, \quad (1.55)$$

with $P_i = P_{\text{lin}}(k_i)$ and the second-order coupling function being defined by Jain & Bertschinger (1994) as

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{2(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{7k_1^2k_2^2} + \frac{1}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right), \quad (1.56)$$

and $F_2(\mathbf{k}, -\mathbf{k}) = 0$.

Chapter 2

The halo model

In the '50s Newman & Scott proposed to describe galaxy clustering by the clustering of discrete matter clumps with an universal radial profile (Neyman & Scott 1952; Neyman et al. 1953). Without being aware of it, they laid the foundation for the *halo model* description of the at the time unknown dark matter distribution. As shown in Fig. 2.1, the complex filamentary distribution of dark matter can be simplified by assuming that all dark matter in the Universe is bound in dark matter halos. This ansatz discretises the problem and splits it in a natural way into the clustering of dark matter halos on large-scales and the distribution of mass within the halos on small scales. Further, it allows an analytical description of all n -point correlations from small to large-scales, from the linear to the non-linear regime, and for dark and visible matter. For a general review see Cooray & Sheth (2002).

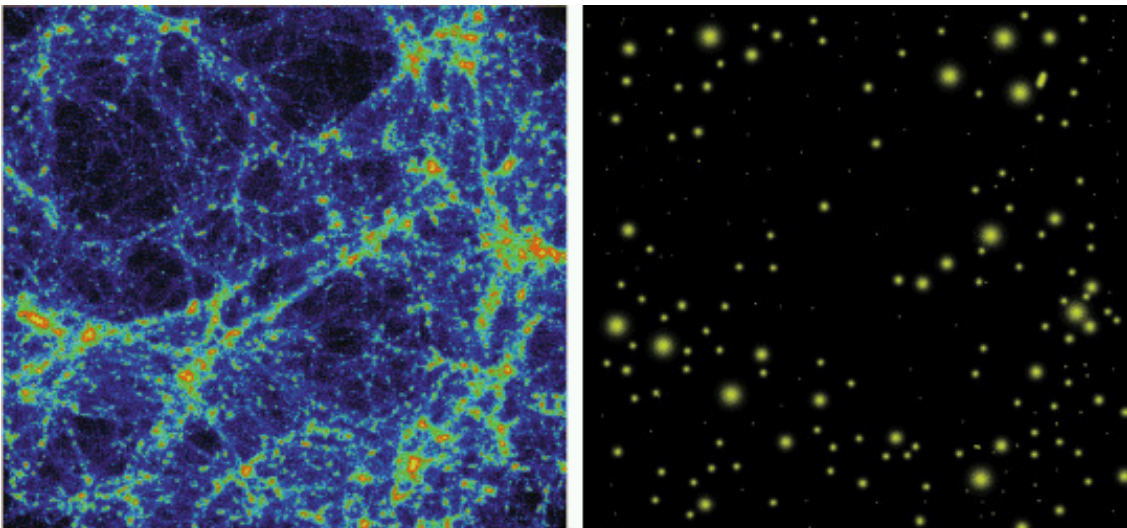


Figure 2.1: The complex distribution of dark matter can be discretised within the framework of the halo model and is then fully described by the number of halos of a certain mass, the density profile of the halos and their clustering. [Source: Cooray & Sheth (2002)]

Twenty five years after the work by Newman & Scott, [White & Rees \(1978\)](#) extended the theory by include galaxies in the dark matter halo model framework based on the physical assumption that galaxies can only form in dark matter halos, because baryonic matter needs the gravitational potential of the dark matter overdensities for radiative gas cooling and star formation. In the halo model picture this translates to the *halo occupation distribution* (HOD): once a dark matter halo has virialised and crossed a certain threshold mass, it can host a first (central) galaxy. While if the dark matter halo is massive enough, more (satellite) galaxies will populate the halo, where the number of galaxies per halo depends solely on the halo mass.

Although this simplified analytical description neglects the potential influence of the halo environment and the halo history (‘assembly bias’; e.g. [Gao & White 2007](#), [Zentner et al. 2014](#), [Montero-Dorta et al. 2017](#), [Mao et al. 2018](#)), its predictions fit the results obtained from dark matter N -body simulations of cosmological structure formation on second-order level remarkably well ([Kravtsov et al. 2004](#), [Zheng et al. 2005](#)). More notably, it also explains the observed distribution of galaxies on the same statistical level ([Zheng et al. 2007](#), [Zehavi et al. 2011](#), [Coupon et al. 2012](#)), as well as the cross-correlation of galaxies and dark matter ([Leauthaud et al. 2012](#), [Velandar et al. 2014](#), [Coupon et al. 2015](#), [Clampitt et al. 2017](#)). The HOD formalism – being based on a physical motivation – offers a unique opportunity to learn about the coupled, but not identical evolution of baryonic and dark matter overdensities, without the need to take all the highly non-trivial hydrodynamical processes into account.

However, at third-order level and beyond, the validity of the halo model in combination with the HOD formalism is not that clear anymore. Although, [Scoccimarro et al. \(2001\)](#) and [Wang et al. \(2004\)](#) have shown a reasonable match between dark matter halo model predictions and results obtained from dark matter N -body simulations, [Lazanu et al. \(2017\)](#) found that although the halo model description delivers satisfying results in the highly non-linear regime, it underestimates the power on intermediate scales, particularly at higher redshifts, when compared to dark matter simulations. The search for optimised halo models has thus started ([Mead et al. 2015](#), [Lazanu et al. 2017](#)).

This work is the first of its kind to investigate the scope of the halo model to describe the galaxy-dark matter cross-correlation on second- and third-order level consistently. It is not the aim of this work to test the most advanced halo models (e.g. [Leauthaud et al. 2011](#)), but rather a ‘standard’ one (e.g. [Kravtsov et al. 2004](#), [Zheng et al. 2007](#), [Coupon et al. 2012](#)). Besides some modifications, the halo model employed in this work is largely motivated by and based on the work of [Rödiger \(2009\)](#).

In this chapter the employed halo model is introduced and analytical expressions for the three-dimensional halo model galaxy-matter power spectrum and bispectrum are given. It is shown how these quantities are projected to their two-dimensional analogues, which are ultimately transformed into galaxy-galaxy lensing (G2L) and galaxy-galaxy-galaxy lensing (G3L) as described in Chapter 3. All equations given in this chapter are implemented into a numerical code which was written as part of this work to generate halo model predictions for the two-dimensional galaxy-matter power spectrum and bispectrum.

2.1 The dark matter halo model

A dark matter halo is defined by the criterion for a spherical overdense region of dark matter to collapse and virialise, namely that its density reaches 178 times the critical density of the universe $\rho_{\text{crit}}(z)$ (Eq. 1.17) at the time of virialisation (Gunn & Gott 1972, Press & Schechter 1974). This threshold value is typically translated into the density contrast δ (Eq. 1.32) of the region at the time of virialisation, assuming that the initial overdensity had evolved according to linear theory, i.e. $\delta_{\text{lin}}^{(\text{EdS})} = 1.686$. Both values are derived for the case of an Einstein-de-Sitter universe but generally depend on cosmology and redshift. Here the fitting formula by Henry (2000) is used as a correction to the Einstein-de-Sitter values, providing expressions for the density ratio $\Delta_c(z)$ and the linear density contrast $\delta_c(z)$ at virialisation:

$$\begin{aligned}\Delta_c(z) &= \Delta^{(\text{EdS})} (1 + 0.4093 x^{2.71572}) , \\ \delta_c(z) &= \delta_{\text{lin}}^{(\text{EdS})} [1 - 0.0123 \ln(1 + x^3)] ,\end{aligned}\tag{2.1}$$

with $x = (\Omega_m^{-1} - 1)^{1/3} / (1 + z)$. Given the definition of a halo, the overall distribution of dark matter is fully described by the comoving number density $n(M_h, z)$ of dark matter halos of mass M_h at redshift z , by the radial distribution of dark matter within the halos $\rho(r|M_h)$, and by the clustering of the halos given by the dark matter halo bias $b^h(M_h)$.

2.1.1 Halo mass function

Acting like a census, the halo mass function $n(M_h, z)$ gives the comoving number density of halos with masses in the range M_h and $M_h + dM_h$ at redshift z (Bond et al. 1991):

$$n(M_h, z) = \frac{\bar{\rho}}{M_h^2} \nu f(\nu) \frac{d \ln \nu}{d \ln M_h} ,\tag{2.2}$$

with $\bar{\rho} = \rho_{m,0}$ denoting the current mean background density. The so-called multiplicity function $\nu f(\nu)$ accounts for the finding from numerical dark matter simulations that the halo mass function has a universal shape. The dimensionless variable ν is defined as

$$\nu = \frac{\delta_c(z)}{D_+(z) \sigma(M_h, z)} ,\tag{2.3}$$

where the growth factor $D_+(z)$ is defined by Eq. (1.38) and the variance $\sigma(M_h, z)$ by Eq. (C.1). Assuming an ellipsoidal mass collapse, Sheth & Tormen (1999) found that $\nu f(\nu)$ can be expressed as,

$$\nu f(\nu) = A \sqrt{\frac{2}{\pi}} [1 + (q\nu^2)^{-p}] \sqrt{q\nu^2} \exp\left(\frac{-q\nu^2}{2}\right) ,\tag{2.4}$$

with the amplitude

$$A = \left[1 + 2^{-p} \Gamma\left(\frac{1}{2} - p\right) / \sqrt{\pi}\right]^{-1} .\tag{2.5}$$

Here Γ is the gamma function. The parameters $q = 0.707$ and $p = 0.3$ fit the results of N -body simulations best. The amplitude A is fixed by the restriction that

$$\frac{1}{\bar{\rho}} \int_0^\infty n(M_h, z) M_h dM_h = \int_0^\infty f(\nu) d\nu = 1. \quad (2.6)$$

In practice the integrations over mass cannot be performed over the whole mass range up to infinity, but reasonable assumptions on the integration limits can be made. In this work the mass of dark matter halos is assumed to lie within the range of $0 - 10^{16} h^{-1} M_\odot$, with the upper limit exceeding the highest measured cluster masses (Buddendiek et al. 2015, and references therein) and the highest halo masses from dark matter only and hydrodynamical simulations (Warren et al. 2006, Tinker et al. 2008, Vogelsberger et al. 2014). All integrals describing dark matter, or the mixed case of galaxies and dark matter, are computed with these limits. In contrast, integrals which are concerned with the mass of galaxies only, are computed for the range $10^{10} - 10^{15} h^{-1} M_\odot$, because the threshold mass for halos to host a first galaxy is approximately $10^{11} h^{-1} M_\odot$ (e.g. Zheng et al. 2005). An analytical derivation of $\frac{d \ln \nu}{d \ln M_h}$ appearing in Eq. (2.2) is given in Appendix D.

2.1.2 Halo density profile

In this work the density profile of dark matter halos is assumed to follow a Navarro-Frenk-White (NFW) profile (Navarro et al. 1996),

$$\rho(r|M_h) = \frac{\rho_s}{(r/r_s)[1 + (r/r_s)]^2}, \quad (2.7)$$

where ρ_s is the central density parameter and r_s the scale radius. Integrating $\rho(r|M_h)$ up to the radius where the halo is in virial equilibrium, i.e. the virial radius

$$r_{\text{vir}}(z) = \left[\frac{3 M_h}{4\pi \Delta_c(z) \bar{\rho}} \right]^{1/3}, \quad (2.8)$$

gives the total mass of a halo (e.g. Takada & Jain 2003):

$$M_h = \int_0^{r_{\text{vir}}} dr 4\pi r^2 \rho(r|M_h) = \frac{4\pi \rho_s r_{\text{vir}}^3}{c^3} \left[\ln(1+c) - \frac{c}{1+c} \right], \quad (2.9)$$

with the concentration parameter $c = r_{\text{vir}}/r_s$. Bullock et al. (2001) found from dark matter simulations the following fitting formula for c for NFW halos:

$$c(M_h, z) = \frac{c_0}{1+z} \left[\frac{M_h}{m_\star(z=0)} \right]^{-\alpha}, \quad (2.10)$$

with the parameters $c_0 = 9$ and $\alpha = 0.13$. The mass m_\star is defined by $\sigma^2(m_\star) := \delta_c^2(z=0)$, with the variance given by Eq. (C.1), and the linear density contrast in Eq. (2.1).

For what follows, it is convenient to define the dimensionless normalised dark matter halo density profile in Fourier space,

$$\tilde{u}(\mathbf{k}, M_h) = \frac{\int d^3r \rho(\mathbf{r}, M_h) e^{i\mathbf{k}\cdot\mathbf{r}}}{\int d^3r \rho(\mathbf{r}, M_h)}, \quad (2.11)$$

which can be expressed as (Scoccimarro et al. 2001)

$$\begin{aligned}\tilde{u}(k, M_h) &= f(c) (\sin(\eta) [\text{Si}(\eta(1+c)) - \text{Si}(\eta)]) \\ &+ f(c) (\cos(\eta) [\text{Ci}(\eta(1+c)) - \text{Ci}(\eta)]) \\ &- f(c) \left[\frac{\sin(\eta c)}{\eta(1+c)} \right],\end{aligned}\quad (2.12)$$

with the dimensionless quantity $\eta = (k r_{\text{vir}})/c$ and the dimensionless function for a NFW profile

$$f(c) = \left[\ln(1+c) - \frac{c}{1+c} \right]^{-1}.\quad (2.13)$$

The sine- and cosine-integral are defined as

$$\text{Si}(x) = \int_0^x dt \frac{\sin t}{t} \quad \text{and} \quad \text{Ci}(x) = - \int_x^\infty dt \frac{\cos t}{t}.\quad (2.14)$$

2.1.3 Dark matter halo bias

The clustering of dark matter halos is different from the clustering of the general dark matter density field as a consequence of the existence of a threshold density contrast δ_c to form a halo (e.g. Mo & White 1996). The scale-dependent difference is described by the halo bias functions $b_i^h(M_h)$, which are defined as the coefficients of a power-series expansion of the density contrast of halos $\delta_h(\mathbf{x}; M_h)$ as a function of the dark matter density contrast $\delta(\mathbf{x})$ (e.g. Mo et al. 1997):

$$\delta_h(\mathbf{x}; M_h) = b_1^h(M_h)\delta(\mathbf{x}) + \frac{b_2^h(M_h)}{2}\delta^2(\mathbf{x}) + \frac{b_3^h(M_h)}{6}\delta^3(\mathbf{x}) + \dots\quad (2.15)$$

For this work only the first order bias factor is considered, which is for a Sheth-Tormen mass function given by (Scoccimarro et al. 2001)

$$b_1^h(M_h) = 1 + \epsilon_1 + E_1,\quad (2.16)$$

with

$$\epsilon_1 = \frac{q\nu^2 - 1}{\delta_c} \quad \text{and} \quad E_1 = \frac{2p}{1 + (q\nu^2)^p} \frac{1}{\delta_c}.\quad (2.17)$$

2.2 The halo occupation distribution

The HOD provides an analytical extension to the dark matter halo model with respect to a description for the distribution of galaxies within dark matter halos (Cooray & Sheth 2002). The combination of the dark matter halo model and the HOD formalism, referred to as the *halo model*, is one among several approaches to modelling the galaxy-halo connection; for a general review see Wechsler & Tinker (2018).

2.2.1 Modelling the relation between dark matter halos and galaxies

The HOD is fully specified by, first, the probability distribution $P(N|M_h)$ that a dark matter halo of mass M_h contains N galaxies, second, by the radial distribution of these N galaxies within the host halo, and, third, by the velocity dispersion of these N galaxies (Berlind & Weinberg 2002).

It is commonly assumed that the normalised distribution of galaxies in Fourier space follows the one of dark matter, i.e. $\tilde{u}_g(k, M_h) = \tilde{u}_{\text{dm}}(k, M_h) \equiv \tilde{u}(k, M_h)$ given by Eq. (2.12). However, since deviations and a dependence on redshift have been found (e.g. Budzynski et al. 2012; Muzzin et al. 2007; van der Burg et al. 2014), in this work an additional free parameter was introduced to allow a deviation of the galaxy density profile from the one of dark matter. Nevertheless, the model fits to the CFHTLenS G2L indicate a preference for the standard assumption of identical profiles (Sect. 4.3.2).

Another standard assumption made in this work is that the number of galaxies N hosted by a halo of mass M_h depends solely on the halo mass, albeit in reality $P(N|M_h)$ depends on the environment and formation history as well (Gao & White 2007, Zentner et al. 2014, Montero-Dorta et al. 2017, Mao et al. 2018). Following Zehavi et al. (2011) (and references therein, e.g. Kravtsov et al. (2004), Zheng et al. (2005) and Zheng et al. (2007)) the mean number of galaxies $\langle N(M_h) \rangle$ is here expressed as the sum of the contribution of central galaxies $\langle N_{\text{cen}}(M_h) \rangle$ and the contribution of satellite galaxies $\langle N_{\text{sat}}(M_h) \rangle$. The former is parametrised by a step function with a soft cut-off profile to account for the scatter $\sigma_{\log(M_h)}$ in the luminosity-halo mass (or stellar mass-halo mass) relation:

$$\langle N_{\text{cen}}(M_h) \rangle = \alpha_{\text{cen}} \times \frac{1}{2} \left[1 + \text{erfc} \left(\frac{\log_{10}(M_h) - \log_{10}(M_{\text{th}})}{\sigma_{\log(M_h)}} \right) \right]. \quad (2.18)$$

Usually, M_{th} is defined as the mass scale at which 50% of the halos host a central galaxy, i.e. $\langle N_{\text{cen}}(M_{\text{th}}) \rangle = 0.5$. However, in this work the galaxy samples in stellar mass and luminosity are defined within bins (rather than by lower thresholds), so that not necessarily every halo hosts a central galaxy. Thus, an additional parameter α_{cen} is introduced to allow for halos without a central galaxy (Clampitt et al. 2017). This changes the meaning of the parameter M_{th} in the sense that M_{th} is the minimum halo mass crucial to potentially host in every second case a first central galaxy, i.e. $\langle N_{\text{cen}}(M_{\text{th}}) \rangle = 0.5 \times \alpha_{\text{cen}}$. The distribution of satellites is expressed by a power-law at the high halo mass end with the normalisation M_1 and a slope parameter β ,

$$\langle N_{\text{sat}}(M_h) \rangle = \left(\frac{M_h}{M_1} \right)^\beta \times \frac{1}{2} \left[1 + \text{erfc} \left(\frac{\log_{10}(M_h) - \log_{10}(M_{\text{th}})}{\sigma_{\log(M_h)}} \right) \right]. \quad (2.19)$$

The complete HOD,

$$\begin{aligned} \langle N(M_h) \rangle &= \langle N_{\text{cen}}(M_h) \rangle + \langle N_{\text{sat}}(M_h) \rangle \\ &= \frac{1}{2} \left[1 + \text{erfc} \left(\frac{\log_{10}(M_h) - \log_{10}(M_{\text{th}})}{\sigma_{\log(M_h)}} \right) \right] \times \left[\alpha_{\text{cen}} + \left(\frac{M_h}{M_1} \right)^\beta \right], \end{aligned} \quad (2.20)$$

thus depends on a total of five model parameters $\sigma_{\log(M_h)}$, M_{th} , M_1 , β , and α_{cen} .

2.2.2 Derived parameters

From the HOD properties of the galaxy sample can be derived. The mean number density of galaxies for a given sample is determined by the completeness relation,

$$\bar{n}_g(z) = \int dM_h \langle N(M_h) \rangle n(M_h, z), \quad (2.21)$$

with the halo mass function $n(M_h, z)$ defined by Eq. (2.2). The fraction of central galaxies per sample is given by

$$f_{\text{cen}}(z) = \frac{\int dM_h \langle N_{\text{cen}}(M_h) \rangle n(M_h, z)}{\bar{n}_g(z)}, \quad (2.22)$$

and the mean halo mass per sample by

$$\langle M_h(z) \rangle = \int dM_h M_h n(M_h, z) \frac{\langle N(M_h) \rangle}{\bar{n}_g(z)}. \quad (2.23)$$

All three quantities depend on redshift. Usually, they are computed for the mean redshift \bar{z} of the considered galaxy sample, which is also done for the mean halo mass $\langle M_h^{\bar{z}} \rangle$ in this work. However, it is more accurate to calculate the number of galaxies in the sample by integrating over the redshift interval $[z_{\text{min}}; z_{\text{max}}]$ (Coupon et al. 2012):

$$N^V = \Omega \int_{z_{\text{min}}}^{z_{\text{max}}} f_K^2(\chi(z)) \frac{d\chi}{dz} \bar{n}_g(z) dz. \quad (2.24)$$

Here Ω is the survey area in units of square degrees, and $f_K^2(\chi(z)) d\chi/dz$ is the volume element. While Eq. (2.24) is strictly only valid for spectroscopic redshifts, in this work the assumption is made that the photometric redshift distribution can be approximated by a top-hat function for the interval $[z_{\text{min}}; z_{\text{max}}]$. The respective mean number density in the redshift bin is then given by:

$$\bar{n}_g^V = N^V \left[\Omega \int_{z_{\text{min}}}^{z_{\text{max}}} f_K^2(\chi(z)) \frac{d\chi}{dz} dz \right]^{-1}. \quad (2.25)$$

Finally, the fraction of centrals for the redshift bin can be written as

$$f_{\text{cen}}^V = \frac{\bar{n}_{g,\text{cen}}^V}{\bar{n}_g^V}. \quad (2.26)$$

2.2.3 Conversion to the ratio of stellar to halo mass

In Chapter 4 halo model fits are performed to the CFHTLenS G2L signal and the number of galaxies for seven galaxy samples binned by stellar mass M_* . The resulting best-fit parameter set for each stellar mass sample determines the HOD $\langle N(M_h) | M_* \rangle_i$, which is the mean number of galaxies of stellar mass M_* hosted by halos of mass M_h in the mass bin i .

For comparison to other works, $\langle N(M_h)|M_* \rangle_i$ can be converted to the ratio of stellar mass to halo mass $M_*(M_h)/M_h$ as follows (Patrick Simon, priv. comm.). The average ratio of stellar to halo mass in halos of mass M_h is given by $M_*/M_h \times \langle N(M_h)|M_* \rangle_i$. Given N_* number of mass bins, the total fraction of stellar to halo mass across the full stellar mass range is given by

$$\frac{M_*}{M_h}(M_h) = \sum_{i=1}^{N_*} \int \frac{dM_*}{M_h} M_* p_i(M_*) \langle N(M_h)|M_* \rangle_i, \quad (2.27)$$

where $p_i(M_*)$ is the probability to find a galaxies of stellar mass M_* inside the i th bin. To estimate the total stellar mass in galaxies it is common to simultaneously fit a model of the stellar mass function to the data (Leauthaud et al. 2011). However, in this work M_* is assumed to be on average provided by galaxies of the considered stellar mass samples. This is, however, only a lower limit on the total stellar mass inside a halo because of the incompleteness of the galaxy selection in the CFHTLenS.

A Monte-Carlo algorithm is used to estimate the statistical error on the result. For each bin i with halo model parameters \mathbf{p}_i , j random realisations are drawn from the model parameter probability distributions to give new model parameter sets $\{\mathbf{p}_i\}_j$. These correspond to a new set of HODs: $\{\langle N(M_h)|M_* \rangle_i\}_j$, which give j estimates of $M_*(M_h)/M_h$. The mean and variance of these estimates at a given M_h is adopted as the 1σ uncertainty of $M_*(M_h)/M_h$.

2.3 Halo model galaxy-matter power- and bispectrum

The strength of the halo model is that it can be used to derive analytical expressions for the n -point (cross-)correlation functions of galaxies and dark matter. For this work the two- and three-point galaxy-matter (cross-)correlation functions are relevant, to be precise their Fourier space analogues, the galaxy-matter power spectrum and bispectrum, as they are probed by galaxy-galaxy(-galaxy) lensing. Conversely, the galaxy-matter power spectrum and bispectrum can be transformed into halo model predictions for G2L and G3L as described in Chapter 3. In the following the definitions for the halo model galaxy-matter power and bispectrum are given. For a derivation of these equations see Rödiger (2009).

2.3.1 Halo model galaxy-dark matter power spectrum

The three-dimensional galaxy-dark matter power spectrum $P^{\delta g}(k)$ is defined according to Eq. (1.50) by the two-point correlator of the dark matter density contrast $\delta(\mathbf{k})$ and the galaxy density contrast $\delta_g(\mathbf{k})$:

$$\langle \delta(\mathbf{k}) \delta_g(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') P^{\delta g}(k). \quad (2.28)$$

The power spectrum consists of two terms, one accounting for contributions of correlations originating from the same halo, and the other for contributions originating from two different halos:

$$P^{\delta g}(k) = P_{1h}^{\delta g}(k) + P_{2h}^{\delta g}(k), \quad (2.29)$$

with

$$\begin{aligned}
P_{1h}^{\delta g}(k) &= \int dM_h n(M_h) \left(\frac{M_h}{\bar{\rho}} \right) \tilde{u}_{\text{dm}}(k, M_h) \left[\frac{\bar{N}_{\text{cen}} + \bar{N}_{\text{sat}} \tilde{u}_g(k, M_h)}{\bar{n}_g} \right], \\
P_{2h}^{\delta g}(k) &= \left[\int dM_h n(M_h) \left(\frac{M_h}{\bar{\rho}} \right) \tilde{u}_{\text{dm}}(k, M_h) b_1^h(M_h) \right] \\
&\times \left[\int dM_h n(M_h) \frac{[\bar{N}_{\text{cen}} + \bar{N}_{\text{sat}} \tilde{u}_g(k, M_h)]}{\bar{n}_g} b_1^h(M_h) \right] \times P_{\text{lin}}(k). \quad (2.30)
\end{aligned}$$

Here $n(M_h)$ is the halo mass function (Eq. 2.2), $\bar{\rho}$ the mean density of the universe (Eq. 1.22), \bar{n}_g the mean number density of galaxies (Eq. 2.21), and \bar{N}_{cen} and \bar{N}_{sat} the central and satellite parts of the HOD (Eqs. 2.18 and 2.19). Moreover, $\tilde{u}_g(k, M_h)$ is the normalised galaxy density profile, which is assumed to be identical with the one of dark matter $\tilde{u}_{\text{dm}}(k, M_h)$ (Eq. 2.12) in this work. Finally, the power spectrum depends on the dark matter halo bias $b_1^h(M_h)$ (Eq. 2.16) and on the linear dark matter power spectrum from perturbation theory $P_{\text{lin}}(k)$ (Eq. 1.54).

2.3.2 Halo model galaxy-dark matter bispectrum

The three-dimensional galaxy-dark matter bispectrum $B^{\text{gg}\delta}(\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_3)$ is defined according to Eq. (1.52) by the three-point correlator of the dark matter and galaxy density contrast:

$$\langle \delta_g(\mathbf{k}_1) \delta_g(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle = (2\pi)^3 \delta_{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B^{\text{gg}\delta}(\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_3). \quad (2.31)$$

Assuming a statistically homogeneous and isotropic Universe, the bispectrum can be parametrised by the length of the three \mathbf{k} -vectors that build a triangle in Fourier space. The bispectrum consists of three terms: the first one accounting for contributions of correlations originating from the same halo, the second one accounting for contributions originating from two different halos, and the third one accounting for contributions originating from three different halos:

$$B^{\text{gg}\delta}(k_1, k_2; k_3) = B_{1h}^{\text{gg}\delta}(k_1, k_2; k_3) + B_{2h}^{\text{gg}\delta}(k_1, k_2; k_3) + B_{3h}^{\text{gg}\delta}(k_1, k_2; k_3), \quad (2.32)$$

with

$$\begin{aligned}
B_{1h}^{\text{gg}\delta} &= \int dM_h n(M_h) \frac{1}{\bar{n}_g^2} \left[\bar{N}_{\text{sat}}^2 + \bar{N}_{\text{cen}} \bar{N}_{\text{sat}} \left(\frac{1}{\tilde{u}_g(k_1, M_h)} + \frac{1}{\tilde{u}_g(k_2, M_h)} \right) \right] \\
&\times \left(\frac{M_h}{\bar{\rho}} \right) \tilde{u}_g(k_1, M_h) \tilde{u}_g(k_2, M_h) \tilde{u}_{\text{dm}}(k_3, M_h), \\
B_{2h}^{\text{gg}\delta} &= G_{11}(k_1) \mathcal{G}_{12}^{\delta}(k_2; k_3) P_{\text{lin}}(k_1) + I_{11}(k_3) G_{12}(k_2, k_1) P_{\text{lin}}(k_3) \\
&+ G_{11}(k_2) \mathcal{G}_{12}^{\delta}(k_1; k_3) P_{\text{lin}}(k_2), \\
B_{3h}^{\text{gg}\delta} &= G_{11}(k_1) G_{11}(k_2) I_{11}(k_3) B_{\text{lin}}(k_1, k_2, k_3). \quad (2.33)
\end{aligned}$$

Here $B_{\text{lin}}(k_1, k_2, k_3)$ is the underlying dark matter bispectrum from perturbation theory (Eq. 1.55). The functions $G_{11}(k)$, $G_{12}(k_1, k_2)$, $\mathcal{G}_{12}^{\text{g}\delta}(k_1; k_2)$, and $I_{11}(k)$ are given by

$$\begin{aligned}
G_{11}(k) &= \int dM_{\text{h}} n(M_{\text{h}}) \left[\frac{\bar{N}_{\text{cen}} + \bar{N}_{\text{sat}} \tilde{u}_{\text{g}}(k, M_{\text{h}})}{\bar{n}_{\text{g}}} \right] b_1^{\text{h}}(M_{\text{h}}), \\
G_{12}(k_1, k_2) &= \frac{1}{\bar{n}_{\text{g}}^2} \int dM_{\text{h}} n(M_{\text{h}}) \tilde{u}_{\text{g}}(k_1, M_{\text{h}}) \tilde{u}_{\text{g}}(k_2, M_{\text{h}}) \\
&\quad \times \left[\bar{N}_{\text{sat}}^2 + \bar{N}_{\text{cen}} \bar{N}_{\text{sat}} \left(\frac{1}{\tilde{u}_{\text{g}}(k_1, M_{\text{h}})} + \frac{1}{\tilde{u}_{\text{g}}(k_2, M_{\text{h}})} \right) \right] b_1^{\text{h}}(M_{\text{h}}), \\
\mathcal{G}_{12}^{\text{g}\delta}(k_1; k_2) &= \frac{1}{\bar{n}_{\text{g}}} \int dM_{\text{h}} n(M_{\text{h}}) [\bar{N}_{\text{sat}}(M_{\text{h}}) \tilde{u}_{\text{g}}(k_1, M_{\text{h}}) + \bar{N}_{\text{cen}}(M_{\text{h}})] \\
&\quad \times \left(\frac{M_{\text{h}}}{\bar{\rho}} \right) \tilde{u}_{\text{dm}}(k_2, M_{\text{h}}) b_1^{\text{h}}(M_{\text{h}}), \\
I_{11}(k) &= \int dM_{\text{h}} n(M_{\text{h}}) \left(\frac{M_{\text{h}}}{\bar{\rho}} \right) \tilde{u}_{\text{dm}}(k, M_{\text{h}}) b_1^{\text{h}}(M_{\text{h}}). \tag{2.34}
\end{aligned}$$

2.3.3 Projecting the spectra from 3D to 2D - Limber's Equation

The power spectrum $P^{\delta\text{g}}(k)$ and the bispectrum $B^{\text{gg}\delta}(\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_3)$ are three-dimensional quantities. However, what is measured on the sky are estimates for the projected versions of these three-dimensional quantities. To be able to compare the halo model predictions for G2L and G3L to the measurements from CFHTLenS, the three-dimensional spectra are transformed into two-dimensional spectra using Limber's approximation (Limber 1953, Kaiser 1998).

The basic idea is that for every modulus ℓ of the two-dimensional Fourier vector ℓ or the vector itself, the integral over the comoving distance χ is performed by using the following weighting function:

$$G(\chi) = \frac{3}{2} \Omega_{\text{m}} \left(\frac{H_0}{c} \right)^2 a^{-1}(\chi) \int_{\chi}^{\chi_{\text{H}}} d\chi_{\text{s}} p_{\text{s}}(z) \frac{dz}{d\chi_{\text{s}}} \frac{f_K(\chi_{\text{s}} - \chi)}{f_K(\chi_{\text{s}})}. \tag{2.35}$$

Here Ω_{m} is the matter density parameter, H_0 the Hubble constant (Eq. 1.3), $a(\chi)$ the scale factor (Eq. 1.4), and $f_K(\chi)$ the comoving angular diameter distance (Eq. 1.9). To compute this function the redshift distribution of source galaxies $p_{\text{s}}(z)$ needs to be known. In this work it is assumed that all source galaxies have approximately the same redshift, which simplifies Eq. (2.35) as the integral does not need to be performed. For the cross-spectra considered in this work, the two-dimensional expressions are then given by:

$$\begin{aligned}
P^{\kappa\text{g}}(\ell) &= \int_0^{\chi_{\text{H}}} d\chi \frac{G(\chi) p_{\text{f}}(\chi)}{f_K(\chi)} P^{\delta\text{g}} \left(k = \frac{\ell}{f_K(\chi)}; \chi \right), \\
B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3) &= \int_0^{\chi_{\text{H}}} d\chi \frac{G(\chi) p_{\text{f}}^2(\chi)}{\chi^3} B^{\text{gg}\delta} \left(\mathbf{k}_1 = \frac{\ell_1}{\chi}, \mathbf{k}_2 = \frac{\ell_2}{\chi}; \mathbf{k}_3 = \frac{\ell_3}{\chi}; \chi \right), \tag{2.36}
\end{aligned}$$

where for the redshift distribution of lens galaxies $p_{\text{f}}(\chi)$ the distribution measured from CFHTLenS (normalised to unity) is used.

2.3.4 Reduced bispectrum

In the numerical implementation of the transformation of the galaxy-matter bispectrum to **G3L**, the $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ -function is called very often. To save computation time, the bispectrum is therefore numerically interpolated. However, $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ shows a cusp feature for squeezed triangles ($\ell_1 = \ell_2$, $\ell_3 = 0$ or $\varphi = \pi$) as also found by, e.g. [Jeong & Komatsu \(2009\)](#), which originates from the second-order coupling function $F_2(\mathbf{k}_1, \mathbf{k}_2)$ (Eq. 1.56) and the linear power spectrum from perturbation theory $P_{\text{lin}}(k)$ (Eq. 1.54) vanishing for $\mathbf{k}_3 = 0$. This feature is tricky to interpolate. The reduced bispectrum, defined as ([Cooray & Sheth 2002](#))

$$Q^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3) = \frac{B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)}{P^{\kappa\text{g}}(\ell_1)P^{\kappa\text{g}}(\ell_2) + P^{\text{gg}}(\ell_1)P^{\kappa\text{g}}(\ell_3) + P^{\text{gg}}(\ell_2)P^{\kappa\text{g}}(\ell_3)}, \quad (2.37)$$

contains the same information as $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$, but shows a less pronounced cusp feature. Therefore, numerically $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ is transformed to $Q^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$, for which the interpolation grid is constructed. When $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ is called in the code it is computed from the interpolation values of $Q^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$.

The computation of $Q^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ requires the two-dimensional galaxy power spectrum:

$$P^{\text{gg}}(\ell) = \int_0^{\chi_{\text{H}}} d\chi \frac{G(\chi)p_{\text{f}}(\chi)}{f_K(\chi)} P^{\text{gg}}\left(k = \frac{\ell}{f_K(\chi)}; \chi\right), \quad (2.38)$$

which is obtained by projecting the three-dimensional galaxy power spectrum given by

$$\begin{aligned} P^{\text{gg}}(k) &= P_{\text{1h}}^{\text{gg}}(k) + P_{\text{2h}}^{\text{gg}}(k), \\ P_{\text{1h}}^{\text{gg}}(k) &= \int dM_{\text{h}} n(M_{\text{h}}) \left[\frac{[\bar{N}_{\text{sat}} \tilde{u}_{\text{g}}(k, M_{\text{h}})]^2 + 2\bar{N}_{\text{cen}} \bar{N}_{\text{sat}} \tilde{u}_{\text{g}}(k, M_{\text{h}})}{\bar{n}_{\text{g}}^2} \right], \\ P_{\text{2h}}^{\text{gg}}(k) &= \left[\int dM_{\text{h}} n(M_{\text{h}}) \frac{[\bar{N}_{\text{cen}} + \bar{N}_{\text{sat}} \tilde{u}_{\text{g}}(k, M_{\text{h}})]}{\bar{n}_{\text{g}}} b_1^{\text{h}}(M_{\text{h}}) \right]^2 \times P_{\text{lin}}(k). \end{aligned} \quad (2.39)$$

Chapter 3

Galaxy-galaxy(-galaxy) lensing

Light rays emitted by background objects (‘sources’) are differentially deflected by the gravitational potential of the intervening matter distribution (‘lenses’) on their path to Earth. This so-called *gravitational lensing effect* distorts the intrinsic shapes of the source images, thereby encoding information about the matter environment of the lenses in the sheared source images. This work is concerned with subtle distortions attributed to the *weak lensing effect*, where a significant lensing signal can only be obtained by statistically superposing the signals of many source-lens pairs using correlation functions. For general reviews on weak gravitational lensing see [Bartelmann & Schneider \(2001\)](#) and [Schneider \(2006\)](#).

One application of weak gravitational lensing is *galaxy-galaxy lensing (G2L)*, which considers the cross-correlation between the position of a lens galaxy and the shear of a source galaxy, thereby probing the radial profile of the surface mass density around an average lens (e.g. [Mandelbaum et al. 2006a](#), [van Uitert et al. 2011](#), [Clampitt et al. 2017](#)). Thus, **G2L** provides not only a possibility to map dark matter, but also to infer the statistical properties of the galaxy-halo connection (e.g. [Zheng et al. 2007](#), [Zehavi et al. 2011](#), [Leauthaud et al. 2011](#), [Coupon et al. 2012](#), [Velandar et al. 2014](#)). The extension of the concept of **G2L** to third-order is *galaxy-galaxy-galaxy lensing (G3L)*, particularly the lens-lens-shear correlator \mathcal{G} which cross-correlates the positions of two lens galaxies with the shear of a source galaxy. This **G3L** correlator thus probes the dark matter environment about pairs of galaxies ([Schneider & Watts 2005](#)).

Two ways to express \mathcal{G} are *excess mass maps* and the *aperture statistics* $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$. Excess mass maps depict, for a fixed separation of lenses on the sky, the average distribution of surface mass around lens pairs in excess of the average surface mass around two individual lenses as measured with **G2L** ([Simon et al. 2008](#), [Simon et al. 2012](#), [Simon et al. 2013](#), [Simon et al. 2019](#)). Thus, excess mass maps provide an intuitive visualisation of the information contained in \mathcal{G} . An alternative way of representing \mathcal{G} is via the aperture statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$, which provides a localised measurement of the projected galaxy-galaxy-matter bispectrum ([Schneider et al. 1998](#), [Simon et al. 2013](#)). Aperture statistics are useful in weak lensing as they are comparatively easy to measure, while containing all the information in their related correlation function or spectrum ([Schneider & Watts 2005](#)).

Over the last decade successful **G3L** measurements have demonstrated a correlation between the properties of the lens galaxies and their matter environment (Simon et al. 2008, Simon et al. 2013, Simon et al. 2019). By confronting halo model predictions with these measurements, the ability of the halo model to consistently describe second- and third-order statistics, as probed by **G2L** and **G3L**, is tested in this work. On the other hand, the halo model is used to study dependencies of **G3L** on galaxy and model properties in order to build up a knowledge base on which to interpret present and future measurements of **G3L**.

To facilitate these two investigations in the following chapters, this chapter summarises the theory of **G2L** and **G3L**. First, in Sect. 3.1 the observables of weak lensing are introduced and the **G2L** and **G3L** correlation functions are defined, including estimators for the excess mass maps and the aperture statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$. Second, Sect. 3.2 outlines how **G2L** and **G3L**, again in terms of excess mass maps and aperture statistics, are predicted based on the halo model formalism introduced in Chapter 2. This chapter is based on Schneider & Watts (2005), Simon et al. (2013), Simon et al. (2019), notes by Patrick Simon (priv. comm.), and the doctoral thesis by Rödiger (2009).

3.1 Definitions

Galaxy-galaxy(-galaxy) lensing probes the distribution of matter around galaxies (pairs of galaxies) by considering the cross-correlation of the projected total matter distribution and the projected lens galaxy distribution.

The three-dimensional matter distribution in the Universe is described by the density contrast $\delta(\mathbf{x})$ (Eq. 1.32), which is projected along the line-of-sight by applying the lensing efficiency,

$$G(\chi) = \frac{3H_0^2 \Omega_m}{2c^2} a(\chi)^{-1} \int_{\chi}^{\chi_h} d\chi' p_s(\chi') \frac{f_K(\chi' - \chi)}{f_K(\chi')}, \quad (3.1)$$

which is the weight function introduced in Eq. (2.35). Furthermore, positions \mathbf{x} are projected into angular positions $\boldsymbol{\theta}$ by defining the transverse, comoving separation of a light ray from a reference line-of-sight by $f_K(\chi)\boldsymbol{\theta}$ with $\boldsymbol{\theta} = 0$ in the direction of the reference line-of-sight. The outcome is the two-dimensional analogue to $\delta(\mathbf{x})$, the lensing *convergence* (see e.g. Bartelmann & Schneider 2001)

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_h} d\chi f_K(\chi) G(\chi) \delta(f_K(\chi)\boldsymbol{\theta}, \chi). \quad (3.2)$$

The actual observable in weak lensing surveys is not $\kappa(\boldsymbol{\theta})$ itself, but the Cartesian shear $\gamma_c(\boldsymbol{\theta})$, for which a noisy estimator is given by the ellipticities $\epsilon(\boldsymbol{\theta})$ of the lensed source images:

$$\epsilon(\boldsymbol{\theta}) \approx \gamma_c(\boldsymbol{\theta}) + \epsilon_s. \quad (3.3)$$

Here ϵ_s accounts for the intrinsic, that means unlensed, unknown shapes of the source galaxies. Galaxies are assumed to have a random intrinsic orientation, so that the ensemble average over the intrinsic ellipticities vanishes and the ellipticity is an unbiased estimator of the shear, that means $\langle \epsilon(\boldsymbol{\theta}) \rangle = \langle \gamma_c(\boldsymbol{\theta}) \rangle$. As each measured elliptical source image is the

result of the numerous differential deflections the corresponding light bundle experiences by being exposed to various gravitational potentials along the path from its source to the observer, the resulting shear field $\gamma_c(\boldsymbol{\theta})$ contains the valuable information on the projected mass distribution. Using the [Kaiser & Squires relation \(1993\)](#) the shear $\gamma_c(\boldsymbol{\theta})$ can be converted into the lensing convergence $\kappa(\boldsymbol{\theta})$ up to a constant κ_0 :

$$\kappa(\boldsymbol{\theta}) - \kappa_0 = \frac{1}{\pi} \int d^2\theta' \mathcal{D}^*(\boldsymbol{\theta} - \boldsymbol{\theta}') \gamma_c(\boldsymbol{\theta}'), \quad (3.4)$$

with the kernel

$$\mathcal{D}^*(\boldsymbol{\theta}) = \frac{\theta_2^2 - \theta_1^2 + 2i\theta_1\theta_2}{|\boldsymbol{\theta}|^4}. \quad (3.5)$$

In this context θ_1 and θ_2 are the real and imaginary parts of the direction $\boldsymbol{\theta}$, which are expressed as complex number in a Cartesian frame. In Fourier space this convolution simplifies to

$$\tilde{\gamma}_c(\boldsymbol{\ell}) = e^{2i\beta} \tilde{\kappa}(\boldsymbol{\ell}), \quad (3.6)$$

where $e^{2i\beta} = (\ell_1 + i\ell_2)^2/\ell^2$. Here the representation of the two-dimensional vector $\boldsymbol{\ell}$ as complex number is used, with the corresponding polar angle β .

The shear at position $\boldsymbol{\theta}$ has two independent components, which can be expressed as tangential shear γ_t and cross shear γ_\times relative to an orientation φ . These components are defined as the real and the imaginary part of the cartesian shear $\gamma_c(\boldsymbol{\theta})$ rotated by the polar angle φ :

$$\gamma(\boldsymbol{\theta}; \varphi) = -e^{-2i\varphi} \gamma_c(\boldsymbol{\theta}) = \gamma_t(\boldsymbol{\theta}; \varphi) + i\gamma_\times(\boldsymbol{\theta}; \varphi). \quad (3.7)$$

Similar to the lensing convergence, the number density contrast of lenses on the sky $\kappa_g(\boldsymbol{\theta})$ is obtained by projecting the three-dimensional number density contrast of galaxies $\delta_g(\boldsymbol{x})$ along the line-of-sight (e.g. [Hoekstra et al. 2002](#)),

$$\kappa_g(\boldsymbol{\theta}) = \int_0^{\chi_h} d\chi p_d(\chi) \delta_g(f_K(\chi)\boldsymbol{\theta}, \chi), \quad (3.8)$$

where the probability density $d\chi p_d(\chi)$ of lens distances is estimated from the probability density function $dz p_z(z)$ of lens redshifts. With the number density of lens galaxies on the sky denoted by $N_g(\boldsymbol{\theta})$, and the mean number density of lens galaxies by \overline{N}_g , $\kappa_g(\boldsymbol{\theta})$ is alternatively given by

$$\kappa_g(\boldsymbol{\theta}) = \frac{N_g(\boldsymbol{\theta}) - \overline{N}_g}{\overline{N}_g}. \quad (3.9)$$

3.1.1 Galaxy-galaxy lensing

[G2L](#) probes the the average (surface-)matter density around galaxies by considering the second-order cross-correlation between the projected number density contrast of lens galaxies $\kappa_g(\boldsymbol{\theta})$ and the shear $\gamma(\boldsymbol{\theta} + \boldsymbol{\vartheta}; \varphi)$. Here the observed position of the source relative to the lens at $\boldsymbol{\theta}$ is denoted by $\boldsymbol{\vartheta} = \vartheta e^{i\varphi}$. The geometry of [G2L](#) is shown in [Fig. 3.1](#). The [G2L](#)

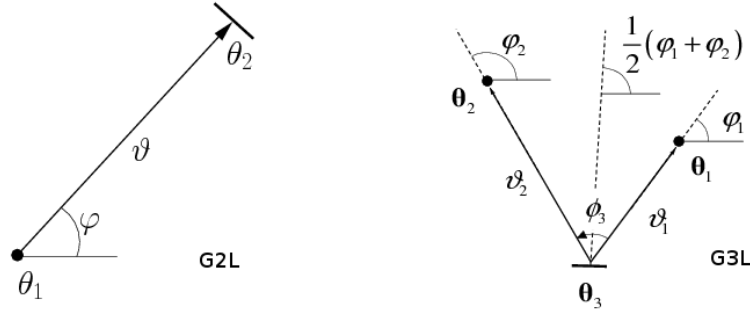


Figure 3.1: *Left:* Illustration of the parametrization of the **G2L** correlator $\langle \gamma_t \rangle(\vartheta)$. *Right:* Illustration of the parametrization of the **G3L** galaxy-galaxy-shear correlator $\mathcal{G}(\vartheta_1, \vartheta_2, \phi_3)$ with one lens at θ_1 , the other lens at θ_2 , and the source at θ_3 . [Source: [Schneider & Watts \(2005\)](#)]

correlator is defined by the average of the tangential shear over many lens-source pairs, formally the ensemble average

$$\langle \gamma_t \rangle(\vartheta) = \langle \kappa_g(\theta) \gamma(\theta + \vartheta; \varphi) \rangle, \quad (3.10)$$

(e.g. [Hoekstra et al. 2002](#), [Schneider & Watts 2005](#)). Note that γ_\times vanishes in the statistical average due to parity invariance ([Schneider 2003](#)). Due to statistical homogeneity and isotropy the correlator depends only on the separation ϑ . The physical interpretation of **G2L** is that of stacking the convergences around individual lens galaxies, thereby probing the average convergence profile $\bar{\kappa}(\vartheta) = \langle \kappa_g(\theta) \kappa(\theta + \vartheta) \rangle$ around lens galaxies ([Kaiser 1995](#)):

$$\langle \gamma_t \rangle(\vartheta) = \left(\frac{2}{\vartheta^2} \int_0^\vartheta d\vartheta' \vartheta' \bar{\kappa}(\vartheta') \right) - \bar{\kappa}(\vartheta). \quad (3.11)$$

In practice $\langle \gamma_t \rangle(\vartheta)$ is estimated by averaging the tangential shear, for a selected population of lenses, over many lens-source pairs in bins of ϑ .

3.1.2 Galaxy-galaxy-galaxy lensing

[Schneider & Watts \(2005\)](#) extended the concept of **G2L** to third-order and introduced two new classes of galaxy-shear correlation functions. Whereas the first class considers the cross-correlation of the shear of two background galaxies and the position of one foreground galaxy, the second class considers the cross-correlation of the positions of two lenses at θ_1 and θ_2 , and the shear at position θ_3 . This work is concerned with the latter case, for which the lens-lens-shear correlator \mathcal{G} is defined as:

$$\begin{aligned} \mathcal{G}(\vartheta_1, \vartheta_2) = \mathcal{G}(\vartheta_1, \vartheta_2, \phi_3) &= \left\langle \kappa_g(\theta_1) \kappa_g(\theta_2) \gamma \left(\theta_3; \frac{\varphi_1 + \varphi_2}{2} \right) \right\rangle \\ &= -e^{-i(\varphi_1 + \varphi_2)} \langle \kappa_g(\theta_1) \kappa_g(\theta_2) \gamma_c(\theta_3) \rangle. \end{aligned} \quad (3.12)$$

In this configuration the rotated shear is defined with respect to the line that bisects the angle $\phi_3 = \varphi_2 - \varphi_1$ spanned by the two separation vectors $\vartheta_1 = \theta_1 - \theta_3$ and $\vartheta_2 = \theta_2 - \theta_3$

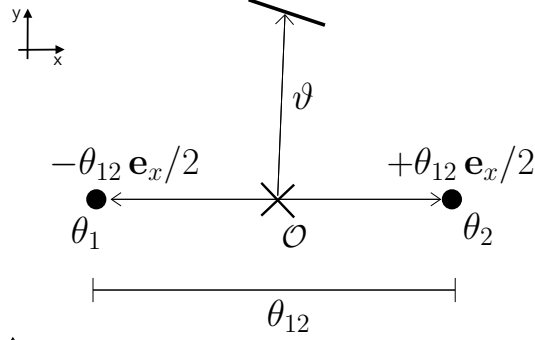


Figure 3.2: Illustration of the coordinate frame for the excess mass maps. [Source: Simon et al. (2019)]

describing the lens positions relative to the source position (see Fig. 3.1). The correlator \mathcal{G} contains no contributions from correlators smaller than third-order, depends only on the galaxy-matter bispectrum and thus vanishes for Gaussian random fields (e.g. Mo et al. 2010).

Excess mass maps

As previously mentioned, the G2L correlator $\langle \gamma_t \rangle(\vartheta)$ stacks the convergence measured around individual lenses and thereby relates to the average convergence profile about an average lens (Eq. 3.11). Similarly, the G3L correlator \mathcal{G} stacks the convergence measured around individual lens pairs, providing a measure of the average convergence profile about lens pairs, which can be expressed by the excess mass map (Saghiha 2017, Simon et al. 2019):

$$\begin{aligned} \overline{\Delta \kappa_{\text{emm}}}(\vartheta; \theta_{12}) &= \langle \kappa_{\text{g}}(\boldsymbol{\theta}_1) \kappa_{\text{g}}(\boldsymbol{\theta}_2) \kappa(\boldsymbol{\vartheta}) \rangle \\ &= \left(1 + w_{\text{gg}}(\theta_{12})\right) \bar{\kappa}_{\text{pair}}(\vartheta; \theta_{12}) - \bar{\kappa}_{\text{indiv}}(|\boldsymbol{\vartheta} - \boldsymbol{\theta}_1|) - \bar{\kappa}_{\text{indiv}}(|\boldsymbol{\vartheta} - \boldsymbol{\theta}_2|) + \kappa_0. \end{aligned} \quad (3.13)$$

A new coordinate system has been introduced for the excess mass maps, where the lenses are located at $\boldsymbol{\theta}_1 = -\theta_{12} \mathbf{e}_x / 2$ and at $\boldsymbol{\theta}_2 = +\theta_{12} \mathbf{e}_x / 2$ (see Fig. 3.2). The first term in Eq. (3.13) originates from the convergence stacking,

$$\bar{\kappa}_{\text{pair}}(\vartheta; \theta_{12}) = \frac{\langle N_{\text{g}}(\boldsymbol{\theta}_1) N_{\text{g}}(\boldsymbol{\theta}_2) \kappa(\boldsymbol{\vartheta}) \rangle}{\langle N_{\text{g}}(\boldsymbol{\theta}_1) N_{\text{g}}(\boldsymbol{\theta}_2) \rangle}, \quad (3.14)$$

which is weighted by the angular clustering correlation function of galaxies,

$$1 + w_{\text{gg}}(\theta_{12}) = \frac{\langle N_{\text{g}}(\boldsymbol{\theta}_1) N_{\text{g}}(\boldsymbol{\theta}_2) \rangle}{\bar{N}_{\text{g}}^2}. \quad (3.15)$$

The average convergence around individual lenses as measured by G2L is captured in the terms $\bar{\kappa}_{\text{indiv}}(|\boldsymbol{\vartheta} - \boldsymbol{\theta}_i|)$ with $i = 1, 2$:

$$\bar{\kappa}_{\text{indiv}}(|\boldsymbol{\vartheta} - \boldsymbol{\theta}_i|) = \frac{\langle N_{\text{g}}(\boldsymbol{\theta}_i) \kappa(\boldsymbol{\vartheta} - \boldsymbol{\theta}_i) \rangle}{\bar{N}_{\text{g}}}. \quad (3.16)$$

The relation between the excess mass maps $\overline{\Delta\kappa_{\text{emmm}}}(\boldsymbol{\vartheta}; \theta_{12})$ and the correlator \mathcal{G} is only defined up to the constant κ_0 (see Eq. 3.4). Therefore, all maps in this work are offset such that the average excess-convergence over the map vanishes.

Aperture statistics

Another way of expressing the lens-lens-shear correlations is by the third-order moment of the aperture statistics (Schneider & Watts 2005). Aperture statistics quantify the moments of fluctuations in the projected matter density field $\kappa(\boldsymbol{\theta})$ and in the projected galaxy number density field $\kappa_{\text{g}}(\boldsymbol{\theta})$ within apertures of angular scale θ and for a specific radial filter function. Here the filter function from Crittenden et al. (2002),

$$U_{\theta}(|\boldsymbol{\vartheta}|) = \frac{1}{\theta^2} u\left(\frac{|\boldsymbol{\vartheta}|}{\theta}\right), \quad (3.17)$$

with

$$u(x) = \frac{1}{2\pi} (1 - x^2/2) e^{-x^2/2}, \quad (3.18)$$

is used, which peaks in Fourier space at the angular wave number $\ell = \sqrt{2}/\theta$. The filter is compensated such that

$$\int_0^{\infty} d\vartheta \vartheta U_{\theta}(\vartheta) = 0. \quad (3.19)$$

The aperture mass is given by

$$M_{\text{ap}}(\theta) = \int_0^{\infty} d^2\vartheta U_{\theta}(|\boldsymbol{\vartheta}|) \kappa(\boldsymbol{\vartheta}), \quad (3.20)$$

and the aperture number count by

$$\mathcal{N}(\theta) = \int_0^{\infty} d^2\vartheta U_{\theta}(|\boldsymbol{\vartheta}|) \kappa_{\text{g}}(\boldsymbol{\vartheta}). \quad (3.21)$$

Assuming that the apertures are centred on $\boldsymbol{\theta}_c = 0$, the aperture statistics is defined as:

$$\begin{aligned} \langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta_1; \theta_2; \theta_3) &= \langle \mathcal{N}(\theta_1) \mathcal{N}(\theta_2) M_{\text{ap}}(\theta_3) \rangle \\ &= \int d^2\vartheta_1 U_{\theta_1}(\vartheta_1) \int d^2\vartheta_2 U_{\theta_2}(\vartheta_2) \int d^2\vartheta_3 U_{\theta_3}(\vartheta_3) \\ &\quad \times \langle \kappa_{\text{g}}(\boldsymbol{\vartheta}_1) \kappa_{\text{g}}(\boldsymbol{\vartheta}_2) \kappa(\boldsymbol{\vartheta}_3) \rangle. \end{aligned} \quad (3.22)$$

In this work, only case with $\theta_1 = \theta_2 = \theta_3$ are considered and for convenience the shorthand notation $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta) = \langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta; \theta; \theta)$ is used.

3.2 Halo model interpretation

To predict G2L and G3L using the halo model formalism introduced in Chapter 2, the relation between the observables $\langle \gamma_{\text{t}} \rangle(\vartheta)$, $\overline{\Delta\kappa_{\text{emmm}}}(\boldsymbol{\vartheta}; \theta_{12})$, and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta_1; \theta_2; \theta_3)$ and the galaxy-matter cross-spectra of second- and third-order is explored in the following two subsections.

3.2.1 Modelling galaxy-galaxy lensing

The real-space observable $\langle \gamma_t \rangle(\vartheta)$ can be related to the statistics of the underlying galaxy and dark matter density fields, in particular to the projected galaxy-matter power spectrum $P^{\kappa g}(\ell)$ (Eq. 2.36) derived in the halo model framework in Chapter 2. Since the galaxy distribution is a homogeneous random field, Eq. (3.10) can be rewritten as

$$\langle \gamma_t \rangle(\vartheta) = \langle \kappa_g(0) \gamma(\boldsymbol{\vartheta}; \varphi) \rangle. \quad (3.23)$$

Applying Eq. (3.7) allows one to separate the φ -dependence, and to transform $\kappa_g(0)$ and $\gamma_c(\boldsymbol{\vartheta})$ into Fourier space:

$$\begin{aligned} \langle \gamma_t \rangle(\vartheta) &= -e^{-2i\varphi} \langle \kappa_g(0) \gamma_c(\boldsymbol{\vartheta}) \rangle \\ &= -e^{-2i\varphi} \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} \int \frac{d^2\boldsymbol{\ell}'}{(2\pi)^2} e^{-i\boldsymbol{\vartheta} \cdot \boldsymbol{\ell}'} \langle \tilde{\kappa}_g(\boldsymbol{\ell}) \tilde{\gamma}_c(\boldsymbol{\ell}') \rangle \\ &= - \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} \int \frac{d^2\boldsymbol{\ell}'}{(2\pi)^2} e^{-i\boldsymbol{\vartheta} \cdot \boldsymbol{\ell}'} e^{2i(\beta-\varphi)} \langle \tilde{\kappa}_g(\boldsymbol{\ell}) \tilde{\kappa}(\boldsymbol{\ell}') \rangle. \end{aligned} \quad (3.24)$$

In the last step, the shear was transformed into the convergence applying the Kaiser-Squires relation $\tilde{\gamma}_c(\boldsymbol{\ell}') = e^{2i\beta} \tilde{\kappa}(\boldsymbol{\ell}')$. Identifying the ensemble average $\langle \tilde{\kappa}_g(\boldsymbol{\ell}) \tilde{\kappa}(\boldsymbol{\ell}') \rangle$ with the angular cross-power spectrum $P^{\kappa g}(\ell)$,

$$\langle \tilde{\kappa}_g(\boldsymbol{\ell}) \tilde{\kappa}(\boldsymbol{\ell}') \rangle = (2\pi)^2 \delta_D(\boldsymbol{\ell} + \boldsymbol{\ell}') P^{\kappa g}(\ell), \quad (3.25)$$

Eq. (3.24) can be written as

$$\begin{aligned} \langle \gamma_t \rangle(\vartheta) &= - \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} e^{i\boldsymbol{\vartheta} \cdot \boldsymbol{\ell}} e^{2i(\beta-\varphi)} P^{\kappa g}(\ell) \\ &= - \frac{1}{(2\pi)^2} \int_0^\infty d\ell \ell \int_0^{2\pi} d\varphi' e^{i\vartheta \ell \cos\varphi'} e^{2i\varphi'} P^{\kappa g}(\ell) \\ &= \frac{1}{(2\pi)} \int_0^\infty d\ell \ell P^{\kappa g}(\ell) J_2(\ell\vartheta). \end{aligned} \quad (3.26)$$

Here $J_2(\ell\vartheta)$ is the second-order Bessel function of first kind,

$$J_2(z) = -\frac{1}{2\pi} \int_0^{2\pi} d\phi e^{iz\cos(\phi)} e^{2i\phi}. \quad (3.27)$$

By applying Eq. (3.26), the halo model predictions for $P^{\kappa g}(\ell)$ can be transformed into the mean tangential shear around lenses. For this work, this step is realised using a numerical code written by Patrick Simon (priv. comm.).

3.2.2 Modelling galaxy-galaxy-galaxy lensing

Similarly, the observables of G3L, the excess mass maps $\overline{\Delta\kappa_{\text{emmm}}}(\boldsymbol{\vartheta}; \theta_{12})$ and the apertures statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta_1; \theta_2; \theta_3)$, can be related to the angular galaxy-matter bispectrum $B^{\text{gg}\kappa}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2; \boldsymbol{\ell}_3)$ defined in Eq. (2.36).

Excess mass maps

The excess mass defined in Eq. (3.13) is related to the bispectrum $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ through

$$\begin{aligned}
\overline{\Delta\kappa_{\text{emmm}}(\boldsymbol{\vartheta}; \theta_{12})} &= \langle \kappa_{\text{g}}(\boldsymbol{\theta}_1) \kappa_{\text{g}}(\boldsymbol{\theta}_2) \kappa(\boldsymbol{\vartheta}) \rangle \\
&= \int \frac{d^2\ell_1}{(2\pi)^2} \int \frac{d^2\ell_2}{(2\pi)^2} \int \frac{d^2\ell_3}{(2\pi)^2} e^{-i\boldsymbol{\theta}_1 \cdot \ell_1} e^{-i\boldsymbol{\theta}_2 \cdot \ell_2} e^{-i\boldsymbol{\vartheta} \cdot \ell_3} \langle \tilde{\kappa}_{\text{g}}(\ell_1) \tilde{\kappa}_{\text{g}}(\ell_2) \tilde{\kappa}(\ell_3) \rangle \\
&= \int \frac{d^2\ell_1 d^2\ell_2 d^2\ell_3}{(2\pi)^4} e^{-i\boldsymbol{\theta}_1 \cdot \ell_1 - i\boldsymbol{\theta}_2 \cdot \ell_2 - i\boldsymbol{\vartheta} \cdot \ell_3} \delta_{\text{D}}(\ell_1 + \ell_2 + \ell_3) B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3) \\
&= \int \frac{d^2\ell_1 d^2\ell_2}{(2\pi)^4} e^{-i\ell_1 \cdot (\boldsymbol{\theta}_1 - \boldsymbol{\vartheta}) - i\ell_2 \cdot (\boldsymbol{\theta}_2 - \boldsymbol{\vartheta})} B^{\text{gg}\kappa}(\ell_1, \ell_2; -\ell_1 - \ell_2), \quad (3.28)
\end{aligned}$$

with

$$\langle \tilde{\kappa}_{\text{g}}(\ell_1) \tilde{\kappa}_{\text{g}}(\ell_2) \tilde{\kappa}(\ell_3) \rangle = (2\pi)^2 \delta_{\text{D}}(\ell_1 + \ell_2 + \ell_3) B^{\text{gg}\kappa}(\ell_1, \ell_2; -\ell_1 - \ell_2). \quad (3.29)$$

The integral transformation of $B^{\text{gg}\kappa}(\ell_1, \ell_2; -\ell_1 - \ell_2)$ to $\overline{\Delta\kappa_{\text{emmm}}(\boldsymbol{\vartheta}; \theta_{12})}$ is, for this work, done numerically by using an approximate multipole expansion of the bispectrum to save computation time. This approach is based on a numerical code and on notes by Patrick Simon (priv. comm.), which are reproduced with his permission in Appendix E.

Aperture statistics

The aperture statistics defined in Eq. (3.22) is related to $B^{\text{gg}\kappa}(\ell_1, \ell_2; -\ell_1 - \ell_2)$ through

$$\begin{aligned}
\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta_1; \theta_2; \theta_3) &= \langle \mathcal{N}(\theta_1) \mathcal{N}(\theta_2) M_{\text{ap}}(\theta_3) \rangle \quad (3.30) \\
&= \int d^2\vartheta_1 U_{\theta_1}(\vartheta_1) \int d^2\vartheta_2 U_{\theta_2}(\vartheta_2) \int d^2\vartheta_3 U_{\theta_3}(\vartheta_3) \\
&\quad \times \langle \kappa_{\text{g}}(\boldsymbol{\vartheta}_1) \kappa_{\text{g}}(\boldsymbol{\vartheta}_2) \kappa(\boldsymbol{\vartheta}_3) \rangle \\
&= \left[\prod_{i=1}^3 \int d^2\vartheta_i U_{\theta_i}(\vartheta_i) e^{-i\ell_i \cdot \vartheta_i} \right] \int \frac{d^2\ell_1}{(2\pi)^2} \int \frac{d^2\ell_2}{(2\pi)^2} \int \frac{d^2\ell_3}{(2\pi)^2} \\
&\quad \times \langle \tilde{\kappa}_{\text{g}}(\ell_1) \tilde{\kappa}_{\text{g}}(\ell_2) \tilde{\kappa}(\ell_3) \rangle \\
&= \int \frac{d^2\ell_1 d^2\ell_2}{(2\pi)^4} \tilde{u}(\ell_1 \theta_1) \tilde{u}(\ell_2 \theta_2) \tilde{u}(|\ell_1 + \ell_2| \theta_3) B^{\text{gg}\kappa}(\ell_1, \ell_2; -\ell_1 - \ell_2),
\end{aligned}$$

where $|\ell_1 + \ell_2| = \sqrt{\ell_1^2 + \ell_2^2 + 2\ell_1\ell_2 \cos \phi}$ is the modulus of ℓ_3 . In this derivation the following relation has been used:

$$\int d^2\vartheta U_{\theta}(\vartheta) e^{-i\ell \cdot \vartheta} = \frac{1}{\theta^2} \int d^2\vartheta u(\vartheta/\theta) e^{-i\ell \cdot \vartheta} = \int d^2x u(x) e^{-i\theta\ell \cdot x} = \tilde{u}(\ell\theta). \quad (3.31)$$

Similar to the numerical transformation of $B^{\text{gg}\kappa}(\ell_1, \ell_2; -\ell_1 - \ell_2)$ to the excess mass maps, the transformation to the aperture statistics (Eq. 3.30) is performed using a multipole expansion of the bispectrum (Appendix E). The transformation is realised using the mentioned numerical code by Patrick Simon (priv. comm.).

3.2.3 Galaxy-galaxy-galaxy lensing for a Gaussian bispectrum

To test the numerical implementation of the integral transformations (Eqs. 3.28 and 3.30), a multivariate Gaussian bispectrum of the form

$$B(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2; \boldsymbol{\ell}_3) = \frac{\eta^4}{2\pi} \left(e^{-\frac{\eta^2 \ell_1^2}{2} - \frac{\eta^2 \ell_2^2}{2}} + e^{-\frac{\eta^2 \ell_1^2}{2} - \frac{\eta^2 \ell_3^2}{2}} + e^{-\frac{\eta^2 \ell_2^2}{2} - \frac{\eta^2 \ell_3^2}{2}} \right), \quad (3.32)$$

is assumed for $B^{\text{gg}\kappa}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2; \boldsymbol{\ell}_3)$. The parameter η (in arcmin) defines the angular scale of the bispectrum. For this particular bispectrum Eqs. (3.28) and (3.30) can be solved analytically. A lengthy analytical derivation results in:

$$\overline{\Delta\kappa_{\text{emmm}}(\boldsymbol{\vartheta}; \theta_{12})} = \frac{1}{(2\pi)^3} \left[e^{-\frac{(\theta_1 - \boldsymbol{\vartheta})^2 - (\theta_2 - \boldsymbol{\vartheta})^2}{2\eta^2}} + e^{-\frac{(\theta_1 - \theta_2)^2 - (\theta_2 - \boldsymbol{\vartheta})^2}{2\eta^2}} + e^{-\frac{(\theta_1 - \theta_2)^2 - (\theta_1 - \boldsymbol{\vartheta})^2}{2\eta^2}} \right], \quad (3.33)$$

and

$$\begin{aligned} \langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta_1; \theta_2; \theta_3) = \mathcal{P} \times & \left[h_1(\theta_1, \theta_2, \theta_3) + h_1(\theta_2, \theta_1, \theta_3) + h_1(\theta_1, \theta_3, \theta_2) \right. \\ & + h_1(\theta_3, \theta_1, \theta_2) + h_1(\theta_2, \theta_3, \theta_1) + h_1(\theta_3, \theta_2, \theta_1) \\ & \left. - h_2(\theta_1, \theta_2, \theta_3) - h_2(\theta_1, \theta_3, \theta_2) - h_2(\theta_2, \theta_3, \theta_1) \right], \end{aligned} \quad (3.34)$$

with the prefactor $\mathcal{P} = \frac{\eta^4}{(2\pi)^4} \frac{\theta_1^2 \theta_2^2 \theta_3^2}{8}$. The two auxiliary functions h_1 and h_2 are defined as,

$$h_1(x, y, z) = \frac{128\pi \left(\frac{y^2 + \eta^2 + z^2}{2} \right) \left[2 \left(\frac{x^2 + \eta^2 + z^2}{2} \right) \left(\frac{y^2 + \eta^2 + z^2}{2} \right) + z^4 \right]}{\left[-4 \left(\frac{x^2 + \eta^2 + z^2}{2} \right) \left(\frac{y^2 + \eta^2 + z^2}{2} \right) + z^4 \right]^4}, \quad (3.35)$$

and

$$h_2(x, y, z) = \frac{64\pi z^2 \left[8 \left(\frac{x^2 + \eta^2 + z^2}{2} \right) \left(\frac{y^2 + \eta^2 + z^2}{2} \right) + z^4 \right]}{\left[-4 \left(\frac{x^2 + \eta^2 + z^2}{2} \right) \left(\frac{y^2 + \eta^2 + z^2}{2} \right) + z^4 \right]^4}. \quad (3.36)$$

The transformation code is tested by comparing the numerical results for $\overline{\Delta\kappa_{\text{emmm}}(\boldsymbol{\vartheta}; \theta_{12})}$ and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta_1; \theta_2; \theta_3)$ to the analytically computed results.

Chapter 4

Halo model fits to the Millennium Simulation & the CFHTLenS G2L

To test whether a ‘standard’ halo model, as introduced in Chapter 2, can describe the observed second- and third-order galaxy-galaxy lensing signal consistently, the model is in a first step fitted to the galaxy-galaxy lensing (G2L) signal measured with the Canada-France Hawaii Telescope Lensing Survey (CFHTLenS) published in Saghiha et al. (2017; hereafter Saghiha+17). In this chapter the best-fit results for 29 lens galaxy samples in stellar mass, luminosity, and galaxy-type, further differentiating between low- and high- z samples, are presented. The best-fit parameters for the different galaxy samples are then used in Chapter 6 to predict the corresponding galaxy-galaxy-galaxy lensing (G3L), which is in a final step compared to the CFHTLenS G3L measurements by Simon et al. (2013; hereafter Simon+13). Additionally, the best-fit parameters for the different galaxy samples are used to derive the mean parent halo mass $\langle M_h \rangle$, the fraction of central galaxies f_{cen} , the first moment of the halo occupation distribution (HOD) $\langle N(M_h) \rangle$, and the ratio of stellar to halo mass $M_*(M_h)/M_h$. From the results conclusions are drawn about the distribution of stellar mass among halos.

Before testing the capability of the halo model to describe the clustering of dark matter and galaxies as probed by the CFHTLenS G2L and G3L, it is necessary to determine limitations of the model and the fitting procedure independently of the CFHTLenS data. To this end, in the first part of this chapter the accuracy of the model predictions is explored using mock data. More specifically, the model is fitted to the G2L signal measured by Saghiha+17 from the Millennium Simulation (Springel et al. 2005) in which galaxies have been included using the semi-analytical model (SAM) from Henriques et al. (2015; hereafter Henriques+15). For the mock data the mean parent halo mass, the fraction of central galaxies as well as the HOD are known for each stellar mass sample. The accuracy of the model and the fitting routine can thus be quantified by comparing the model predictions for these quantities to their ‘true’ values.

The content of this chapter is being prepared for submission to *Astronomy & Astrophysics*.

Table 4.1: Selection criteria of the lens samples for the **G2L** and **G3L** analysis adopted in this work (based on Simon+13). The stellar mass bins (**smx**), luminosity bins (**Lx**) and galaxy-type bins (**ETG**: early-type galaxies; **LTG**: late-type galaxies) are further subdivided into ‘low- z ’ ($0.2 \leq z_{\text{ph}} < 0.44$) and ‘high- z ’ ($0.44 \leq z_{\text{ph}} < 0.6$) samples. For the stellar mass and the luminosity estimates $h = 0.73$ is assumed.

Sample	Selection criteria	Sample	Selection criteria	Sample	Selection criteria
sm1	$0.5 \leq M_*/10^{10}M_{\odot} < 1.0$	L1	$-18 \leq M_r < -17$	ETG	$0 \leq \text{T.B} < 2 \mid -23 \leq M_r < -21$
sm2	$1.0 \leq M_*/10^{10}M_{\odot} < 2.0$	L2	$-19 \leq M_r < -18$	LTG	$2 \leq \text{T.B} < 6 \mid -23 \leq M_r < -21$
sm3	$2.0 \leq M_*/10^{10}M_{\odot} < 4.0$	L3	$-20 \leq M_r < -19$		
sm4	$4.0 \leq M_*/10^{10}M_{\odot} < 8.0$	L4	$-21 \leq M_r < -20$		
sm5	$8.0 \leq M_*/10^{10}M_{\odot} < 16.0$	L5	$-22 \leq M_r < -21$		
sm6	$16.0 \leq M_*/10^{10}M_{\odot} < 32.0$	L6	$-23 \leq M_r < -22$		
sm7	$32.0 \leq M_*/10^{10}M_{\odot} < 64.0$				

4.1 Data

4.1.1 CFHTLenS

The **CFHTLenS** is a 154 deg^2 multi-colour ($u^*g'r'i'z'$) optical survey performed with the MegaPrime/MegaCam of the Canada-France Hawaii Telescope at Mauna Kea, optimised for weak lensing analyses (Heymans et al. 2012, Erben et al. 2013). The **CFHTLenS** data comprises publicly available shear and photometric redshift catalogues, allowing the computation of the **G2L** and the **G3L** signal for different lens galaxy samples (Simon+13, Velander et al. 2014, Fu et al. 2014, Hudson et al. 2015, Coupon et al. 2015).

The **G2L** signal of the **CFHTLenS** data used in this work was measured and provided by Patrick Simon (priv. comm.) and is identical with the one published in Saghiha+17. Following Simon+13, 129 of 171 pointings appropriate for a weak lensing analysis are used (Heymans et al. 2012), where the survey area is further reduced by masking, leading to an effective area of 94.5 deg^2 . Source and lens galaxies are selected by their brightness, $i'_{\text{source}} \leq 24.7$ and $i'_{\text{lens}} \leq 22.5$, and by their photometric redshift z_{ph} . Source galaxies cover the redshift interval $0.65 \leq z_{\text{ph}} < 1.2$. Lens galaxies are subdivided into a low (‘low- z ’, $0.2 \leq z_{\text{ph}} < 0.44$) and a high photometric redshift sample (‘high- z ’, $0.44 \leq z_{\text{ph}} < 0.6$). With respect to their physical properties, galaxies are subdivided into stellar mass (**smx**), luminosity (**Lx**) as well as an early-type (**ETG**) and a late-type galaxy (**LTG**) sample as given in Table 4.1. Estimates for the stellar masses are obtained from fitting a galaxy spectral energy distribution (SED) model to the galaxy photometry as described in Velander et al. (2014). The same is done for the classification of galaxy-types, for which an additional cut in luminosity ($-23 \leq M_r < -21$) is made to obtain a volume-limited sample. The classification of galaxies as early- or late-type galaxies is based on the value of the T.B parameter in Bayesian photometric redshift (BPZ) estimation, which provides the most likely galaxy SED (for details see Benítez 2000). The galaxy luminosities are estimated from the absolute r' -band magnitudes. Due to very faint limits the L1 high- z sample contains too few lenses for **G2L** and **G3L** measurements. The L2 high- z sample is affected by incompleteness. The **G2L** signal is measured for $b = 13$ logarithmic bins of the lens-source separation θ , ranging from 0.5 arcmin to 20 arcmin. To fit the **CFHTLenS** data, the WMAP9 cosmology (Hinshaw et al. 2013) is adopted (Table A.1).

4.1.2 Mock data – Millennium Simulation & Semi-Analytical Model

The Millennium Simulation is a cosmological N -body simulation of the dark matter density field (Springel et al. 2005). Assuming a Λ CDM cosmology based on the first-year WMAP data ($\Omega_\Lambda = 0.75$, $\Omega_b = 0.045$, $\Omega_m = 0.25$, $\sigma_8 = 0.9$, $n_s = 1$ and $H_0 = 73 \text{ km s}^{-1} \text{ Mpc}^{-1}$; Spergel et al. 2003) the simulation traces the evolution of 2160^3 dark matter particles from redshift $z = 127$ to today in a comoving volume of $500 h^{-1} \text{ Mpc}$ side length. The spatial comoving resolution of the simulation is $5 h^{-1} \text{ kpc}$, allowing the resolution of individual dark matter halos. To implement galaxies into the dark matter simulation, a SAM is used (e.g. Bower et al. 2006, Guo et al. 2011, Henriques+15). For the resulting simulated realisation of the Universe, the lensing distortions are obtained by tracing light-rays from source galaxies on their path through the large-scale dark matter distribution associated with lens galaxies using a multiple-lens-plane approach (Hilbert et al. 2009). The output are convergence and shear maps for 64 fields of view of $4 \times 4 \text{ deg}^2$ each, along with galaxy redshifts, positions and additional properties, such as parent halo mass, stellar mass, and specifications on whether a galaxy is a central or a satellite galaxy.

The mock G2L data, used for the analysis in this work, is based on the Millennium Simulation employing the SAM by Henriques+15 and is published in Saghiiha+17 for the stellar mass samples sm1 to sm6 defined in Table 4.1. To ensure that the galaxy samples have the same lensing efficiency as those of CFHTLenS, Saghiiha+17 apply a rejection method to reproduce the redshift distributions of galaxies in CFHTLenS. Specifically, mock galaxies are randomly discarded to match the number density distribution of galaxies dn_{CFHTLenS}/dz . Contrary to CFHTLenS, the mock G2L signal does not contain shape noise, so that the errors result solely from the cosmic variance of the 64 fields.

For the accuracy test of the halo model using mock data, the total number of galaxies N_{gal} , and the fraction of centrals f_{cen} are computed for each stellar mass sample defined in Table 4.1. Errors on N_{gal} and f_{cen} are estimated from the cosmic variance of the 64 fields. The estimate for the mean halo mass per sample $\langle M_h \rangle$ is taken from Saghiiha+17. The reference mock HODs for the different samples $\langle N(M_h) \rangle$ were kindly computed by Stefan Hilbert (2016; priv. comm.). The HODs are for two mean redshifts of $\bar{z}_{\text{ph}} = 0.3, 0.5$ by counting the number of galaxies per parent halo mass M_h in bins of 0.1 dex in $\log_{10}(M_{200})$ for the stellar mass samples given in Table 4.1. A statistical error of 0.3 dex in stellar mass is assumed to be consistent with the selection in CFHTLenS (Velander et al. 2014).

4.2 Method

Halo model fits are performed to each of the mock and CFHTLenS G2L signals $\gamma_t(\theta)$ measured for the 29 galaxy samples. In addition to the G2L signal, the total number of galaxies per sample N_{gal} is used for the fitting. The data vector is therefore given by $\mathbf{d} = (\gamma_{t,1}^{\text{obs}}, \dots, \gamma_{t,b}^{\text{obs}}, N_{\text{gal}}^{\text{obs}})^T$ with $b = 13$ being the number of bins in angular separation θ . Following the same notation, the model vector reads $\mathbf{m}(\mathbf{p}) = (\gamma_{t,1}^{\text{model}}(\mathbf{p}), \dots, \gamma_{t,b}^{\text{model}}(\mathbf{p}), N_{\text{gal}}^{\text{model}}(\mathbf{p}))^T$ with the vector of model parameters $\mathbf{p} = (\sigma_{\log(M_h)}, \beta, \log_{10}M_1, \log_{10}M_{\text{th}}, \alpha_{\text{cen}})^T$. The best-fit model parameters and their uncertainties are obtained using the maximum-likelihood method (e.g. Barlow 1989).

4.2.1 Fitting procedure

The model fits to the **CFHTLenS** data are performed using the MINUIT package¹, an analysis tool for function minimisation developed at the Conseil Européen pour la Recherche Nucléaire (James & Roos 1975). Given the data vector \mathbf{d} , the fitting routine aims for finding the maximum likelihood (ML) point by minimising the negative of the log-likelihood function $\mathcal{L}(\mathbf{d}|\mathbf{p}) := -2\ln L(\mathbf{d}|\mathbf{p})$, which in the Gaussian case is up to a constant normalisation given by:

$$\begin{aligned}\chi^2(\mathbf{d}|\mathbf{p}) &= (\mathbf{d} - \mathbf{m}(\mathbf{p}))^T (\mathbf{C}^{\text{obs}})^{-1} (\mathbf{d} - \mathbf{m}(\mathbf{p})) \\ &= \sum_{i,j}^b [\gamma_{t,i}^{\text{obs}} - \gamma_{t,i}^{\text{model}}(\mathbf{p})] \cdot \left[(\mathbf{C}_{\gamma_t}^{\text{obs}})^{-1} \right]_{ij} \cdot [\gamma_{t,j}^{\text{obs}} - \gamma_{t,j}^{\text{model}}(\mathbf{p})] \\ &\quad + \frac{[N_{\text{gal}}^{\text{obs}} - N_{\text{gal}}^{\text{model}}(\mathbf{p})]^2}{[\sigma_{N_{\text{gal}}}^{\text{obs}}]^2}.\end{aligned}\tag{4.1}$$

The 14×14 covariance matrix of the measurement \mathbf{C}^{obs} is built up of the covariance matrix of the **G2L** measurement $\mathbf{C}_{\gamma_t}^{\text{obs}}$ and of the variance in the total number of galaxies $\sigma_{N_{\text{gal}}}^{\text{obs}}$:

$$\mathbf{C}^{\text{obs}} = \begin{pmatrix} & & 0 \\ & \mathbf{C}_{\gamma_t}^{\text{obs}} & \vdots \\ & & 0 \\ 0 & \dots & 0 & (\sigma_{N_{\text{gal}}}^{\text{obs}})^2 \end{pmatrix}.\tag{4.2}$$

Both, $\mathbf{C}_{\gamma_t}^{\text{obs}}$ and $\sigma_{N_{\text{gal}}}^{\text{obs}}$, are obtained by *Jackknifing* (e.g. Efron & Stein 1981, Knight 1999, Shirasaki et al. 2017), which ideally accounts for source shape noise, sampling noise and cosmic variance (Simon+13).

Jackknifing is a resampling technique according to which an estimate of the statistical error on a parameter X is determined by splitting the data into S subsamples, and then leaving out the α -th subsample for an estimate $X_{(\alpha)}^{\text{JK}}$ of X . With the mean of the S resampled estimators $X_{(\alpha)}^{\text{JK}}$,

$$\bar{X}^{\text{JK}} = \frac{1}{S} \sum_{\alpha=1}^S X_{(\alpha)}^{\text{JK}},\tag{4.3}$$

the estimator of the Jackknife error of X given by

$$\sigma_X^{\text{JK}} = \sqrt{\frac{S-1}{S} \sum_{\alpha=1}^S (X_{(\alpha)}^{\text{JK}} - \bar{X}^{\text{JK}})^2}.\tag{4.4}$$

¹<http://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/>

In the case of **CFHTLenS** $S = 129$ pointings are used for Jackknifing, whereas in the case of the simulations $S = 64$ fields are employed.

Applying the Jackknifing method to the total number of galaxies gives

$$\sigma_{N_{\text{gal}}}^{\text{obs}} = \sqrt{\frac{S-1}{S} \sum_{\alpha=1}^S \left(N_{\text{gal},(\alpha)}^{\text{JK}} - \bar{N}_{\text{gal}}^{\text{JK}} \right)^2}. \quad (4.5)$$

This expression follows from the assumption that each subsample (pointing/field) contains on average the same number of galaxies, so that an estimator for the total number of galaxies is given by

$$N_{\text{gal},(\alpha)}^{\text{JK}} = \frac{S}{S-1} \sum_{j \neq \alpha}^S N_{\text{gal},j} \approx \sum_{j \neq \alpha}^S N_{\text{gal},j}, \quad (4.6)$$

if the number of subsamples S is large. The error of neglecting the prefactor in the last step is less than 1% for **CFHTLenS** and less than 2% for the mock data.

The Jackknife estimate for the covariance matrix of the **G2L** observation $\mathbf{C}_{\gamma_t}^{\text{obs}}$ is analogously given by the matrix elements

$$C_{\gamma_t,ij}^{\text{obs}} = \frac{S-1}{S} \sum_{\alpha=1}^S \left[\gamma_{t,(\alpha)}^{\text{JK}}(\theta_i) - \bar{\gamma}_t^{\text{JK}}(\theta_i) \right] \left[\gamma_{t,(\alpha)}^{\text{JK}}(\theta_j) - \bar{\gamma}_t^{\text{JK}}(\theta_j) \right]. \quad (4.7)$$

For the mock data, however, the covariance matrix $\mathbf{C}_{\gamma_t}^{\text{obs}}$ is estimated by the sample covariance of the S subsamples, with elements

$$C_{\gamma_t,ij}^{\text{obs}} = \frac{1}{S-1} \sum_{\alpha=1}^S \left[\gamma_{t,\alpha}(\theta_i) - \bar{\gamma}_t(\theta_i) \right] \left[\gamma_{t,\alpha}(\theta_j) - \bar{\gamma}_t(\theta_j) \right], \quad (4.8)$$

where $\gamma_{t,\alpha}(\theta_i)$ is the **G2L** measurement of the i -th bin in separation θ in the α -th simulated field and $\bar{\gamma}_t(\theta_i)$ the sample mean of the S realisations in the i -th bin in θ . In order to obtain an unbiased estimate for the inverse of the covariance matrix $\mathbf{C}_{\gamma_t}^{\text{obs}}$, as needed for Eq. (4.1), the inverse of Eq. (4.7) and (4.8) need to be rescaled by the factor $(S-2-b)/(S-1)$ for Gaussian statistics, such as assumed here (Hartlap et al. 2007).

By minimising the χ^2 -function (Eq. 4.1) the best estimate for the model parameter values at the maximum likelihood point $\mathbf{p}_{\text{ML}} = (\sigma_{\log(M_h)}, \beta, \log_{10} M_1, \log_{10} M_{\text{th}}, \alpha_{\text{cen}})^T$ is determined. For the fitting, restrictions on the parameter space are set in reference to Coupon et al. (2012) and Clampitt et al. (2017) with $\sigma_{\log(M_h)} \in [0.01; 0.6]$, $\beta \in [0.6; 2]$, $\log_{10} M_1 \in [11; 16]$, $\log_{10} M_{\text{th}} \in [10; 15]$, $\alpha_{\text{cen}} \in [0; 1]$.

4.2.2 Error estimates on best-fit parameters and model predictions

The error on the best-fit parameters is set by the width of the probability distribution function of the model parameters \mathbf{p} around the **ML** point given the data vector \mathbf{d} . The

so-called posterior distribution $\text{post}(\mathbf{p}|\mathbf{d})$ describes the posterior parameter uncertainty and is according to the *Bayes' theorem* given by

$$\text{post}(\mathbf{p}|\mathbf{d}) \propto e^{-\frac{1}{2}\mathcal{L}(\mathbf{d}|\mathbf{p})} \cdot e^{-\frac{1}{2}\mathcal{P}(\mathbf{p})}, \quad (4.9)$$

with the log-likelihood $\mathcal{L}(\mathbf{d}|\mathbf{p})$ and the logarithmic prior distribution $\mathcal{P}(\mathbf{p})$. The posterior distribution is here first approximated by a Gaussian, the importance function $\text{imp}(\mathbf{p}|\mathbf{d})$, and then resampled to its ‘true’ distribution using *importance sampling* (Barlow 1989).

The exact method is as follows. For a Gaussian approximation, Eq. (4.1) can be written in terms of the Fisher information matrix of the model parameters $\mathcal{F}^{\text{model}}$,

$$\mathcal{L}(\mathbf{d}|\mathbf{p}) = (\mathbf{d} - \mathbf{m}(\mathbf{p}))^T (\mathbf{C}^{\text{obs}})^{-1} (\mathbf{d} - \mathbf{m}(\mathbf{p})) \approx (\mathbf{p}_{\text{ML}} - \mathbf{p})^T \mathcal{F}^{\text{model}} (\mathbf{p}_{\text{ML}} - \mathbf{p}), \quad (4.10)$$

for which

$$\mathcal{F}_{ij}^{\text{model}} := \left(\frac{\partial \mathbf{m}(\mathbf{p})}{\partial p_i} \right)^T [\mathbf{C}^{\text{obs}}]^{-1} \left(\frac{\partial \mathbf{m}(\mathbf{p})}{\partial p_j} \right), \quad (4.11)$$

(Cramér 1946, Rao 1965, Tegmark et al. 1997) where \mathbf{C}^{obs} is defined in Eq. (4.2). The prior distribution also uses a Gaussian model,

$$\mathcal{P}(\mathbf{p}) = (\mathbf{p}_{\text{ML}} - \mathbf{p})^T \mathcal{F}^{\text{prior}} (\mathbf{p}_{\text{ML}} - \mathbf{p}). \quad (4.12)$$

Assuming that the priors are uncorrelated, $\mathcal{F}^{\text{prior}}$ contains on the diagonal σ_{ii}^{-2} , where $\sigma_{ii} = \Delta p_i / \sqrt{12}$ and Δp_i is the allowed parameter range. The priors are centred on the **ML** point \mathbf{p}_{ML} . Then the importance function reads

$$\text{imp}(\mathbf{p}|\mathbf{d}) \propto e^{-\frac{1}{2}(\mathbf{p}_{\text{ML}} - \mathbf{p})^T (\mathcal{F}^{\text{model}} + \mathcal{F}^{\text{prior}}) (\mathbf{p}_{\text{ML}} - \mathbf{p})} =: e^{-\frac{1}{2}\mathcal{I}(\mathbf{p}|\mathbf{d})}. \quad (4.13)$$

To account for deviations of the posterior distribution from the Gaussian approximation, importance sampling is applied. Following Press et al. (2007) for each best-fit parameter set S random realisations $\mathbf{p}_{\text{ML}}^{r=1, \dots, S}$ are drawn from the importance function $\text{imp}(\mathbf{p}|\mathbf{d})$ with mean \mathbf{p}_{ML} , representing the best-fit result, and with covariance $\mathbf{C}^{\text{model}}$, introducing the errors on the parameters. An estimate for $\mathbf{C}^{\text{model}}$ is given by the inverse of the Fisher matrix $(\mathcal{F}^{\text{model}})^{-1}$ (Cramér 1946, Rao 1965). A random realisation is then given by $\mathbf{p}_{\text{ML}}^r = \mathbf{p}_{\text{ML}} + \mathbf{A}\mathbf{z}^r$, with $\mathbf{C}^{\text{model}} = \mathbf{A}\mathbf{A}^T$. The matrix \mathbf{A} can be obtained by *Cholesky decomposition* as $\mathbf{C}^{\text{model}}$ is hermitian positive-definite. Through the vector \mathbf{z}^r , which contains five independent standard normal variates components, the Gaussian random process enters the procedure (*Box-Mueller method*). Then every realisation r of the importance function is resampled to the posterior distribution by giving it a weight,

$$w^r := \frac{\text{post}(\mathbf{p}^r|\mathbf{d})}{\text{imp}(\mathbf{p}^r|\mathbf{d})} \propto e^{-\frac{1}{2}[\mathcal{L}(\mathbf{d}|\mathbf{p}^r) + \mathcal{P}(\mathbf{p}^r) - \mathcal{I}(\mathbf{p}^r|\mathbf{d})]}. \quad (4.14)$$

The weights w^r are normalised, such that $\sum_{r=1}^S w^r = 1$.

The weights w^r are used to weight every of the S realisations \mathbf{p}_{ML}^r with respect to their importance. Likewise, the realisations of the predicted quantities, e.g. derived halo quantities, **G2L**, and **G3L** are weighted with w^r . The uncertainty on the model parameters and all derived quantities is then given by the variance of the S weighted random realisations.

4.3 Limitations of the model and the fitting procedure

There are two major difficulties that the proposed method – to predict **G3L** based on the best-fit **HOD** parameters obtained from halo model fits to **G2L** of **CFHTLenS** – could run into: limitations of the fitting routine to converge towards the global minimum of χ^2 and the inability of the model to describe the observables. This section explores these potential problems by fitting the model to mock **G2L** data from the Millennium Simulation. Since the mean parent halo mass, the fraction of central galaxies as well as the **HOD** are known for the mock data, their comparison to the values obtained from the best-fits allows one to determine the accuracy of the fitting routine and the ability of the model to describe **G2L**, independently of the **CFHTLenS** data.

4.3.1 Model fits to **G2L** mock data

The **G2L** signal in the Millennium Simulation as a function of separation θ is shown together with the best-fit model in Fig. 4.1 for the different stellar mass samples, for low redshifts in the left column and high redshifts in the right column. The 1σ error bars of the Millennium Simulation **G2L** signal are smaller than the data points as they only account for cosmic variance and not for shape noise. An expanded version of this figure with the 1σ uncertainties on the **G2L** best-fits is shown in the appendix in Figs. F.1 and F.2. A summary of the best-fit **HOD** parameters and the respective best-fit χ^2 values is given in Table 4.2. The best-fit parameters determine the best-fit **HODs** (Eq. 2.21) for the different galaxy samples. Therefore, the best-fits parameters and the **HODs** are discussed simultaneously in the following. The **HODs** are shown in Fig. 4.2, with the colour coding and line style as in Fig. 4.1.

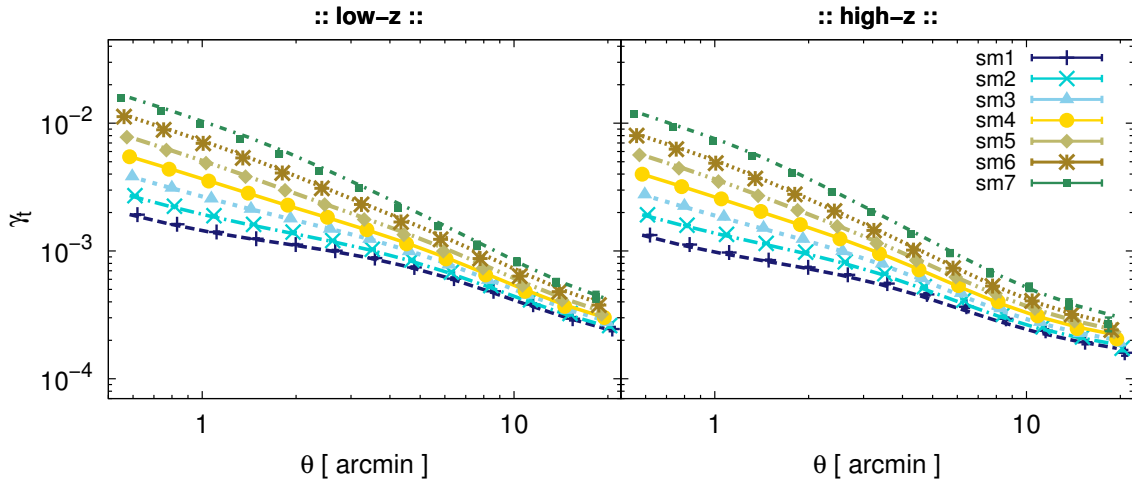


Figure 4.1: Best-fit models (lines) to the **G2L** signals from the Millennium Simulation (symbols) using the **SAM** by Henriques+15 for all stellar mass samples for low- z ($0.2 \leq z_{\text{ph}} < 0.44$; *left*) and high- z ($0.44 \leq z_{\text{ph}} < 0.6$; *right*).

Table 4.2: Best-fit HOD parameters obtained from model fits to G2L from the Millennium Simulation using the SAM by Henriques+15 with $\chi^2/\text{d.o.f.}$ for each stellar mass sample for low- z (*top*) and for high- z (*bottom*). The parameters characterising the halo mass M_1 and M_{th} are given in $h^{-1} M_{\odot}$. Bold values for $\chi^2/\text{d.o.f.}$ indicate a tension between best-fit model and measurement at the 95% confidence level.

low- z						
Sample	$\sigma_{\log(M_h)}$	β	M_1	M_{th}	α_{cen}	$\chi^2/\text{d.o.f.}$
sm1	0.27 ± 0.07	1.06 ± 0.02	$(2.75 \pm 0.26) \times 10^{13}$	$(6.07 \pm 0.36) \times 10^{10}$	0.07 ± 0.01	0.86
sm2	0.27 ± 0.05	1.00 ± 0.01	$(2.83 \pm 0.27) \times 10^{13}$	$(1.24 \pm 0.06) \times 10^{11}$	0.10 ± 0.01	0.84
sm3	0.28 ± 0.04	0.99 ± 0.01	$(2.96 \pm 0.42) \times 10^{13}$	$(2.64 \pm 0.10) \times 10^{11}$	0.16 ± 0.02	2.02
sm4	0.29 ± 0.03	1.02 ± 0.02	$(4.90 \pm 0.44) \times 10^{13}$	$(5.74 \pm 0.24) \times 10^{11}$	0.21 ± 0.02	4.96
sm5	0.20 ± 0.06	1.11 ± 0.02	$(1.12 \pm 0.10) \times 10^{14}$	$(1.07 \pm 0.05) \times 10^{12}$	0.17 ± 0.02	2.48
sm6	0.19 ± 0.08	1.28 ± 0.04	$(2.59 \pm 0.22) \times 10^{14}$	$(2.16 \pm 0.15) \times 10^{12}$	0.12 ± 0.01	2.17
sm7	0.10 ± 0.06	1.37 ± 0.07	$(1.18 \pm 0.22) \times 10^{15}$	$(4.51 \pm 0.23) \times 10^{12}$	0.03 ± 0.01	1.56
high- z						
Sample	$\sigma_{\log(M_h)}$	β	M_1	M_{th}	α_{cen}	$\chi^2/\text{d.o.f.}$
sm1	0.29 ± 0.07	1.09 ± 0.02	$(5.05 \pm 0.71) \times 10^{13}$	$(5.74 \pm 0.39) \times 10^{10}$	0.05 ± 0.01	1.68
sm2	0.33 ± 0.08	0.98 ± 0.02	$(3.32 \pm 0.24) \times 10^{13}$	$(1.18 \pm 0.09) \times 10^{11}$	0.09 ± 0.00	2.14
sm3	0.25 ± 0.05	0.96 ± 0.02	$(3.08 \pm 0.34) \times 10^{13}$	$(2.47 \pm 0.10) \times 10^{11}$	0.15 ± 0.02	2.05
sm4	0.24 ± 0.06	0.95 ± 0.02	$(3.26 \pm 0.37) \times 10^{13}$	$(5.46 \pm 0.28) \times 10^{11}$	0.25 ± 0.02	3.49
sm5	0.40 ± 0.03	1.02 ± 0.01	$(6.21 \pm 0.47) \times 10^{13}$	$(1.32 \pm 0.07) \times 10^{12}$	0.31 ± 0.03	7.15
sm6	0.25 ± 0.04	1.10 ± 0.05	$(2.49 \pm 0.18) \times 10^{14}$	$(2.25 \pm 0.10) \times 10^{12}$	0.13 ± 0.01	3.68
sm7	0.17 ± 0.10	1.37 ± 0.12	$(1.12 \pm 0.18) \times 10^{15}$	$(4.57 \pm 0.42) \times 10^{12}$	0.03 ± 0.00	1.72

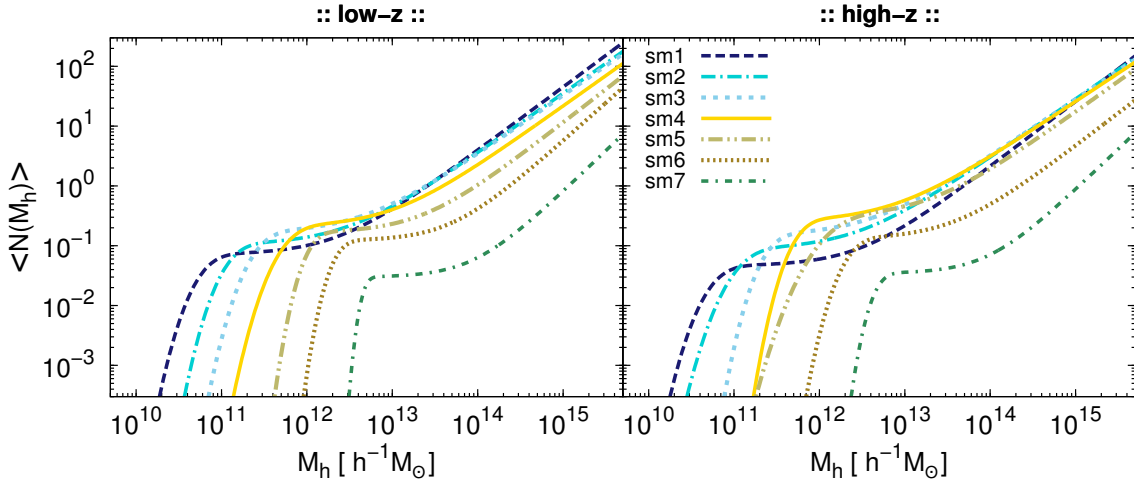


Figure 4.2: Best-fit HODs obtained from model fits to G2L from the Millennium Simulation using the SAM by Henriques+15. Shown is the mean number of galaxies per dark matter halo $\langle N(M_h) \rangle$ as a function of halo mass M_h for all stellar mass samples for low- z (*left*) and high- z (*right*).

The goodness of the fits is quantified by the reduced chi-squared $\chi^2/\text{d.o.f.}$ where the number of degrees of freedom is nine, since 13 G2L data points plus the number of galaxies are fit using a five parameter model. A tension between the model and the measurement at the 95% confidence level is indicated by $\chi^2/9 > 1.88$, which is the case for most of the

stellar mass samples (marked boldface in Table 4.2). The high χ^2 values are, however, not a consequence of the fitting routine failing to converge towards the global minimum of the χ^2 function, but as can be seen in Figs. F.1 and F.2 a consequence of the small error bars of the data without shape noise.

In general all HOD parameters are constrained by the data within an uncertainty of 20% with the exception of the scatter in the mean number of central galaxies at a fixed halo mass $\sigma_{\log(M_h)}$. The best-fit values for $\sigma_{\log(M_h)}$ appear to be largely independent of stellar mass (up to sm4 for low- z), having a mean value of 0.23 for low- z and around 0.28 for high- z . Regarding the halo masses M_{th} and M_1 , the former is always smaller than the latter in agreement with their definition (see Sect. 2.2.1). Both increase with stellar mass, which results in a shift of $\langle N(M_h) \rangle$ towards higher halo masses for higher stellar masses M_* , as can be seen in Fig. 4.2. There is no trend with redshift found for the two mass-related HOD parameters. The parameter β , the slope of the satellite power-law $\langle N_{sat}(M_h) \rangle$, shows a slight increase with increasing stellar mass. Finally, the parameter α_{cen} , which was introduced in this work to allow for halos without a central galaxy, reaches a maximum for medium stellar masses, and decreases towards low and high stellar masses.

4.3.2 Accuracy of predicting the mean halo mass and fraction of centrals

The advantage of the mock data is that the ‘true’ values for the mean parent halo mass $\langle M_h \rangle$, and for the fraction of centrals f_{cen} are known. Thus, the comparison of the predicted values from the best-fit models with the ‘true’ reference values serves as a crucial test for the model. The ‘true’ values of $\langle M_h \rangle$ are taken from Saghiha+17 (Table 7 therein). The ‘true’ f_{cen} values are computed from the mock data as the ratio of the number of central

Table 4.3: Derived quantities from the best-fit parameter sets for the mock lensing data in comparison to the ‘true’ reference values. The results are given for all stellar mass samples for low- z (*top*) and high- z (*bottom*). The comoving galaxy number density is given in $h^3 \text{Mpc}^{-3}$ and the mean halo mass in $h^{-1} \text{M}_\odot$.

low- z						
Sample	N_{tot}^{true}	N_{tot}	f_{cen}^{true}	f_{cen}^V	$\langle M_h \rangle^{true}$	$\langle M_h^z \rangle$
sm1	791888 ± 98216	819985 ± 71801	0.66 ± 0.01	0.79 ± 0.01	(3.8 ± 0.4) × 10 ¹³	(2.37 ± 0.09) × 10 ¹³
sm2	627542 ± 77832	671508 ± 58481	0.65 ± 0.01	0.74 ± 0.01	(4.2 ± 0.4) × 10 ¹³	(2.42 ± 0.06) × 10 ¹³
sm3	492654 ± 61110	552861 ± 65290	0.65 ± 0.01	0.72 ± 0.01	(4.8 ± 0.4) × 10 ¹³	(2.77 ± 0.06) × 10 ¹³
sm4	330867 ± 41054	330543 ± 27222	0.66 ± 0.02	0.75 ± 0.01	(5.5 ± 0.4) × 10 ¹³	(3.06 ± 0.13) × 10 ¹³
sm5	154859 ± 19219	134341 ± 14517	0.69 ± 0.02	0.80 ± 0.01	(6.4 ± 0.4) × 10 ¹³	(3.76 ± 0.13) × 10 ¹³
sm6	45909 ± 5700	47298 ± 4811	0.72 ± 0.02	0.85 ± 0.01	(7.7 ± 0.4) × 10 ¹³	(4.86 ± 0.24) × 10 ¹³
sm7	5510 ± 686	5350 ± 1086	0.77 ± 0.04	0.86 ± 0.01	no data	(6.23 ± 0.65) × 10 ¹³
high- z						
Sample	N_{tot}^{true}	N_{tot}	f_{cen}^{true}	f_{cen}^V	$\langle M_h \rangle^{true}$	$\langle M_h^z \rangle$
sm1	714827 ± 88655	749046 ± 98256	0.70 ± 0.01	0.85 ± 0.01	(2.8 ± 0.4) × 10 ¹³	(2.26 ± 0.10) × 10 ¹³
sm2	844574 ± 104746	844577 ± 50244	0.68 ± 0.01	0.76 ± 0.01	(3.3 ± 0.4) × 10 ¹³	(2.41 ± 0.08) × 10 ¹³
sm3	822437 ± 102015	735547 ± 89614	0.67 ± 0.01	0.72 ± 0.01	(3.8 ± 0.4) × 10 ¹³	(2.78 ± 0.09) × 10 ¹³
sm4	561931 ± 69722	603726 ± 55102	0.67 ± 0.01	0.71 ± 0.01	(4.6 ± 0.4) × 10 ¹³	(3.16 ± 0.10) × 10 ¹³
sm5	270062 ± 33513	326059 ± 22989	0.70 ± 0.01	0.78 ± 0.00	(5.4 ± 0.4) × 10 ¹³	(3.75 ± 0.09) × 10 ¹³
sm6	72471 ± 8990	70904 ± 4407	0.73 ± 0.02	0.82 ± 0.01	(6.3 ± 0.4) × 10 ¹³	(4.49 ± 0.32) × 10 ¹³
sm7	8102 ± 1007	8074 ± 1061	0.77 ± 0.04	0.89 ± 0.01	no data	(6.38 ± 0.83) × 10 ¹³

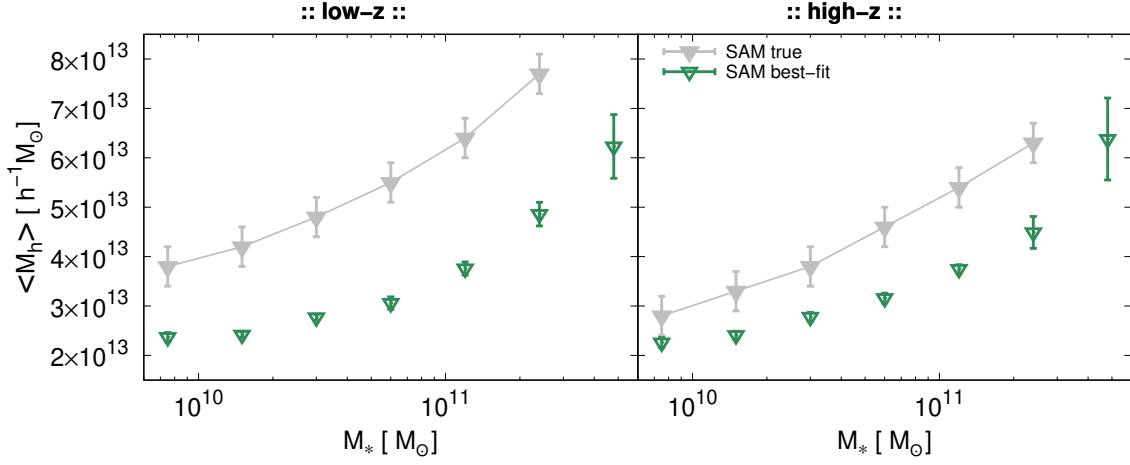


Figure 4.3: Predicted mean halo mass $\langle M_h^{\bar{z}} \rangle$ (open triangles) as a function of stellar mass for the mock data. The ‘true’ values are represented by the filled triangles with the solid line. The error bars represent the standard deviation. For the stellar mass $h = 0.73$ is assumed. There is no ‘true’ $\langle M_h \rangle^{\text{true}}$ value available for sm7.

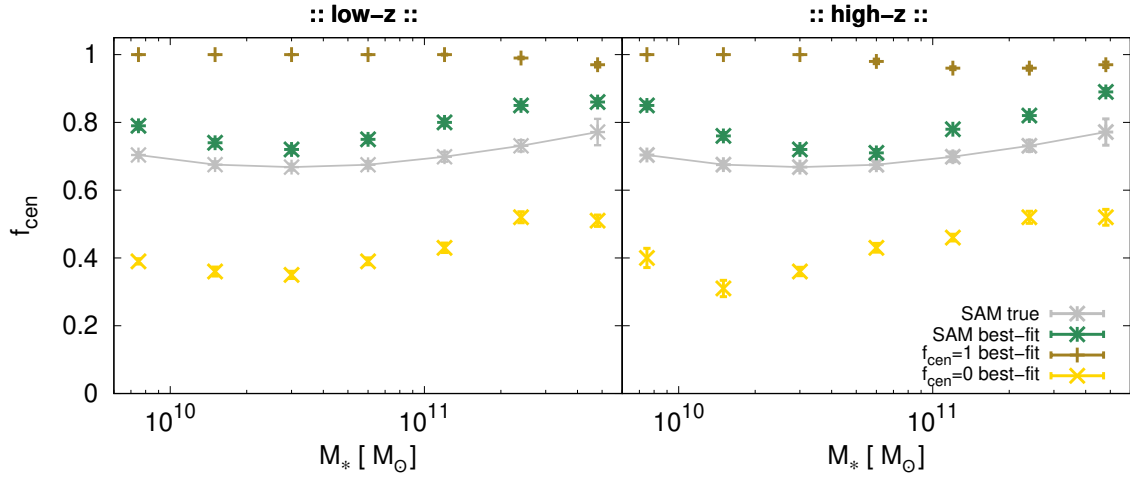


Figure 4.4: Predicted fraction of centrals f_{cen}^V (green stars) as a function of stellar mass obtained for the mock data. The ‘true’ values $f_{\text{cen}}^{\text{true}}$ are represented by the gray stars with solid line. Also shown are the predictions for the centrals-only scenario ($f_{\text{cen}} = 1$, pluses) and the satellites-only scenario ($f_{\text{cen}} = 0$, crosses). The error bars represent the standard deviation. For the stellar mass $h = 0.73$ is assumed.

galaxies to the total number of galaxies per sample. As described in Sect. 2.2.2, the model prediction for the mean halo mass $\langle M_h^{\bar{z}} \rangle$ is estimated for the mean redshift \bar{z} of the galaxy sample, whereas the prediction for the fraction of centrals f_{cen}^V is obtained by integrating over the volume of the respective redshift bin.

The results are presented in Table 4.3, where the reference values from the mock data are listed as ‘true’. Figure 4.3 shows the model prediction for the mean halo mass as a

function of stellar mass in comparison to its ‘true’ reference with 1σ error bars given for both. The model significantly underestimates the halo mass for all stellar masses. The relative deviation is 20 – 45%. Figure 4.4 shows the comparison of the model predictions to the ‘true’ values for the fraction of centrals. The model overpredicts the fraction of centrals, with the relative deviation varying between 5% and 20%.

To clarify the origin of the discrepancies a further test is applied. The model is fitted to the mock G2L for a hypothetical scenario where only centrals exist ($f_{\text{cen}}^{\text{true}} = 1$), and for a scenario where only satellites exist ($f_{\text{cen}}^{\text{true}} = 0$). The resulting best-fit f_{cen} are shown in Fig. 4.4. For the centrals-only scenario only a slight deviation of approximately 5% from $f_{\text{cen}} = 1$ exists for $M_* \gtrsim 10^{11}$. In the case of the satellites-only scenario the discrepancy is much larger reaching up to 55%. This finding is particularly interesting, as it indicates that the description of satellite galaxies in the model is not sufficient to reproduce the mock data.

Ultimately, the relevance of the satellite problem for this work is minor, as the fraction of satellites in simulations and observations is found to be 10% – 50% depending on galaxy-type, stellar mass, and luminosity (Mandelbaum et al. 2006b, 2005). On average this corresponds to the mixed case in Fig. 4.4 for which the deviation of the best-fit from the ‘true’ values is around 15%. However, the inability of the employed halo model to describe satellites correctly opens up opportunities for testing how the halo model needs to be modified in order to represent the galaxy- dark matter connection correctly. Thus, further investigations have been undertaken, which are presented in the following.

Treatment of satellite galaxies in the halo model

Two assumptions underpin the treatment of satellite galaxies in the halo model used in this work. First, the distribution of satellite galaxies in the halo is assumed to follow that of dark matter, i.e. $\tilde{u}_{\text{g}}(k, M_{\text{h}}) = \tilde{u}_{\text{dm}}(k, M_{\text{h}})$. Second, the model assumes that satellite galaxies do not have subhalos. Whether a violation of these assumptions could explain the inability of the model to adequately describe satellites is tested as follows.

The first assumption is relaxed by introducing an additional free model parameter c , such that $[\tilde{u}_{\text{g}}(k, M_{\text{h}})]^c = \tilde{u}_{\text{dm}}(k, M_{\text{h}})$. This allows satellites to have a more or less concentrated distribution inside the halo than dark matter. The fit results, however, show that the concentration parameter c is consistent with unity, implying that the description of the satellites distribution is not the source of the discrepancy.

To test the second assumption, the model is again fitted to mock G2L data for a satellites-only scenario, this time inferring the fraction of centrals in a scenario where all satellite galaxies have a subhalo, versus a scenario where none of the satellite galaxies have a subhalo. The mock G2L data for these two scenarios was kindly generated and provided by Hananeh Saghiha (priv. comm.). As the halo model used in this work ignores subhalos, the model is expected to perform better in the latter scenario. However, this is not the case; the best-fit fraction of centrals with $f_{\text{cen}} \approx 0.2$ is closer to zero for the scenario with subhalos compared to $f_{\text{cen}} \approx 0.5$ for the scenario without subhalos. It seems that the reason why the model counterintuitively describes the scenario with subhalos better is that in this case the G2L signal has a lower amplitude on small scales than in the scenario

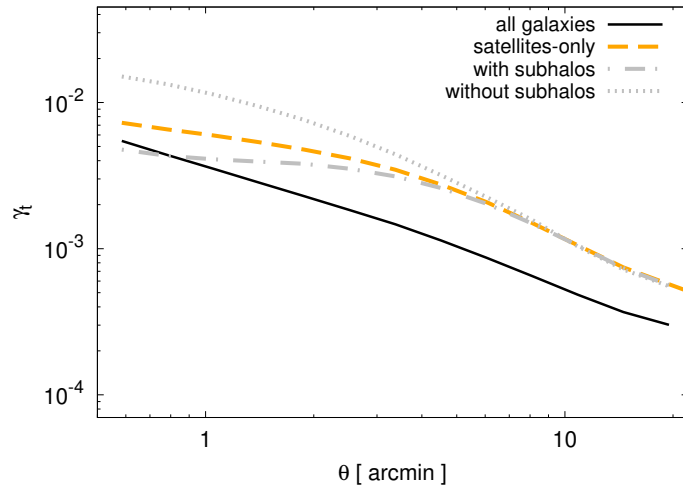


Figure 4.5: Best-fit G2L for all galaxies, satellites-only, satellites-only with subhalos, and satellites-only without subhalos using mock data.

without subhalos (Fig. 4.5). To produce the large central amplitude of the G2L signal found for the scenario without subhalos, the model presumably compensates for the lack of the subhalos by placing a central galaxy in the halo. That might also explain why the model overestimates the fraction of centrals, as seen in Fig. 4.4.

Why the amplitude of G2L for small scales is higher in the scenario without subhalos is not clear. Maybe the satellites with subhalo feel only the potential of the subhalo but not of the parent halo, whereas the satellites without subhalo feel the significantly larger potential of the parent halo. Additionally, satellites without subhalo can be much closer to the centre of the main halo. Rödiger (2009) also studied the G2L signal of satellites-only and found a decrease of the signal for separations smaller than 1 arcmin, which is not observed in this work (Fig. 4.5). The difference might be explained by Rödiger (2009) making only qualitative model predictions and not fitting the model to mock or real data.

4.3.3 Accuracy of the HOD model prediction

Figure 4.6 shows a comparison between $\langle N(M_h) \rangle$ derived from the best-fits and the ‘true’ ones for the stellar mass samples sm1 to sm6 for low- z . The high- z results are presented in the appendix in Fig. F.3. Note that the ‘true’ $\langle N(M_h) \rangle$ are calculated for a single redshift slice of $\bar{z} = 0.32$ for low- z , and $\bar{z} = 0.51$ for high- z (S. Hilbert, priv. comm.).

There are three characteristic differences. First, the amplitude of $\langle N(M_h) \rangle$ is always underestimated by the model for $M_h \gtrsim 10^{12} h^{-1} M_\odot$. This reflects the finding that the fraction of centrals is always overestimated or, equivalently, that the fraction of satellites is always underestimated by the model (Fig. 4.4). Moreover, the model fits are based on the mock data from Saghiiha+17 who apply a selection method to match the observed redshift distributions and thereby the amplitude of $\langle N(M_h) \rangle$ of Simon+13. The ‘true’ $\langle N(M_h) \rangle$ by Stefan Hilbert are, however, given for a single redshift slice resulting in a higher number density and, thus, in a higher number of galaxies. In consequence the

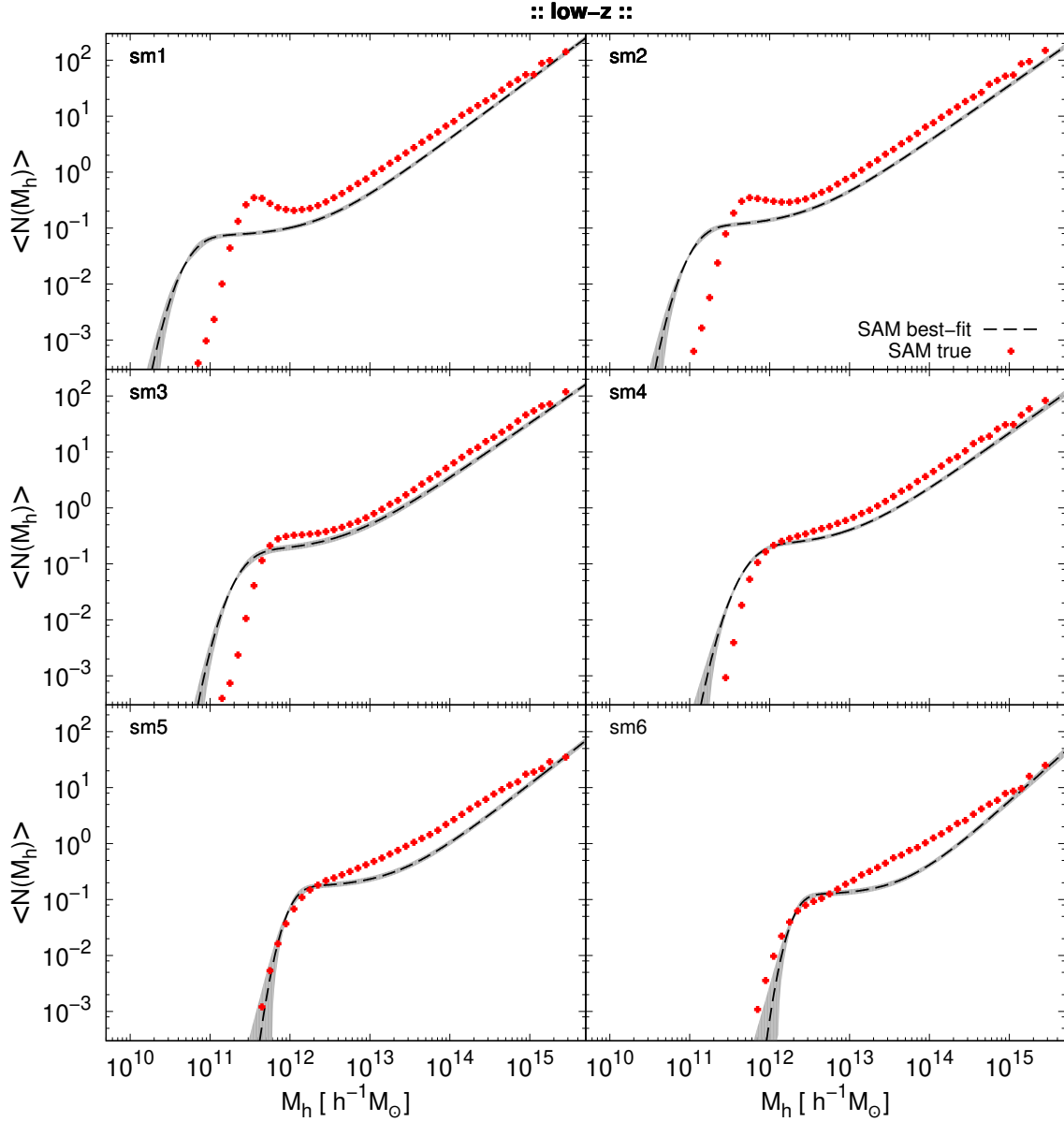


Figure 4.6: Best-fit HODs obtained from the model fits to G2L from the Millennium Simulation using the SAM of Henriques+15 (dashed line with shaded 1σ uncertainty region) in comparison to the ‘true’ HODs (symbols). The results are shown for the various stellar mass samples for low- z .

amplitude of $\langle N(M_h) \rangle$ is higher for the ‘true’ HODs. Second, the best-fit threshold mass M_{th} in the stellar mass samples sm1 to sm4 is found to be smaller than for the ‘true’ HODs. The offset is largest for the smallest stellar mass sample, and the shift of the HOD towards higher halo masses with stellar mass is less pronounced for the ‘true’ HOD compared to the best-fit HOD. Third, the ‘true’ HOD shows a ‘bump’ feature at the position of the threshold mass for sm1 and sm2, which can not be reproduced by the

model, because $\langle N(M_h) \rangle$ is by definition always monotonically increasing in the model. Zheng et al. (2005) study HODs obtained from SAMs and find a bump feature, too, when considering low stellar mass samples defined by bins. They argue that a HOD as the one used in this work does not allow an upper cut-off in $\langle N_{\text{cen}} \rangle$, and is thus not suited to describe samples binned in mass or luminosity.

4.3.4 Conclusions on the limitations of the fitting routine and the model

In this section, possible limitations of the fitting routine and the halo model were investigated using mock G2L data from the Millennium Simulation in combination with the SAM by Henriques+15. In particular, the mean halo mass $\langle M_h \rangle$, the fraction of central galaxies f_{cen} , and the HOD were derived from the best-fit parameter sets. For the mock data these quantities are known, allowing for accuracy tests of the employed halo model.

Tensions between the G2L model predictions and the measurements from the mock data are found for nine out of 14 samples at a confidence level of 95%. Albeit these tensions are a consequence of very small errors on the mock data, the accuracy tests hint at an inability of the halo model to accurately reproduce the mock data: the comparison of the model predictions to the ‘true’ values of $\langle M_h \rangle$ and f_{cen} shows that for some samples the former deviates from the latter by about 45% and 20%, respectively. Because of this large inaccuracy in the $\langle M_h \rangle$ prediction, caution needs to be exercised when characterising halos by their mean halo mass. It is shown that the discrepancy concerning f_{cen} is related to the treatment of satellites in the halo model as the deviations increase for the hypothetical scenario of halos being populated exclusively by satellites. Follow-up investigations on whether the absence of subhalos in the model causes these discrepancies could reveal new insights in the physics of subhalos and satellites. Such investigations are deferred to future work as the satellite-only case is for this work of little relevance: the actual fraction of satellites is around 30%, which leads to an inaccuracy of predicting f_{cen} of on average 15%. Finally, although the model successfully reproduces the overall behaviour of the mock HOD, it overestimates the amplitude of the HOD for smaller halo masses and underestimates the amplitude for larger halo masses.

In conclusion, Fig. 4.1 shows that the model fits are successful and that the model as well as the fitting routine can be applied to fit the CFHTLenS G2L. The discrepancies explored with the accuracy tests indicate where caution needs to be taken when interpreting quantities derived from the best-fit models, as well as promising follow-up research regarding the description of satellites in the model.

4.4 Constraining HODs with CFHTLenS data

In this section, the model fits to the CFHTLenS G2L are presented and discussed for various galaxy samples of stellar mass, luminosity, and galaxy-type. The best-fit parameter sets are used to predict the HOD, the mean halo mass $\langle M_h \rangle$, and the fraction of centrals f_{cen} for each sample. The results are compared to the findings for the Millennium Simulation data presented in Sect. 4.3. Finally, the ratio of stellar to halo mass $M_*(M_h)/M_h$, as derived from the model predictions, is presented.

4.4.1 Model fits to G2L CFHTLenS data

The CFHTLenS G2L signal as a function of separation θ together with the best-fit model predictions is shown in Fig. 4.7 for low- z in the left column and high- z in the right column. The first row shows the results for the stellar mass samples, the second row for the luminosity samples, and the third row for the early- and late-type galaxy samples. The error bars are the 1σ errors obtained from Jackknifing over 129 realisations. The 1σ confidence

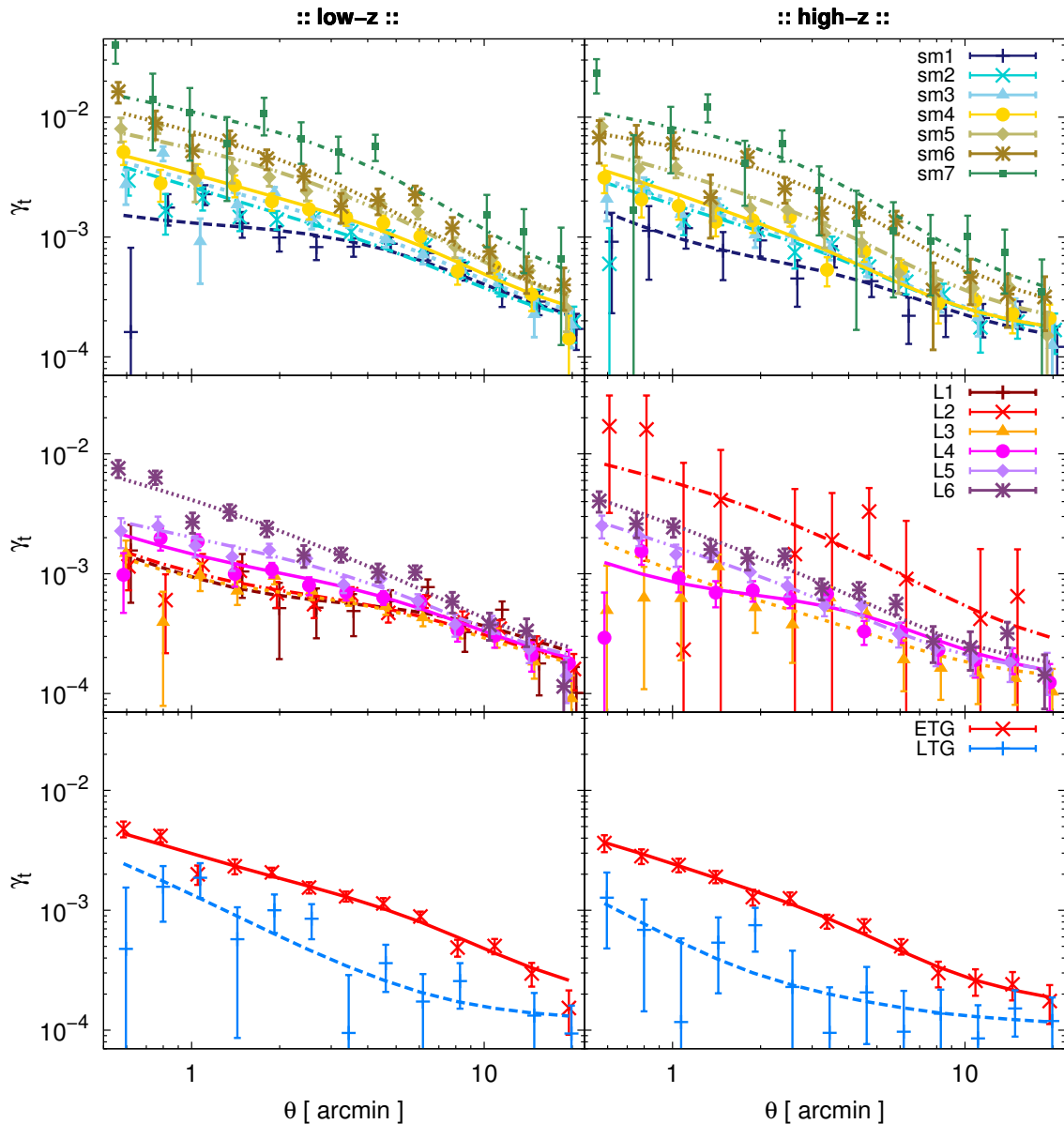


Figure 4.7: Best-fit models (lines) to the CFHTLenS G2L signals (symbols) for the stellar mass (*top*), the luminosity (*middle*), and the galaxy-type samples (*bottom*) for low- z (*left*) and high- z (*right*).

regions around the best-fit G2L signals are omitted here for clarity but can be found in separate figures in the appendix (Figs. F.4 - F.9). A summary of the best-fit HOD parameters and the respective $\chi^2/\text{d.o.f.}$ is given in Table 4.4. The HODs are shown in Fig. 4.8, where the colour coding and line styles are identical to Fig. 4.7 and the respective figures showing the Millennium Simulation results (Figs. 4.1 and 4.2).

For the majority of galaxy samples a good model fit can be stated. In a third of the cases, however, the tension between the model and the observations is at the 95%

Table 4.4: Best-fit HOD parameters obtained from the model fits to the CFHTLenS G2L for low- z (*top*) and high- z (*bottom*). Bold values for $\chi^2/\text{d.o.f.}$ indicate a tension between prediction and observations at the 95% confidence level for nine degrees of freedom. The parameters characterising the halo masses M_1 and M_{th} are given in $h^{-1} M_{\odot}$.

low- z						
Sample	$\sigma_{\log(M_h)}$	β	M_1	M_{th}	α_{cen}	$\chi^2/\text{d.o.f.}$
sm1	0.26 ± 0.15	1.08 ± 0.13	$(2.05 \pm 0.45) \times 10^{13}$	$(3.20 \pm 1.53) \times 10^{10}$	0.03 ± 0.01	1.84
sm2	0.22 ± 0.16	0.92 ± 0.23	$(2.48 \pm 0.70) \times 10^{13}$	$(3.24 \pm 0.41) \times 10^{11}$	0.21 ± 0.04	1.48
sm3	0.30 ± 0.16	0.94 ± 0.20	$(2.72 \pm 0.67) \times 10^{13}$	$(4.28 \pm 1.18) \times 10^{11}$	0.18 ± 0.05	3.77
sm4	0.25 ± 0.16	0.99 ± 0.24	$(3.71 \pm 0.96) \times 10^{13}$	$(5.43 \pm 1.64) \times 10^{11}$	0.16 ± 0.05	1.31
sm5	0.22 ± 0.14	0.83 ± 0.17	$(4.92 \pm 1.21) \times 10^{13}$	$(1.90 \pm 4.28) \times 10^{12}$	0.20 ± 0.06	2.67
sm6	0.22 ± 0.14	0.81 ± 0.17	$(2.58 \pm 2.11) \times 10^{14}$	$(3.61 \pm 0.99) \times 10^{12}$	0.15 ± 0.05	1.67
sm7	0.16 ± 0.16	1.01 ± 0.20	$(8.02 \pm 3.90) \times 10^{14}$	$(1.04 \pm 0.48) \times 10^{13}$	0.05 ± 0.02	1.91
L1	0.35 ± 0.14	1.40 ± 0.28	$(8.44 \pm 1.84) \times 10^{13}$	$(4.29 \pm 1.90) \times 10^{10}$	0.02 ± 0.01	2.96
L2	0.34 ± 0.15	1.19 ± 0.24	$(2.38 \pm 0.93) \times 10^{13}$	$(5.15 \pm 1.54) \times 10^{10}$	0.11 ± 0.03	1.07
L3	0.33 ± 0.16	1.24 ± 0.22	$(1.75 \pm 0.58) \times 10^{13}$	$(4.55 \pm 1.46) \times 10^{10}$	0.14 ± 0.04	1.20
L4	0.21 ± 0.16	1.08 ± 0.18	$(1.56 \pm 0.46) \times 10^{13}$	$(9.07 \pm 3.36) \times 10^{10}$	0.17 ± 0.07	1.24
L5	0.32 ± 0.17	0.85 ± 0.12	$(1.47 \pm 0.40) \times 10^{13}$	$(1.65 \pm 0.54) \times 10^{11}$	0.11 ± 0.04	1.14
L6	0.22 ± 0.13	0.88 ± 0.08	$(4.69 \pm 0.51) \times 10^{13}$	$(9.01 \pm 1.25) \times 10^{11}$	0.27 ± 0.04	2.64
ETG	0.23 ± 0.15	1.09 ± 0.14	$(3.36 \pm 0.61) \times 10^{13}$	$(4.04 \pm 0.83) \times 10^{11}$	0.18 ± 0.05	2.31
LTG	0.31 ± 0.14	0.77 ± 0.26	$(1.96 \pm 1.02) \times 10^{14}$	$(1.56 \pm 0.53) \times 10^{11}$	0.17 ± 0.05	1.75
high- z						
Sample	$\sigma_{\log(M_h)}$	β	M_1	M_{th}	α_{cen}	$\chi^2/\text{d.o.f.}$
sm1	0.23 ± 0.19	1.45 ± 0.25	$(5.86 \pm 1.43) \times 10^{13}$	$(1.18 \pm 0.44) \times 10^{11}$	0.09 ± 0.03	1.29
sm2	0.25 ± 0.17	1.02 ± 0.13	$(2.53 \pm 0.58) \times 10^{13}$	$(3.16 \pm 0.68) \times 10^{11}$	0.21 ± 0.05	2.58
sm3	0.23 ± 0.12	0.97 ± 0.07	$(2.18 \pm 0.22) \times 10^{13}$	$(3.56 \pm 0.64) \times 10^{11}$	0.21 ± 0.03	0.93
sm4	0.25 ± 0.13	0.91 ± 0.21	$(3.37 \pm 0.86) \times 10^{13}$	$(5.50 \pm 1.36) \times 10^{11}$	0.22 ± 0.07	1.52
sm5	0.20 ± 0.11	0.99 ± 0.26	$(4.44 \pm 2.10) \times 10^{13}$	$(1.12 \pm 0.25) \times 10^{12}$	0.18 ± 0.07	1.67
sm6	0.21 ± 0.15	0.91 ± 0.21	$(8.77 \pm 2.47) \times 10^{13}$	$(4.07 \pm 1.38) \times 10^{12}$	0.13 ± 0.07	1.35
sm7	0.19 ± 0.14	1.01 ± 0.22	$(6.16 \pm 5.07) \times 10^{14}$	$(9.72 \pm 4.37) \times 10^{12}$	0.05 ± 0.03	1.32
L1	-	-	-	-	-	-
L2	0.22 ± 0.17	0.94 ± 0.12	$(3.59 \pm 2.06) \times 10^{15}$	$(2.77 \pm 1.38) \times 10^{12}$	0.01 ± 0.00	2.31
L3	0.30 ± 0.17	1.14 ± 0.26	$(7.03 \pm 2.18) \times 10^{13}$	$(1.39 \pm 0.42) \times 10^{11}$	0.10 ± 0.03	1.16
L4	0.16 ± 0.08	1.48 ± 0.14	$(2.89 \pm 0.51) \times 10^{13}$	$(6.84 \pm 1.59) \times 10^{10}$	0.13 ± 0.02	2.26
L5	0.24 ± 0.15	0.83 ± 0.20	$(1.72 \pm 0.76) \times 10^{13}$	$(2.88 \pm 1.12) \times 10^{11}$	0.27 ± 0.14	0.45
L6	0.23 ± 0.16	0.98 ± 0.20	$(3.96 \pm 0.99) \times 10^{13}$	$(6.62 \pm 1.82) \times 10^{11}$	0.27 ± 0.08	1.58
ETG	0.22 ± 0.16	0.99 ± 0.20	$(2.93 \pm 0.75) \times 10^{13}$	$(5.47 \pm 1.42) \times 10^{11}$	0.23 ± 0.07	0.50
LTG	0.21 ± 0.16	1.47 ± 0.33	$(1.22 \pm 1.81) \times 10^{14}$	$(8.38 \pm 3.74) \times 10^{10}$	0.12 ± 0.04	0.55

confidence level, indicated by $\chi^2/9 > 1.88$ (bold values in Table 4.4). Similar to the fits to the Millennium Simulation data (Sect. 4.3.1), no significant trend of $\sigma_{\log(M_h)}$ with stellar mass, luminosity, or galaxy-type is found. This is to be expected with respect to stellar mass, as G2L is sensitive to $\sigma_{\log(M_h)}$ only for very high stellar masses with $M_* > 10^{10.5} M_\odot$ (Leauthaud et al. 2012) corresponding to galaxies from sm3 upwards. With respect to luminosity, Coupon et al. (2012) find a rather constant value of ~ 0.3 for red galaxies, which agrees within the errors with the best-fit value found in this work. In contrast, Zheng et al. (2007) and Zehavi et al. (2011) find a decrease of $\sigma_{\log(M_h)}$ with luminosity ($-22 < M_r < -18$) by more than a factor of two. The quantitative differences possibly result from the different sample selection criteria as Zheng et al. (2007), Zehavi et al. (2011), and Coupon et al. (2012) use luminosity threshold samples. In general, all of the cited works mention difficulties constraining $\sigma_{\log(M_h)}$.

In contrast, the halo masses M_{th} and M_1 are well constrained by the data. Both increase with stellar mass in agreement with the results obtained from the fits to the mock data (Fig. 4.9). As a consequence, with higher stellar mass the HOD $\langle N(M_h) \rangle$ shifts towards higher halo mass, as can be seen in Fig. 4.8. As shown in Fig. 4.10, M_{th} shows a clear increase with respect to increasing luminosity for low- z , and for $M_r < -20$ for high- z , which is in agreement with Zheng et al. (2007), Zehavi et al. (2011), and Coupon et al. (2012). The apparent increase of the halo masses with decreasing luminosity for $M_r > -20$ in the high- z sample cannot be trusted as the data for L1 is not available and the L2 sample is incomplete. In agreement with Coupon et al. (2012), the results of this work indicate that in order to host a first galaxy, a dark matter halo needs to be more massive if the first galaxy is an early-type galaxy (higher M_{th}) instead of a late-type galaxy (lower M_{th}). At the same time, the threshold mass for satellites M_1 is higher for the late-type sample than for the early-type one.

The parameter β , which is the satellite power-law slope, is well constrained to $\beta \approx 1$, in agreement with Zheng et al. (2007), Zehavi et al. (2011) and Coupon et al. (2012).

The amplitude α_{cen} of the ‘step function’ of the HOD for CFHTLenS shows the same behaviour as for the mock data: it peaks for intermediate stellar mass bins. Again no clear trend with luminosity or galaxy-type can be seen for α_{cen} .

4.4.2 Mean halo mass and fraction of centrals

From the best-fit parameters the mean halo mass $\langle M_h^z \rangle$, and the fraction of centrals f_{cen}^V are predicted for all galaxy samples as described in Sect. 2.2.2. A summary of the results is presented in Table 4.5.

The derived mean halo masses for CFHTLenS and the mock data are shown as a function of stellar mass in Fig. 4.9. The mean halo mass increases with stellar mass and, as in the case of M_1 and M_{th} , a good agreement between the results for the observations and the simulations is found. This is not too surprising, because the G2L of the SAM is close to the one from CFHTLenS, as shown in Saghiiha+17. Since Saghiiha+17 use stellar mass samples only, a similar comparison with respect to galaxy luminosity is not possible. Nevertheless, the CFHTLenS results for the halo masses as a function of luminosity are

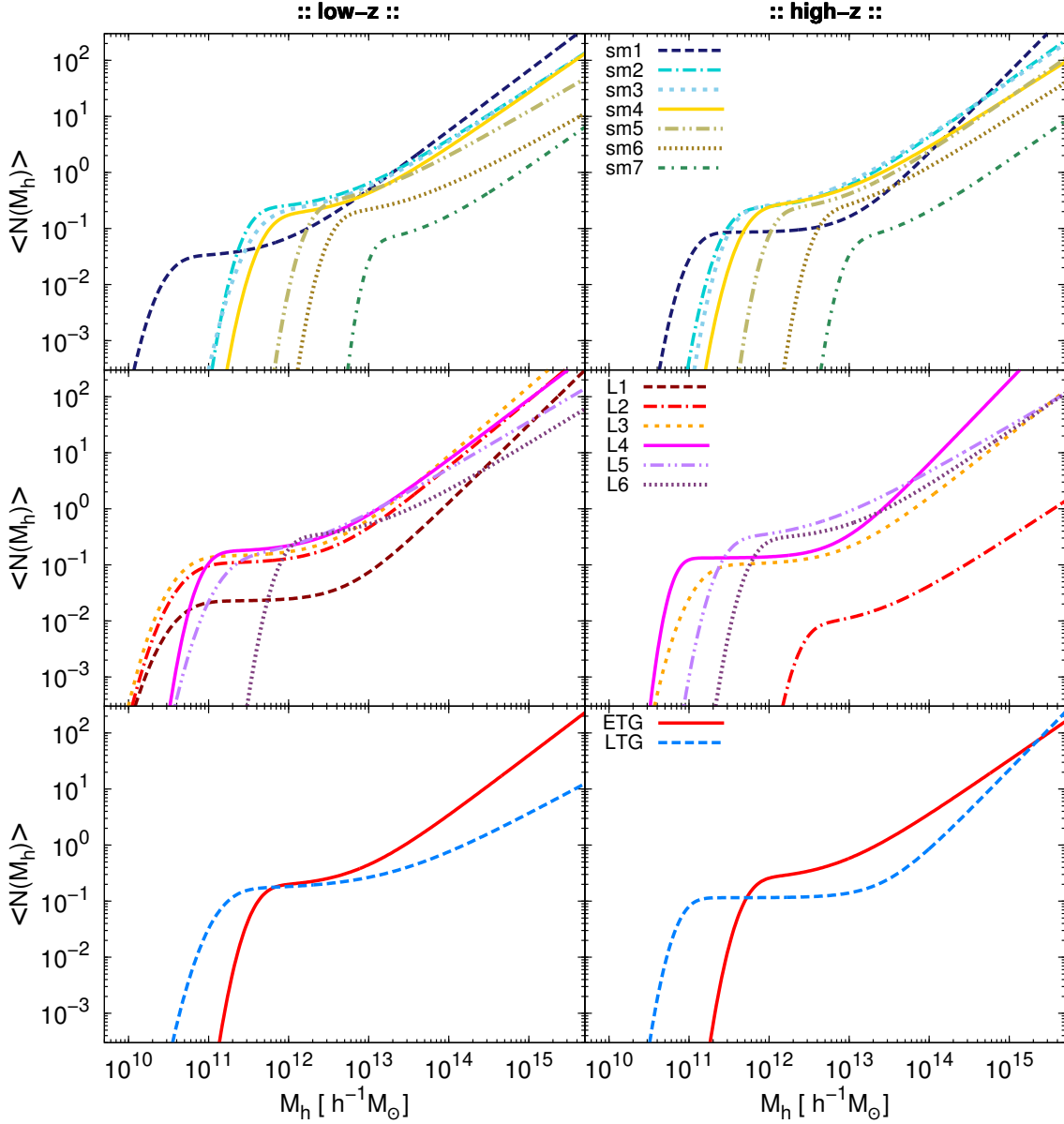


Figure 4.8: Best-fit HODs for CFHTLenS. Shown is the mean number of galaxies per dark matter halo $\langle N(M_h) \rangle$ as a function of halo mass M_h for all stellar mass (*top*), luminosity (*middle*), and galaxy-type samples (*bottom*) for low- z (*left*) and high- z (*right*).

shown in Fig. 4.10.

The mean halo mass $\langle M_h^z \rangle$ for early-type galaxies is found to be about an order of magnitude higher than that of late-type galaxies (Table 4.5). This difference is more pronounced for the high- z sample. This reflects that early-type galaxies are more likely to inhabit galaxy groups and clusters than late-type galaxies (Dressler 1980). This is further supported by the finding that the fraction of centrals is found to be higher for late-type

Table 4.5: Derived quantities from the best-fit parameter sets for the CFHTLenS data. The results are given for all stellar mass, luminosity and galaxy-type samples for low- z (*top*) and high- z (*bottom*). The galaxy number density is given in $h^3\text{Mpc}^{-3}$ and the mean halo mass in $h^{-1}M_\odot$.

low- z					
Sample	N^{obs}	N	\bar{n}_g^V	f_{cen}^V	$\langle M_h^{\bar{z}} \rangle$
sm1	59272 ± 5264	60902 ± 5006	$(4.24 \pm 0.35) \times 10^{-3}$	0.70 ± 0.11	$(2.52 \pm 0.51) \times 10^{13}$
sm2	46609 ± 4157	49125 ± 4025	$(3.42 \pm 0.28) \times 10^{-3}$	0.70 ± 0.19	$(1.88 \pm 0.68) \times 10^{13}$
sm3	36688 ± 3278	36698 ± 3012	$(2.56 \pm 0.21) \times 10^{-3}$	0.65 ± 0.17	$(2.34 \pm 0.68) \times 10^{13}$
sm4	25068 ± 2252	24692 ± 1954	$(1.72 \pm 0.14) \times 10^{-3}$	0.66 ± 0.17	$(2.86 \pm 0.85) \times 10^{13}$
sm5	11908 ± 1078	12104 ± 1002	$(8.43 \pm 0.70) \times 10^{-4}$	0.50 ± 0.15	$(3.49 \pm 1.24) \times 10^{13}$
sm6	3498 ± 328	3616 ± 294	$(2.52 \pm 0.20) \times 10^{-4}$	0.65 ± 0.12	$(3.48 \pm 1.00) \times 10^{13}$
sm7	403 ± 54	412 ± 45	$(2.87 \pm 0.31) \times 10^{-5}$	0.56 ± 0.16	$(8.19 \pm 1.73) \times 10^{13}$
L1	27674 ± 2463	27559 ± 2313	$(1.92 \pm 0.16) \times 10^{-3}$	0.92 ± 0.20	$(2.03 \pm 0.63) \times 10^{13}$
L2	119491 ± 10618	114541 ± 10387	$(7.98 \pm 0.72) \times 10^{-3}$	0.88 ± 0.07	$(1.60 \pm 0.56) \times 10^{13}$
L3	166448 ± 14779	168220 ± 14748	$(1.17 \pm 0.10) \times 10^{-2}$	0.89 ± 0.10	$(1.80 \pm 0.51) \times 10^{13}$
L4	112975 ± 10045	111846 ± 9724	$(7.79 \pm 0.68) \times 10^{-3}$	0.79 ± 0.17	$(1.90 \pm 0.44) \times 10^{13}$
L5	67115 ± 5990	67236 ± 9077	$(4.68 \pm 0.63) \times 10^{-3}$	0.55 ± 0.14	$(1.74 \pm 0.31) \times 10^{13}$
L6	23863 ± 2148	24456 ± 1599	$(1.70 \pm 0.11) \times 10^{-3}$	0.70 ± 0.04	$(2.10 \pm 0.26) \times 10^{13}$
ETG	30261 ± 2733	32891 ± 2340	$(2.29 \pm 0.16) \times 10^{-3}$	0.73 ± 0.11	$(2.89 \pm 0.52) \times 10^{13}$
LTG	60714 ± 5434	60997 ± 5331	$(4.25 \pm 0.37) \times 10^{-3}$	0.92 ± 0.09	$(3.26 \pm 0.23) \times 10^{12}$
high- z					
Sample	N^{obs}	N	\bar{n}_g^V	f_{cen}^V	$\langle M_h^{\bar{z}} \rangle$
sm1	52662 ± 4696	54452 ± 3602	$(2.71 \pm 0.18) \times 10^{-3}$	0.93 ± 0.10	$(3.00 \pm 0.85) \times 10^{13}$
sm2	62124 ± 5523	67042 ± 8124	$(3.33 \pm 0.40) \times 10^{-3}$	0.76 ± 0.14	$(2.47 \pm 0.53) \times 10^{13}$
sm3	61210 ± 5445	63989 ± 1727	$(3.18 \pm 0.09) \times 10^{-3}$	0.69 ± 0.05	$(2.60 \pm 0.31) \times 10^{13}$
sm4	42788 ± 3831	43965 ± 3646	$(2.19 \pm 0.18) \times 10^{-3}$	0.70 ± 0.16	$(2.25 \pm 0.96) \times 10^{13}$
sm5	20806 ± 1886	20339 ± 1458	$(1.01 \pm 0.07) \times 10^{-3}$	0.62 ± 0.14	$(4.01 \pm 1.11) \times 10^{13}$
sm6	5426 ± 511	5473 ± 466	$(2.72 \pm 0.23) \times 10^{-4}$	0.43 ± 0.20	$(6.29 \pm 1.53) \times 10^{13}$
sm7	606 ± 70	588 ± 79	$(2.93 \pm 0.39) \times 10^{-5}$	0.53 ± 0.20	$(8.95 \pm 2.33) \times 10^{13}$
L1	-	-	-	-	-
L2	269 ± 33	304 ± 23	$(1.51 \pm 0.11) \times 10^{-5}$	0.66 ± 0.19	$(4.22 \pm 0.99) \times 10^{13}$
L3	57558 ± 5147	57796 ± 4600	$(2.87 \pm 0.23) \times 10^{-3}$	0.92 ± 0.10	$(1.30 \pm 0.57) \times 10^{12}$
L4	139604 ± 12426	136501 ± 13089	$(6.79 \pm 0.65) \times 10^{-3}$	0.92 ± 0.03	$(3.60 \pm 0.73) \times 10^{13}$
L5	101054 ± 8994	100183 ± 9479	$(4.98 \pm 0.47) \times 10^{-3}$	0.69 ± 0.22	$(1.54 \pm 0.48) \times 10^{13}$
L6	41590 ± 3745	41451 ± 3266	$(2.06 \pm 0.16) \times 10^{-3}$	0.76 ± 0.13	$(2.41 \pm 0.71) \times 10^{13}$
ETG	46387 ± 4182	46228 ± 3663	$(2.30 \pm 0.18) \times 10^{-3}$	0.71 ± 0.12	$(2.80 \pm 0.88) \times 10^{13}$
LTG	96247 ± 8592	93770 ± 8018	$(4.66 \pm 0.40) \times 10^{-3}$	0.99 ± 0.03	$(7.06 \pm 0.48) \times 10^{12}$

galaxies with approximately 95%, compared to only 72% for early-types. Whereas late-type galaxies are more likely to be isolated field galaxies, early-type galaxies share the parent halo with other galaxies, only one of which can be the central galaxy.

The derived fraction of centrals for CFHTLenS as a function of stellar mass is shown in Fig. 4.11 together with the values obtained from the model fits to the mock G2L data. A good agreement is seen for stellar masses smaller than $10^{11} M_\odot$. For higher stellar masses,

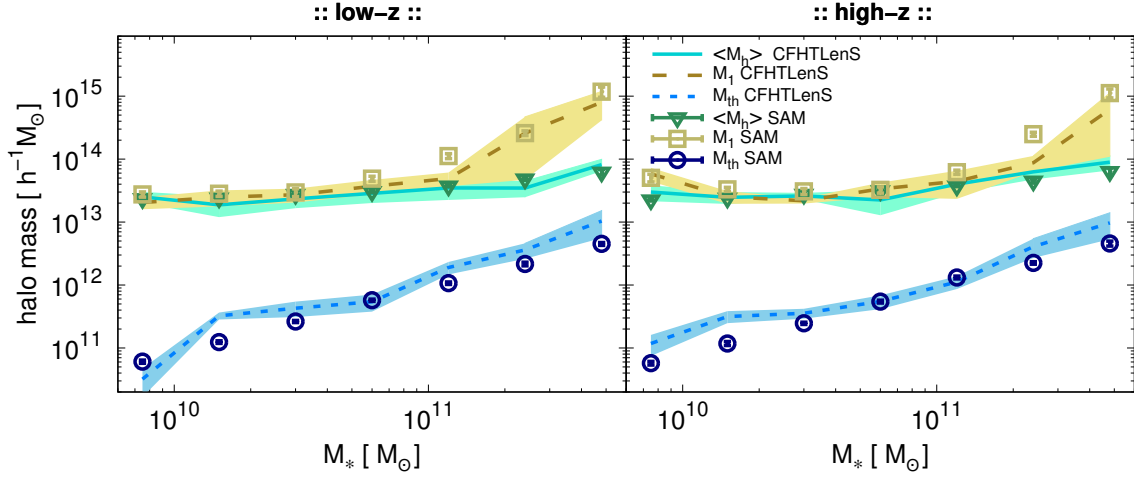


Figure 4.9: Mean halo mass $\langle M_h^z \rangle$ (solid line), satellite threshold mass M_1 (long dashed line) and threshold mass M_{th} (short dashed line) as a function of stellar mass derived from the best-fits to the **CFHTLenS** data. Shaded areas give the 1σ error from Jackknifing. Also shown are the results from the mock data for $h = 0.73$: $\langle M_h^z \rangle$ (triangles), M_1 (squares) and M_{th} (circles). The error bars represent the standard deviation. For the stellar mass $h = 0.73$ is assumed.

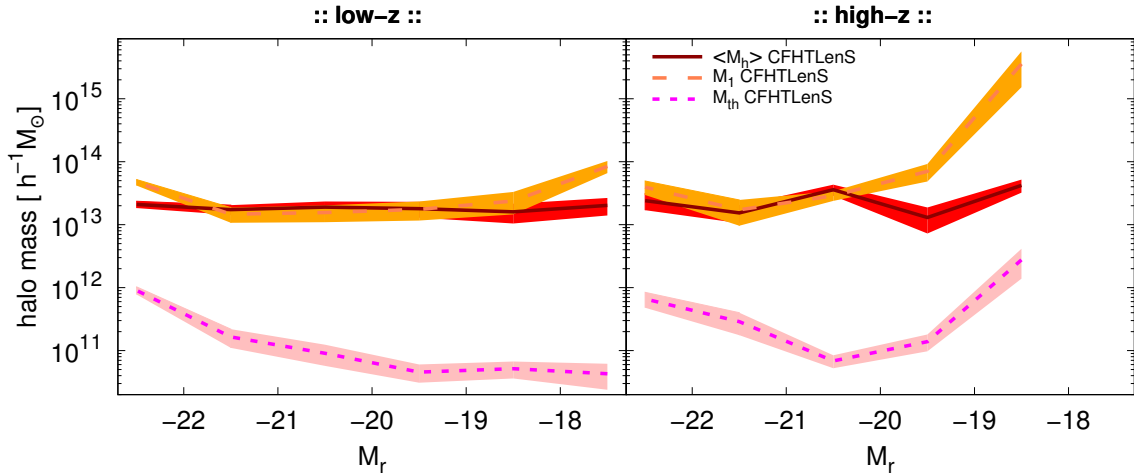


Figure 4.10: Mean halo mass $\langle M_h^z \rangle$ (solid line), satellite threshold mass M_1 (long dashed line) and threshold mass M_{th} (short dashed line) as a function of luminosity derived from the best-fits to the **CFHTLenS** data. Shaded areas indicate 1σ confidence regions. For the luminosity $h = 0.73$ is assumed. There is no L1 data for high-z.

an increasing discrepancy between the two is found, as f_{cen} for **CFHTLenS** decreases with stellar mass in contrast to the mock data. The disagreement can possibly be explained by the bias explored in Fig. 4.4.

Velander et al. (2014) also fit a halo model to the **CFHTLenS G2L** signal. Only for the largest and smallest stellar masses considered, however, their results are found to

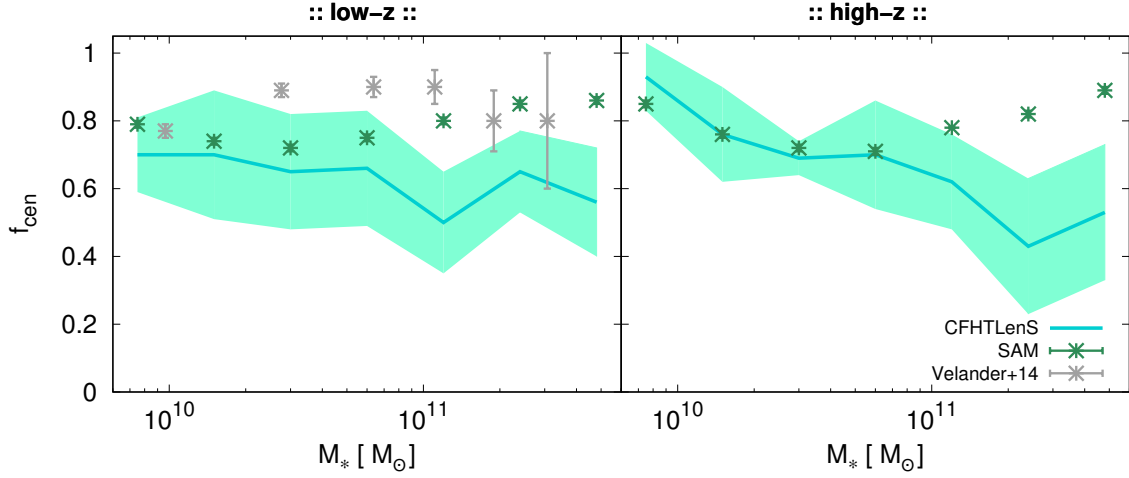


Figure 4.11: Fraction of centrals $f_{\text{cen}}^{\text{V}}$ as a function of stellar mass obtained from the best-fit models to the CFHTLenS G2L signal (solid line, shaded area gives 1σ uncertainty). In comparison the results obtained from the fits to the mock data (green stars), and the results from Velander et al. (2014) (gray stars) are shown. For the stellar mass $h = 0.73$ is assumed.

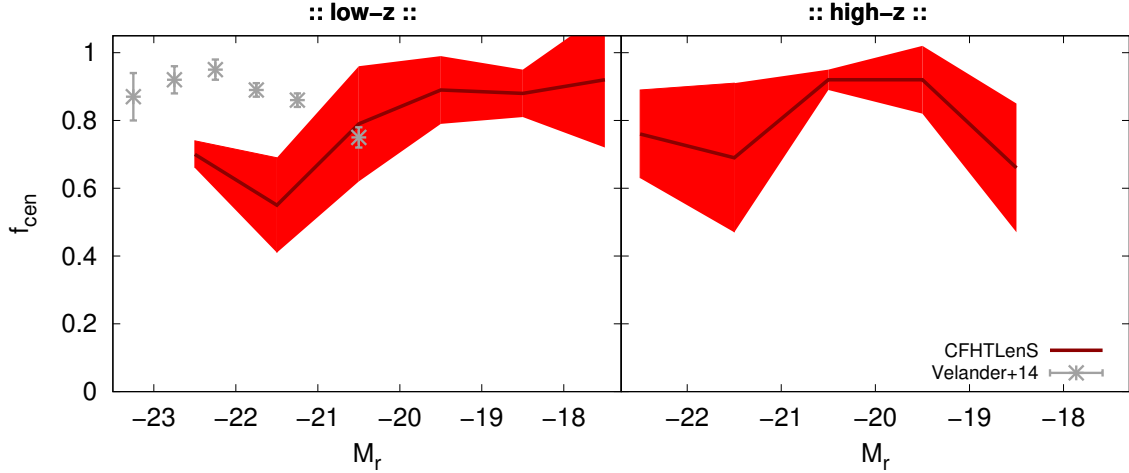


Figure 4.12: Fraction $f_{\text{cen}}^{\text{V}}$ of centrals as a function of luminosity obtained from the best-fit models to the CFHTLenS G2L signal (solid line, shaded area gives 1σ uncertainty). In comparison the results from Velander et al. (2014) (gray stars) are shown. For the luminosity $h = 0.73$ is assumed.

agree within uncertainties with the results of this work. Moreover, application of the same stellar mass bin scatter correction as in their work only increases the disagreement. This disagreement is most likely due to the different halo models employed, as Velander et al. (2014) use a halo model which includes subhalos. Figure 4.12 shows the best-fit f_{cen} as a function of luminosity, again in comparison to Velander et al. (2014). In the overlapping luminosity range ($-22.5 < M_r < -20.5$) the results disagree.

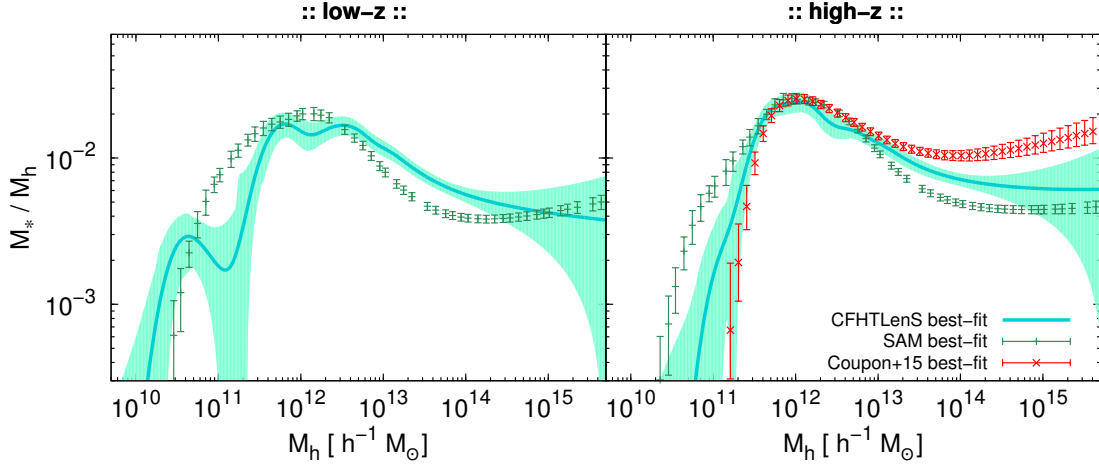


Figure 4.13: Central $M_*(M_h)/M_h$ relationship as a function of halo mass M_h obtained from the best-fit models to the G2L signal from CFHTLenS data (solid line with 1σ error-band). The findings are compared to the best-fit results for the mock data (dash symbols with 1σ error bars), and to the results from Coupon et al. (2015) (cross symbols with 1σ error bars).

4.4.3 Ratio of stellar to halo mass

An important quantity in the halo model context is the stellar to halo mass fraction $M_*(M_h)/M_h$ as it not only carries information about the physical relation of stellar and dark matter, but also about the history of star formation in halos (e.g. Coupon et al. 2015; Leauthaud et al. 2011). The HOD $\langle N(M_h) \rangle = \langle N(M_h) | M_* \rangle$, derived from the best-fits to the CFHTLenS data, can be converted to the ratio of stellar mass to halo mass $M_*(M_h)/M_h$ as described in Sect. 2.2.3. The results for CFHTLenS are presented together with the best-fit results to the mock data in Fig. 4.13. In general a good agreement is seen, particularly around the peak at $M_h \approx 10^{12} h^{-1} M_\odot$. However, the SAM over-predicts (under-predicts) the amount of stellar mass for lower (higher) halo masses. This reflects that the best-fit model for the mock data overestimates the amplitude of the HOD for small halo masses and underestimates the amplitude for large halo masses compared to the ‘true’ HOD (see Fig. 4.6).

Coupon et al. (2015) infer the $M_*(M_h)/M_h$ ratio from a joint lensing, clustering, and abundance analysis using CFHTLenS and VIPERS data. A comparison of their results with this work is interesting, because the CFHTLenS data are the basis of the analysis in both cases. The results of Coupon et al. (2015) are shown in the right panel of Fig. 4.13, although note that their results are for a mean redshift of $\bar{z} = 0.8$, which falls outside of the high- z sample used in this work. A good agreement with the findings of this work can be seen for $M_h \lesssim 3 \times 10^{13} h^{-1} M_\odot$. For more massive halos, the model employed in this work underestimates the fraction of stellar mass compared to the results of Coupon et al. (2015). This lack of stellar mass is most likely due to the incompleteness of the galaxy selection in this analysis.

Besides obtaining the best-fit parameter sets needed for the G3L predictions, this result for the $M_*(M_h)/M_h$ relationship is the main physical results of this chapter and the best

verification that the method, model and fitting routine successfully keep up with state of the art methods in describing the galaxy-dark matter relation.

4.4.4 Conclusions on the model fits to G2L from CFHTLenS

This section is concerned with the halo model fits to G2L measured with CFHTLenS for 29 galaxy samples. The resulting best-fit HOD parameter sets are the basis to generate G3L predictions which are ultimately confronted with their measured counterparts from CFHTLenS (Chapter 6). The results of the G2L model fits are presented and discussed in this section. For two thirds of the 29 galaxy samples a good model fit at a confidence level of 95% can be stated, and the obtained best-fit values for the HOD parameters are in an overall agreement with the literature.

Moreover, for each galaxy sample the mean halo mass $\langle M_h^z \rangle$, and the fraction of centrals f_{cen}^V are derived from the best-fit parameter set. The results are compared to the results obtained from the model fits to the mock data (presented in Sect. 4.3). In concordance with the latter the mean halo mass is found to increase with stellar mass. Furthermore, the halo model predicts the mean halo mass to be more than a magnitude higher for early-type compared to late-type galaxies, with the difference increasing with redshift. This can be interpreted as early-types residing typically in galaxy groups or clusters, whereas late-types are usually found to be field galaxies. The results for the fraction of centrals fit this scenario: the fraction of centrals for late-types is 95% compared to 72% for early-types. Whereas the model predictions for f_{cen}^V agree with the ones from the mock data for stellar masses below $10^{11} M_\odot$, an increasing discrepancy is found for higher stellar masses.

Finally, the HOD $\langle N(M_h) \rangle = \langle N(M_h) | M_* \rangle$ is converted to the ratio of stellar mass to halo mass $M_*(M_h)/M_h$. The results are compared on the one hand to the corresponding prediction obtained from the best-fits for the mock data, and on the other hand to results of another G2L analysis for CFHTLenS by Coupon et al. (2015). For $M_h \lesssim 10^{13} h^{-1} M_\odot$ the results for $M_*(M_h)/M_h$ are in qualitative agreement with the ones obtained from the mock data, and in quantitative agreement with the ones by Coupon et al. (2015). However, for higher halo masses the ratio of stellar mass to halo mass is found to be lower for the mock data and higher for Coupon et al. (2015) than predicted here. The difference to the findings by Coupon et al. (2015) is due to the incompleteness of the galaxy selection in the present analysis.

Chapter 5

First-time prediction of excess mass maps using the halo model – dependence of G3L on galaxy & model properties

This chapter presents excess mass maps as predicted for the first time using a halo model. Excess mass maps were first measured a decade ago in the Red-Sequence Cluster Survey (RCS) (Simon et al. 2008), and subsequently in CFHTLenS (Simon+13, Simon et al. 2019; hereafter Simon+19). These measurements revealed that, first, excess mass can only be found if the two lens galaxies are physically close. Second, excess mass is sensitive to galaxy-type: for a given lens-lens separation a ‘cross-shaped’ structure of excess mass envelops pairs of early-type galaxies, whereas a trough of excess mass is centred between pairs of late-type galaxies. Recently, Simon+19 published a comparison of excess mass maps measured for a combined sm1-sm6 sample of CFHTLenS galaxies to predictions from the SAM of Henriques+15 implanted within the Millennium Simulation. Simon+19 report generally good agreement with some discrepancies, e.g. they find a vertical bulge-like structure in the CFHTLenS maps, which is absent in the SAM maps.

In this chapter it is examined, first, whether a ‘standard’ halo model as introduced in Chapter 2 predicts the same trends with lens-lens separation and galaxy-type for the excess mass as measured by Simon et al. (2008), Simon+13 and Simon+19. The results are presented in Sect. 5.2. Second, to check whether the halo model predicts the map features measured by Simon+19 for CFHTLenS galaxy pairs, halo model predictions for the combined sm1-sm6 sample are confronted with their observational counterpart (Sect. 5.5).

Furthermore, the dependence of excess mass on stellar mass, luminosity, and redshift is explored. These dependencies will eventually be measured in future surveys, and their study is crucial to illuminate the galaxy-halo connection. All predictions are generated based on the best-fit parameter sets obtained from model fits to G2L for the 29 galaxy samples of CFHTLenS. This approach not only allows to investigate how the amplitude and map features depend on galaxy properties but also to quantify what is to be expected for future observations. The results are presented in Sect. 5.2 and are discussed hand in hand with the corresponding predictions of the halo model for the aperture statistic $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$. Although excess mass maps and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are both representations of the galaxy-galaxy-shear correlator \mathcal{G} , and thus contain the same information, each has

particular advantages. Whereas excess mass maps provide a visualisation of the matter environment around galaxy pairs, and display the information contained in non-equilateral triangle configurations, each map contains by definition only information for a specific projected lens-lens separation. Aperture mass statistics are more convenient for investigating the scale-dependence of the third-order galaxy-mass correlation.

The rest of the chapter is dedicated to study the dependence of **G3L** on halo model properties. In Sect. 5.3 the contributions of the one-halo, two-halo and three-halo term of the halo model to the excess mass maps and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are studied, and it is determined which term dominates the total signal on which scales. This knowledge is important for the interpretation of the measurements and for future approximations to the halo model formalism to make the costly computation of model predictions faster and direct model fits to **G3L** feasible.

To explore the sensitivity of **G3L** with respect to the **HOD** parameters, in Sect. 5.4 excess mass maps and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are predicted for the five model parameters being varied individually by $\pm 20\%$ around their best-fit value, while keeping the other four fixed. Particular focus is set on whether map features show changes that are identifiable with specific **HOD** parameters. This part of the analysis shows which **HOD** parameters can be constrained with **G3L** in future surveys, and which galaxy samples are better suited for this purpose.

The content of this chapter is being prepared for submission to *Astronomy & Astrophysics*.

5.1 Method

5.1.1 Predicting **G3L**

The **G3L** model predictions are based on the model fits to the **CFHTLenS G2L** for the galaxy samples defined in Table 4.1. Specifically, the best-fit parameter set for each of the 29 samples is used as input for the halo model. The model predicts the excess mass maps and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ for a given parameter set by numerically implementing the analytical formalism described in Sect. 3.2. Since no **G2L** measurement is available for the L1 high- z sample, there are also no **G3L** model predictions possible for this sample. Except for the five **HOD** parameters, all other parameters of the model are fixed.

The model predictions can in principle be produced for all lens-lens separations and aperture scale radii. For conciseness, only seven lens-lens separations between 0.5 and 3 arcmin are considered for the excess mass maps, which corresponds roughly to a projected physical separation of $250h^{-1}\text{kpc}$ to $1h^{-1}\text{Mpc}$, thereby focussing on pairs of galaxies inside groups and clusters. The $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ predictions are for aperture scale radii between 0.1 and 40 arcmin, which covers the range probed by **CFHTLenS** (1 – 10 arcmin).

5.1.2 Uncertainties in predicting **G3L**

Uncertainties in the **G3L** predictions arise from uncertainties in the measured **CFHTLenS G2L** signal, to which the model is fit to obtain the best-fit parameters. The treatment of these errors, discussed in Sects. 4.2.2 and 6.2.2, is relevant for the quantitative comparison

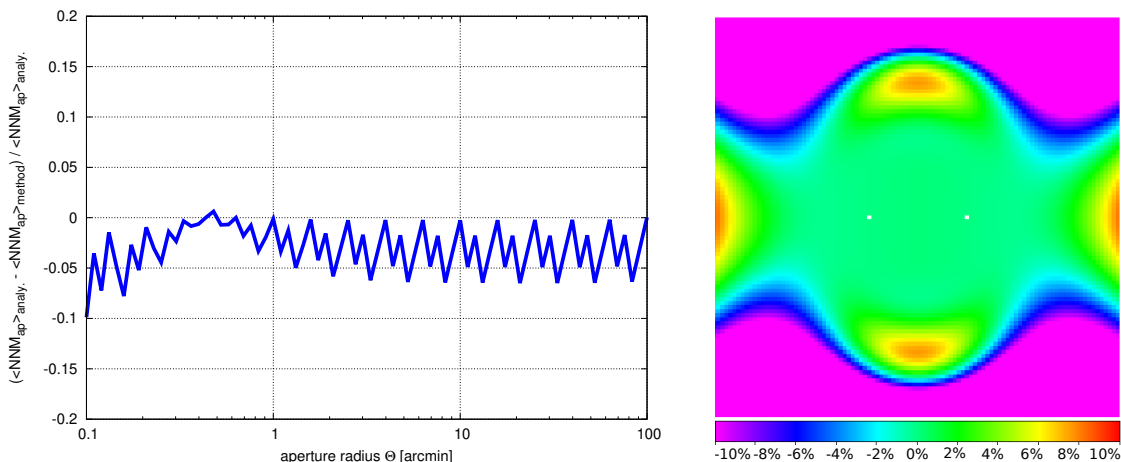


Figure 5.1: Relative deviation of the aperture statistics (*left*) and excess mass map (*right*) model predictions from the exact result using a Gaussian test bispectrum. The size of the map is $4 \times 4 \text{ arcmin}^2$.

of the aperture statistics predictions with the [CFHTLenS](#) measurements in Chapter 6. For the analysis of [G3L](#) presented in this chapter, however, these uncertainties are neglected. Furthermore, errors introduced by uncertainties in the values of adopted fixed model parameters, e.g. the cosmological parameters, are neglected throughout this work.

Additional errors arise from the numerical treatment of the problem. To make the numerically expensive computation of higher-order correlation functions feasible, numerical interpolations have to be used. The relative error from employing interpolations is in general negligible in this work, with one exception: the accuracy of the interpolation of the two-dimensional bispectrum $B^{\text{ggk}}(\ell_1, \ell_2; \ell_3)$ is only 50% for particular triangle configurations. This interpolation is, however, unavoidable as the computation of the bispectrum is expensive but needs to be repeated many times for the transformation to the excess mass maps and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ using a multipole expansion (see Sect. 3.2.2). The high inaccuracy of the interpolation is due to the bispectrum showing a cusp feature for squeezed triangles ($\ell_1 = \ell_2$, $\ell_3 = 0$ or $\varphi = \pi$) as also found by, e.g. [Jeong & Komatsu \(2009\)](#), which originates from the second-order coupling function $F_2(\mathbf{k}_1, \mathbf{k}_2)$ (Eq. 1.56) and the linear power spectrum from perturbation theory $P_{\text{lin}}(k)$ (Eq. 1.54) vanishing for $\mathbf{k}_3 = 0$. The accuracy of interpolating the bispectrum is therefore mainly determined by how well the cusp feature can be interpolated. By comparing the interpolated bispectrum to the original one, the accuracy of the interpolation is found to be 50% for $\varphi = \pi$. The resulting inaccuracy in the excess mass maps and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ is determined using the following two tests.

The general accuracy of the [G3L](#) predictions is tested using the Gaussian test bispectrum defined in Eq. (3.32), for which on one hand an excess mass map and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ can be calculated exactly (Eqs. 3.33 and 3.34). The Gaussian bispectrum is on the other hand also used as input to the transformation code to compute the [G3L](#) predictions. The comparison of the model predictions to the exact results in Fig. 5.1 shows that the relative error of employing the interpolation for the Gaussian two-dimensional bispectrum is below 7% in the range of 1 – 10 arcmin.

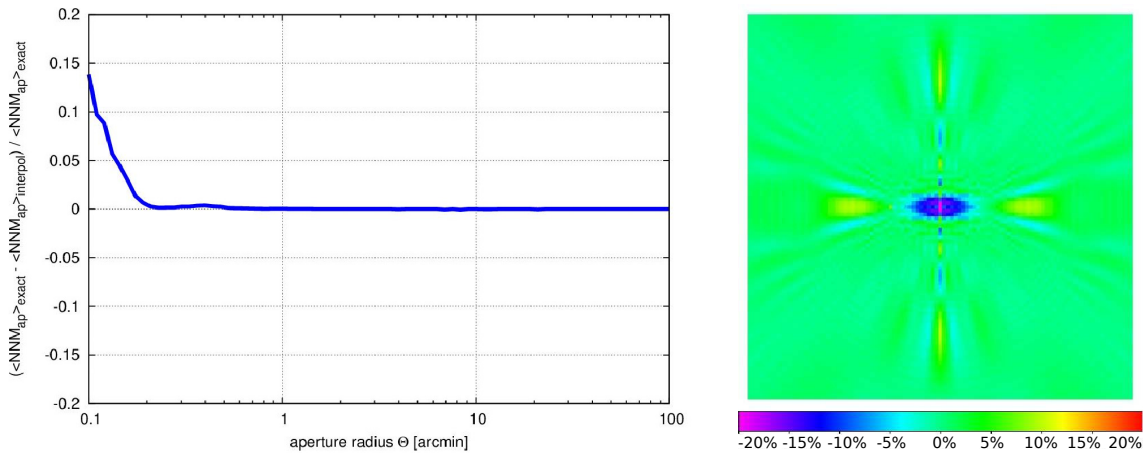


Figure 5.2: Relative error of the model prediction from the exact result for **G3L**. In the latter case, the bispectrum is interpolated over the whole range but close to the cusp feature ($\varphi = \pi$). The size of the map is 4×4 arcmin².

As the Gaussian test bispectrum has no cusp feature, a second test is performed with the real bispectrum where the bispectrum is interpolated for $\varphi \leq 0.998\pi$ and $\varphi \geq 1.002\pi$ (i.e. away from the cusp), and computed without interpolation in the range $0.998\pi < \varphi < 1.002\pi$ (i.e. around the cusp). This is computationally feasible for testing purposes, but not for predicting **G3L** generally. For $\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ the relative error reaches 15% for aperture scales around 0.1 arcmin (Fig. 5.2) but is negligible in the range relevant for this work (1 – 10 arcmin). In contrast, the relative error found for the excess mass map peaks right between the lenses with a value of 20%, but is typically below 5% outside the centre. This is relevant for future quantitative comparisons. The vertical artefacts in the map appear in most of the presented maps and are a consequence of the bispectrum cusp at $\varphi = \pi$.

5.2 Dependence of excess mass on galaxy properties

In this section it is investigated whether the halo model can describe the qualitative trends found for the observed excess mass maps (Simon et al. 2008, Simon+13, Simon+19) regarding lens-lens separation and lens galaxy-types. Moreover, the dependence of excess mass on stellar mass, luminosity, and redshift is studied. The predictions are presented for seven different lens-lens separations ranging from 0.5 arcmin to 3 arcmin for each of the 29 galaxy samples defined in Table 4.1. Figure 5.3 shows the predictions for the stellar mass samples, Fig. 5.4 for the luminosity samples, and Fig. 5.5 for the early- and late-type galaxy samples, all for low- z .

Overall, a clear dependence of the amplitude of the excess mass on the separation of the lens pair is found for all samples. In agreement with the observations, the amplitude of the signal increases with decreasing lens-lens separation. This behaviour relates to the increase of the three-point matter correlation function with decreasing physical scale arguing that galaxies largely trace the dark matter field (Simon+19).

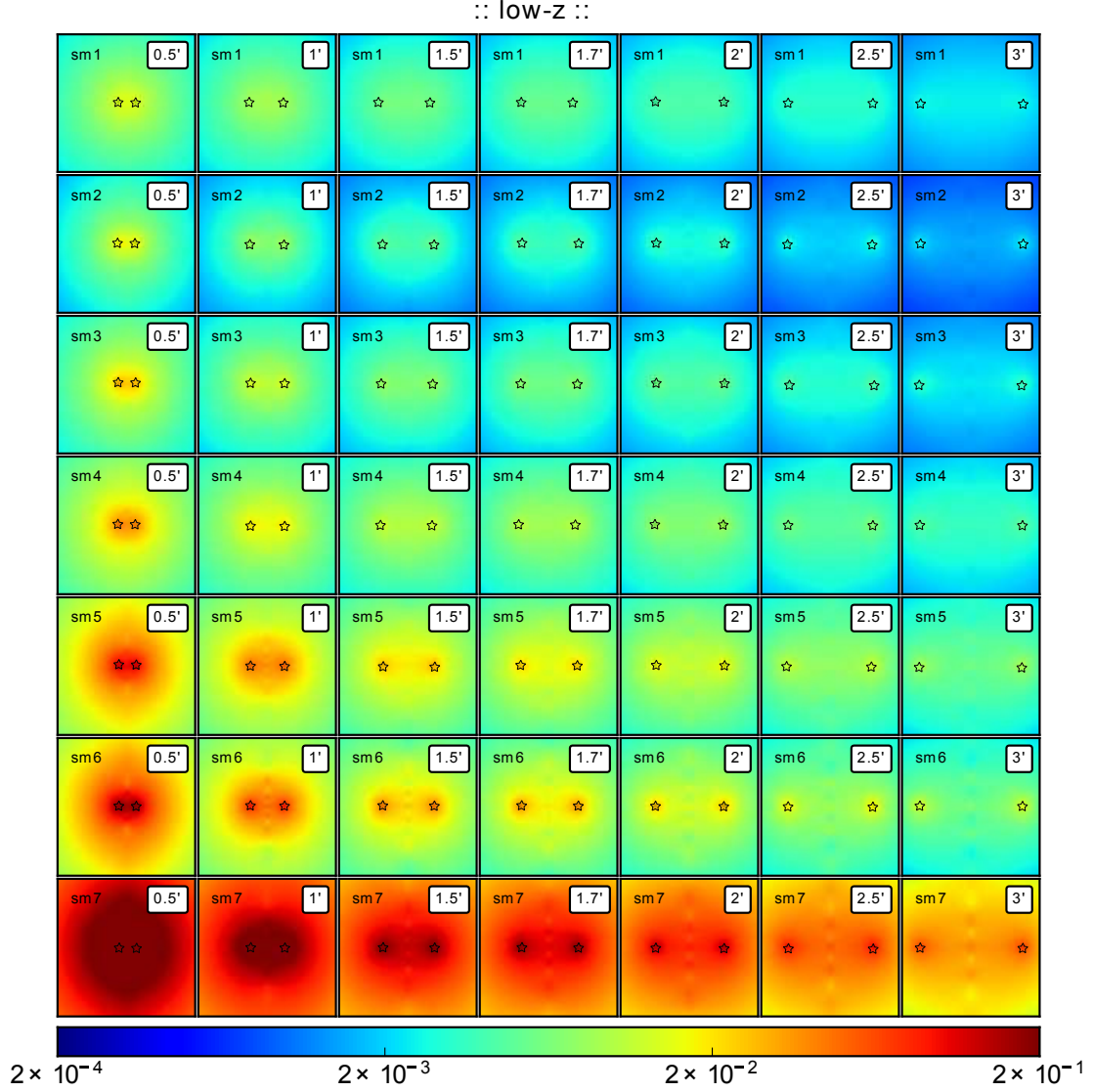


Figure 5.3: Excess mass map predictions for seven stellar mass samples (*top to bottom*) and lens-lens separations ranging from 0.5 arcmin to 3 arcmin (*left to right*). Shown are the low- z results. The lens positions are indicated by stars. The size of each map is 4×4 arcmin².

Regarding the dependence of excess mass on the stellar mass and the luminosity of the lens pairs, the excess mass amplitude generally increases with both, indicating that brighter and more massive galaxies inhabit more massive halos. In contrast to this general trend, the excess mass also increases for the lightest (sm1) and the faintest galaxies (L1, L2) considered. Notably, the HODs for the sm1 and L1 samples (Fig. 4.8), show a distinct behaviour compared to most other samples, i.e. a comparatively low (small α_{cen}) but extended plateau for small halo masses and a steep slope for high halo masses (large β). Also the fraction of centrals is comparatively high for these samples (Table 4.5).

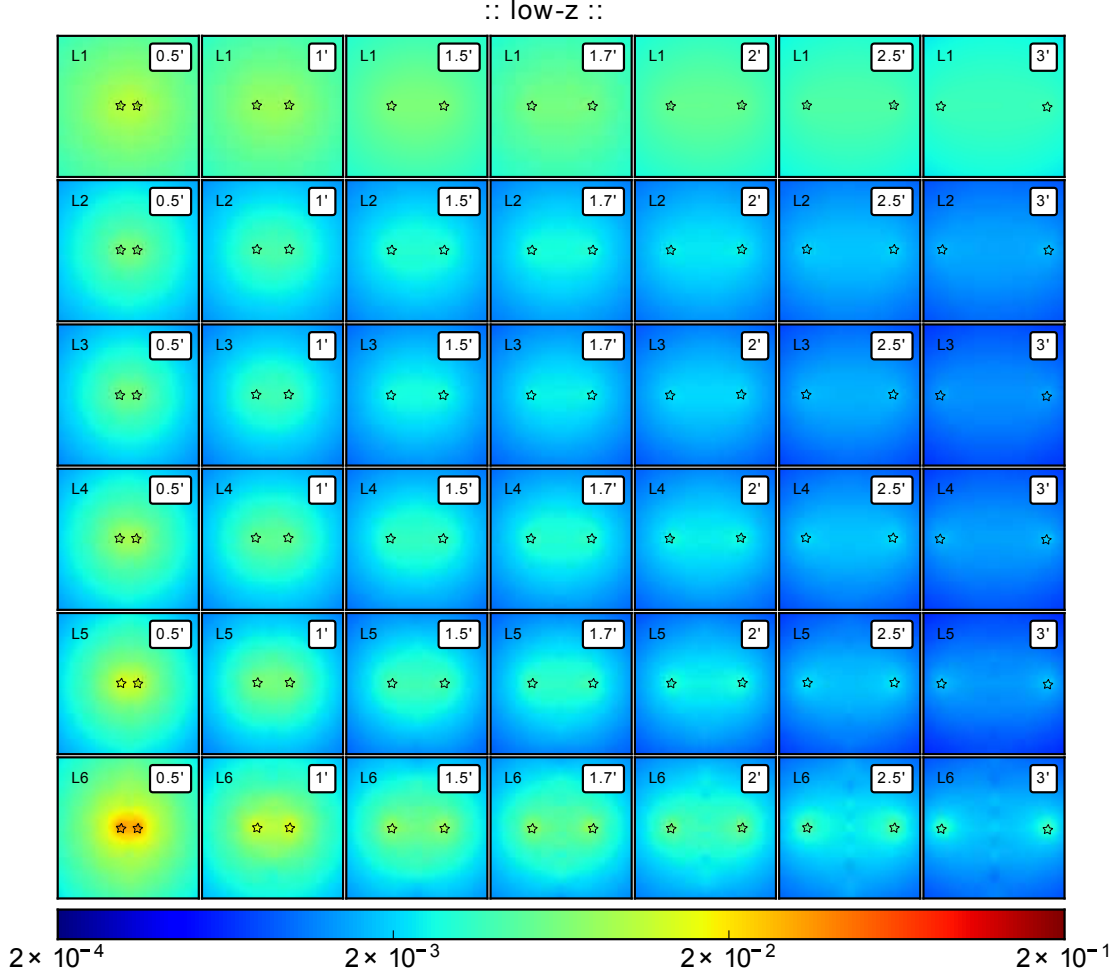


Figure 5.4: Excess mass map predictions for six luminosity samples (*top to bottom*) and various lens-lens separations. Shown are the low- z results. The lens positions are indicated by stars. The size of each map is 4×4 arcmin².

Whether the high amplitude of excess mass found for the sm1, L1, and L2 samples is physical can not be stated as particularly the L1 and L2 samples have a very different redshift distribution $p(z)$ compared to all other samples and therefore different lensing efficiencies (see Simon+13, Fig. 5 therein). Accounting for the different lensing efficiencies, the observed galaxy-galaxy-matter bispectrum still shows an increase for the L1 and L2 samples, however, not for the sm1 sample (Simon+13). Still, the results have to be treated with caution: given that the galaxy samples are flux-limited, the samples with the highest fraction of faint galaxies, i.e. L1 and L2, are most affected by incompleteness effects. Additionally, all samples below $\theta_{\text{ap}} \sim 1$ arcmin are affected by the transformation bias, which is evident for the L1 sample (Simon+13). Finally, also only a relatively poor model fit to G2L is given for the L1 sample (Table 4.4).

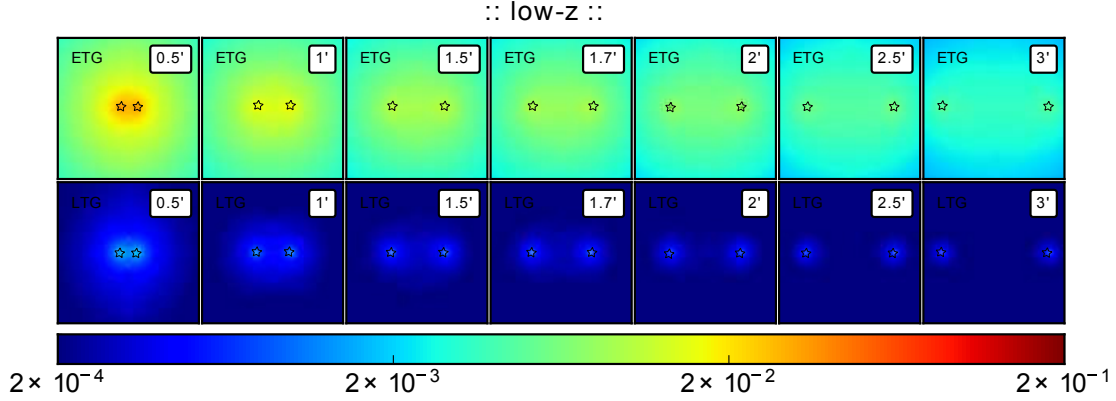


Figure 5.5: Excess mass map predictions for early-type (*top*) and late-type galaxies (*bottom*) and various lens-lens separations. Shown are the low- z results. The lens positions are indicated by stars. The size of each map is 4×4 arcmin².

With regard to its mean stellar mass and mean r -band luminosity, the early-type galaxy sample is basically a combined sample of the stellar mass samples sm3 to sm5 or of the luminosity samples L5 and L6 (Simon+13). It is therefore not surprising, that the model predictions for the early-type sample resemble the results for a typical stellar mass sample, i.e. sm3 or sm4. For example an envelope of excess mass, in the following referred to as ‘common’ halo, becomes increasingly apparent around the lens pair with decreasing lens-lens separation. This is also true for the late-type sample, although the amplitude of the late-type maps is more than a factor of ~ 10 lower than for the early-type maps, which is in agreement with the [CFHTLenS](#) observations (Simon+13) and [SAM](#) predictions ([Saghiha et al. 2012](#)). These results indicate that early-type galaxies inhabit a denser matter environment compared to late-types, which is in concordance with the observation that early-type galaxies constitute about 90% of satellites in galaxy clusters, whereas late-types account for 70% of field galaxies (e.g. [Dressler 1980](#); [Mandelbaum et al. 2006c](#)). This also indicates that late-types have a suppressed one-halo term.

The results for the high- z samples are presented in the appendix in Figs. [G.1 - G.3](#). For all galaxy samples, except for the high- z L2, L4 (high χ^2 for the [G2L](#) fitting; Table [4.4](#)) and sm6 sample, the amplitude decreases compared to the low- z sample, but the general trends found for the low- z sample are also valid in the high- z case. Again, this is to some extent because the lensing efficiency changes with redshift. Additionally, for the same angular lens-lens separation at low- and high- z , the actual physical separation increases with redshift, which results in a decrease of excess mass, too. In future work one can account for these effects, and possibly find for normalised excess mass maps what Simon+13 found for the normalised galaxy-galaxy-matter bispectrum for stellar masses below $\sim 10^{11} M_{\odot}$: a physical increase of the amplitude with redshift, which may be attributed to structure growth in a Λ CDM universe.

Figure [5.6](#) shows the $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ model predictions for all stellar mass, luminosity and galaxy-type samples for low- z and high- z . For the different stellar mass samples one can clearly see the trend already found in the excess mass maps: the amplitude of the

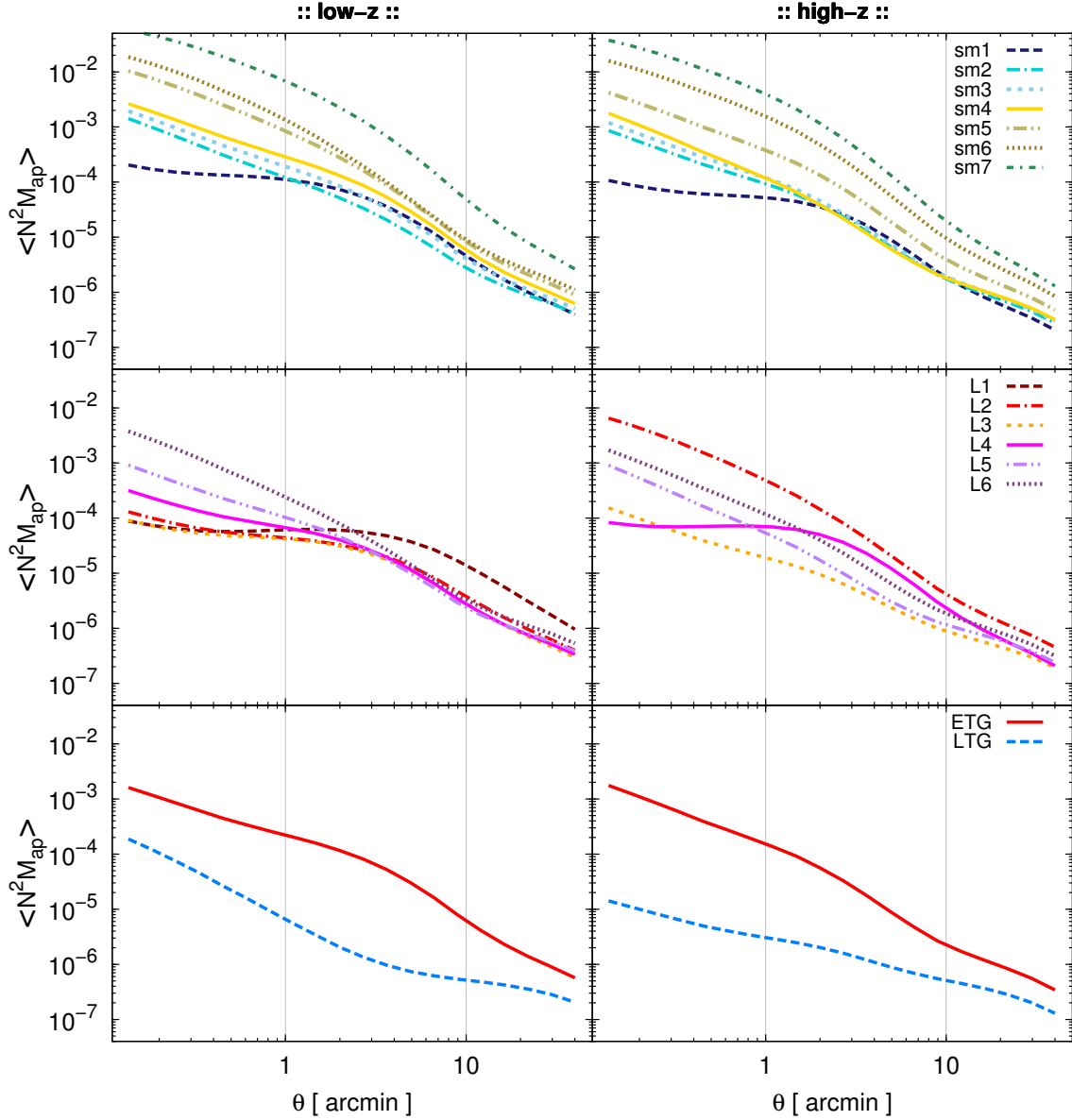


Figure 5.6: $\langle N^2 M_{\text{ap}} \rangle$ predictions for the stellar mass (*top*), luminosity (*middle*), and galaxy-type (*bottom*) samples as a function of aperture scale radius θ . The low- z results are shown on the left, the high- z results on the right. The vertical lines indicate the range considered in the quantitative analysis in Chapter 6.

signal increases with increasing stellar mass or decreasing aperture scale. As in the case of the excess mass maps, the sm1 sample shows a higher amplitude than the sm2 and sm3 samples, albeit only for aperture scales larger than approximately 1 arcmin. Towards smaller aperture scales the signal becomes relatively flat and is found to follow the general trend again. Regarding the luminosity samples, for the excess mass maps only a weak trend with luminosity was found. However, ignoring the L1 low- z as well as the L2 and

L4 high- z samples in Fig. 5.6, an increase of the amplitude with luminosity is visible, particularly for small aperture scales and high- z .

The $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ signal of the early-type sample dominates over that of the late-types on all scales and for both redshift samples. The deviation reaches up to two orders of magnitude, for low- z around $\theta \sim 3$ arcmin and for high- z on the smallest scales considered of $\theta \sim 0.1$ arcmin. In contrast to the late-type sample, the early-type sample is found to resemble the signal of a typical stellar mass or luminosity sample, i.e. sm3/sm4 and L5/L6, in terms of amplitude and shape.

The preceding discussion suggests that the stellar mass and luminosity samples show a similar qualitative behaviour. Therefore, the following discussion of the excess mass maps is restricted to the stellar mass samples sm1 to sm6 and the galaxy-type samples. Additionally, for conciseness only excess mass maps for a lens-lens separation of 1 arcmin are presented, which corresponds roughly to the separation explored in the CFHTLenS excess mass maps. With respect to the aperture statistics, the discussion will focus exemplarily on the early- and late-type samples, as the former resembles a typical stellar mass sample and the latter is found to behave uniquely compared to all other samples.

5.3 One-halo, two-halo, and three-halo term contributions

In this section the halo model is used to split the predicted excess mass and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ signals into contributions of the one-halo, the two-halo, and the three-halo term (see Eq. 2.32), which are due to correlations in a single halo, in two different halos, and in three different halos, respectively. The regimes of dominance of the individual halo model terms are determined, which is crucial to interpret the measurements.

Figure 5.7 shows the relative contributions of the one-, two- and three-halo term to the full excess mass signal for the stellar mass and galaxy-type samples for a lens-lens separation of 1 arcmin. For low- z , the three-halo term contributes at most 15% to the full signal for all samples other than the late-type sample. For all stellar mass samples and the early-type sample the one-halo term dominates with a relative contribution of about 60 – 90%. The excess mass associated with the one-halo term is distributed in a common halo around the lens pair. For the two-halo term a contribution of 15 – 40% is found. Whereas the relative contribution of the one-halo term is highest around the lens pair, the relative contribution of the two- and three-halo terms decrease towards the lens pair. This is expected as these terms typically become important on intermediate to large scales (Rödiger 2009), i.e. on scales displayed towards the corners of the maps and beyond.

A completely different behaviour is found for the low- z late-type sample. First, the two- and the three-halo terms clearly dominate, comprising 20 – 90% and 0 – 70% of the signal, respectively, compared to 5 – 30% of the one-halo term. Second, the relative contribution of the two-halo term peaks around the individual lens positions. Third, the gradient in the relative contribution is higher compared to the other samples. These findings support the assertion that the low amplitude of excess mass found around pairs of late-types (this work Fig. 5.5, Simon+13) is a consequence of late-types being typically field galaxies (Dressler 1980; Mandelbaum et al. 2006c). Namely, for field galaxies correlations stemming

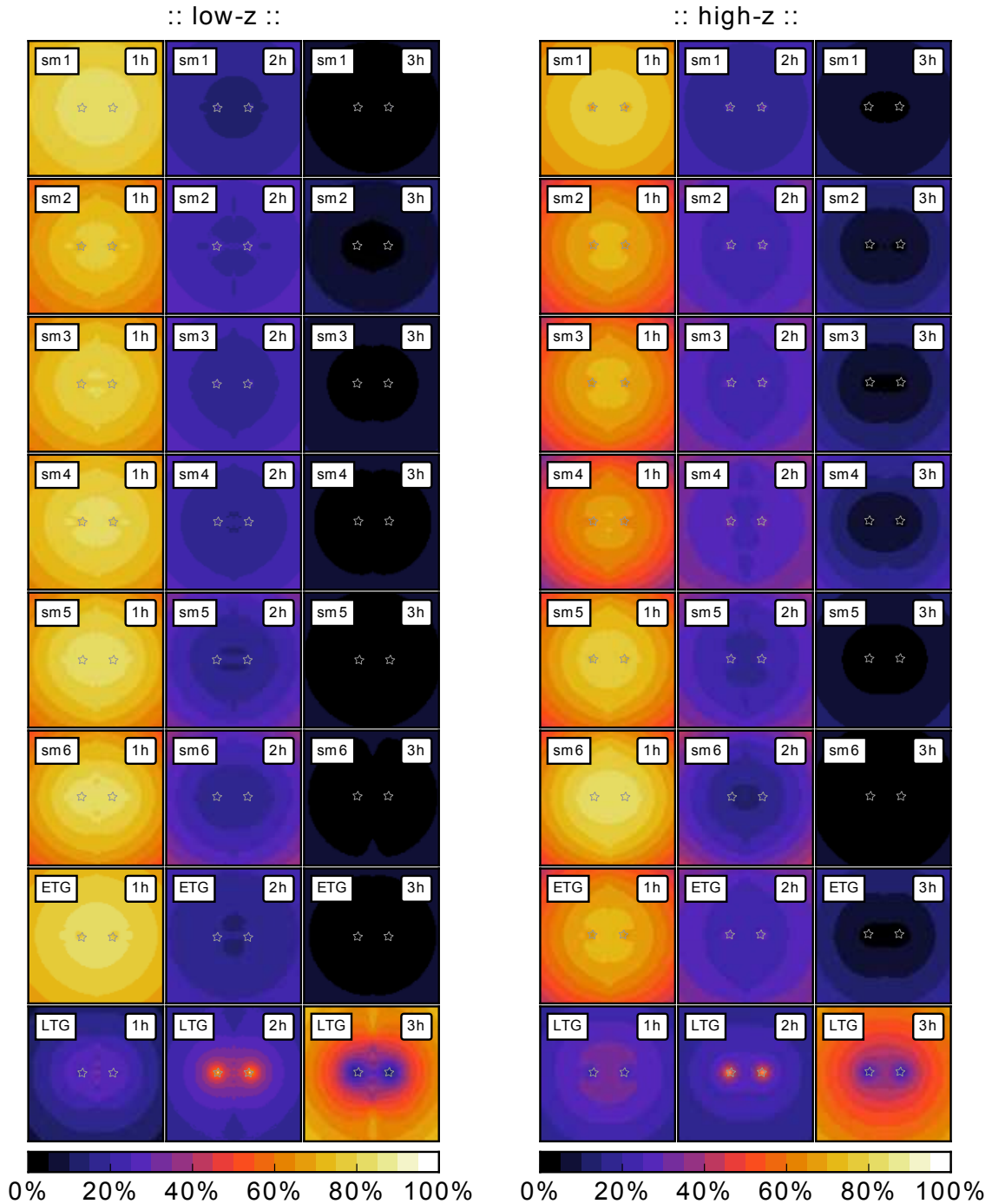


Figure 5.7: Relative contributions of the one-, two- and three-halo term (*left to right*) in the excess mass maps in percent for the stellar mass and galaxy-type samples (*top to bottom*). In the left plot the low- z results are shown, in the right plot the high- z results. All maps are for a lens-lens separation of 1 arcmin. The lens positions are indicated by stars. The size of each map is 4×4 arcmin².

from the same halo are suppressed, but correlations originating from different halos are enhanced.

Regarding redshift, a slight shift in contribution from the two-halo to the one-halo term is found for the late-type sample. In contrast, for the stellar mass and the early-type samples the relative contribution of the one-halo term decreases slightly with redshift, whereas the relative contributions of the two- and three-halo term increase. This could indicate that the excess mass of the common halo about pairs of galaxies from the sm1 to sm6, and the early-type samples increases with time, and vice versa for the late-type sample. Whether this finding reflects a physical trend can only be answered by comparing the maps at a fixed physical instead of angular separation.

The trends found for the early- and late-type galaxies become even clearer when examining the aperture statistics predictions (Fig. 5.8). First of all, the crossover between the dominance of the one-halo term and the dominance of the two/three-halo terms happens for the early-type low- z sample only at an aperture scale of approximately 10 arcmin, whereas for the late-type low- z sample the three-halo term starts to dominate already around 4 arcmin. Being typically cluster galaxies, early-type galaxies live in more extended halos, which allows to find pairs of early-types sharing the same halo even at comparably large lens-lens separations. Therefore, the one-halo term contributes to much larger radii than in the case of late-types. Second, the crossover point shifts for early-types towards smaller aperture scales with increasing redshift. The observed increase of the influence of the one-halo term towards larger scales with time is partially due to objects getting smaller with redshift. Partially, it may also reflect the bottom-up scenario of structure formation in a Λ CDM universe, where large structures are assumed to form by mergers of smaller structures. For the late-type sample the crossover point between one- and three-halo term stays roughly at the same aperture radius with changing redshift, but the two-halo term dominates over the one-halo term only at larger aperture scales for high- z . Third, the total amplitude increases with time: for the early-type sample particularly on scales larger than 1 arcmin and for the late-type sample on scales smaller than 4 arcmin. This is to some extent due to the effect of the samples having different redshift distributions. Additionally, this may reflect the accretion of mass with time in a cold dark matter model. Finally, Fig. 5.8 shows that, although the two-halo term adds a significant contribution to the signal particularly in the range tested by CFHTLenS, it is actually never the dominant term among the three halo terms.

As a final note, the computation of the full G3L signal is computationally very expensive. This is also why in this work it is refrained from fitting the model simultaneously to G2L and G3L to constrain the HOD. A possible approximation of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ by the sum of the one-halo term and the large-scale limit of the three-halo term is discussed in Rödiger (2009). The results presented here are in agreement with Rödiger (2009): the one-halo term is generally a good approximation up to scales of 2 – 3 arcmin. However, care must be taking for the late-type sample. Whereas the signal for the early-type and stellar mass samples is with a maximal deviation of 10% well approximated by the one-halo term up to scales of 3 arcmin, for late-types deviations of 10% occur already for 0.1 arcmin in the high- z case due to the suppression of the one-halo term.

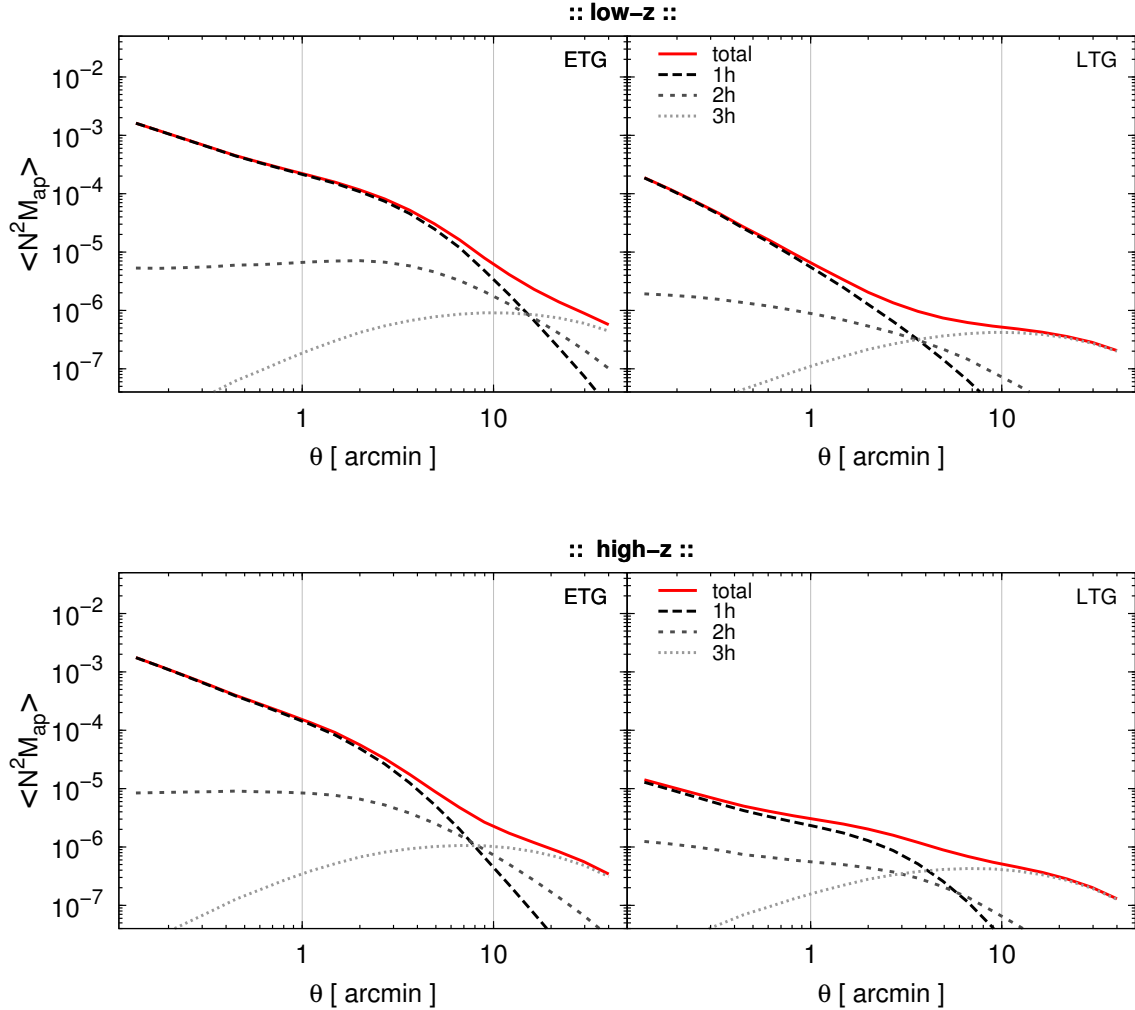


Figure 5.8: Contributions of the one-, two- and three-halo term to $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ for the early-type (*left*) and the late-type (*right*) samples as a function of aperture scale radius θ . Shown are the low- z (*top*), and the high- z results (*bottom*). The vertical lines indicate the range considered in the analysis in Chapter 6.

5.4 Sensitivity of G3L with respect to HOD parameters

In this section the sensitivity of excess mass and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ on the employed parametrisation of the HOD (Sect. 2.2) is studied by varying the five model parameters $\sigma_{\log(M_h)}$, β , M_1 , M_{th} , and α_{cen} individually by $\pm 20\%$. As reference the best-fit model to the CFHTLenS G2L for each sample of lenses is used (Table 4.4). It is investigated which HOD parameters can potentially be constrained, e.g. by fitting the model directly to G3L measurements, and which galaxy samples serve this purpose best. The focus is particularly on the excess mass map representation of G3L, as the maps offer the opportunity to study whether map

features change in an identifiable manner when varying HOD parameters.

The resulting maps are shown in the appendix in Figs. G.4 - G.7. Figures 5.9 and 5.10 show the relative change from the best-fit reference map in percent for the stellar mass and galaxy-type samples for low- and high- z . As in the case of G2L, the parameter $\sigma_{\log(M_h)}$, which quantifies the dispersion of the mean number of central galaxies at a fixed halo mass, has little influence on the signal. Specifically, when varying $\sigma_{\log(M_h)}$ by $\pm 20\%$ around its best-fit value, a change of the respective excess mass maps from the reference map of at most 10% is found for either redshift sample. Since the uncertainty in determining $\sigma_{\log(M_h)}$ from the G2L model fits is larger than 50% (Table 4.4), varying $\sigma_{\log(M_h)}$ by $\pm 20\%$ is hardly meaningful, which should be taken into account in future analyses.

Variations in the parameter β lead to the strongest changes in the excess mass maps where deviations of up to at least $\pm 50\%$ are found. This is not surprising as β is the parameter that controls the satellite power-law slope, and small changes in β lead to large changes in the mean number of satellites $\langle N_{\text{sat}}(M_h) \rangle$. For all samples but the low- z late-type sample, a higher value for β increases the excess mass, whereas a lower value for β decreases the excess mass; the effect is more pronounced for the smallest stellar mass samples and less pronounced for the highest stellar mass samples. The effect of varying β can be explained from the definition of the galaxy-galaxy-matter bispectrum $B^{gg\delta}$ (Eq. 2.32): a higher value of β gives more weight to massive halos, and consequently the one-halo term contributes to larger scales, compensating the exponential decrease of the halo mass function for massive halos. The increase of $B^{gg\delta}$ translates into the increase of \mathcal{G} (Eq. 3.28). The effect is strongest for the samples containing galaxies of the lowest stellar masses because the best-fit reference value for β is the highest for these samples (Table 4.4). Interestingly, for the low- z late-type galaxy sample the change is less than 15% and opposite to what has been found for all other samples: an increase of β leads to a lower amplitude of the excess mass, a decrease to a higher amplitude. This point is picked up later when discussing the sensitivity analysis for the aperture statistics.

For the remaining three HOD parameters M_1 , M_{th} , and α_{cen} all galaxy samples show basically the same trends, although the size of the relative change depends on the sample. A lower satellite threshold mass M_1 enhances the amplitude of $\langle N_{\text{sat}}(M_h) \rangle$, and thereby the amplitude of the excess mass by 10% for, e.g. the low- z sm6 and late-type sample, and by more than 50% for, e.g. the high- z sm1 sample. However, a reduced threshold mass for central galaxies M_{th} results in a decrease of the amplitude of $\langle N_{\text{sat}}(M_h) \rangle$, and thus in a decrease of the excess mass by 15% (late-type sample) to 50% (low- z sm6 and early-type sample, and high- z sm1 sample). Reducing the parameter α_{cen} increases the number of halos with only satellites. This is similar to lowering the mass threshold M_1 , populating small halos with one satellite only, and accordingly the same trends are found: the excess mass increases by 10% to 40% with reduced α_{cen} .

In summary, although for four out of five HOD parameters variations cause significant changes in the maps, the changes do not have a clear characteristic of a specific parameter. Moreover, for simplicity, correlations between the HOD parameters have not been taken into account in this analysis, but realistically every change in the signal is a superposition of multiple parameters changing at the same time, which has to be taken into account in future analyses.

Figure 5.9: Relative deviation of excess mass maps from best-fit reference maps for the case that one HOD parameter is varied by $\pm 20\%$ while keeping the other four fixed at their best-fit value. Shown are the results for the different stellar mass and galaxy-type samples (*top to bottom*) for a lens-lens separation of 1 arcmin and low- z . The size of each map is 4×4 arcmin².

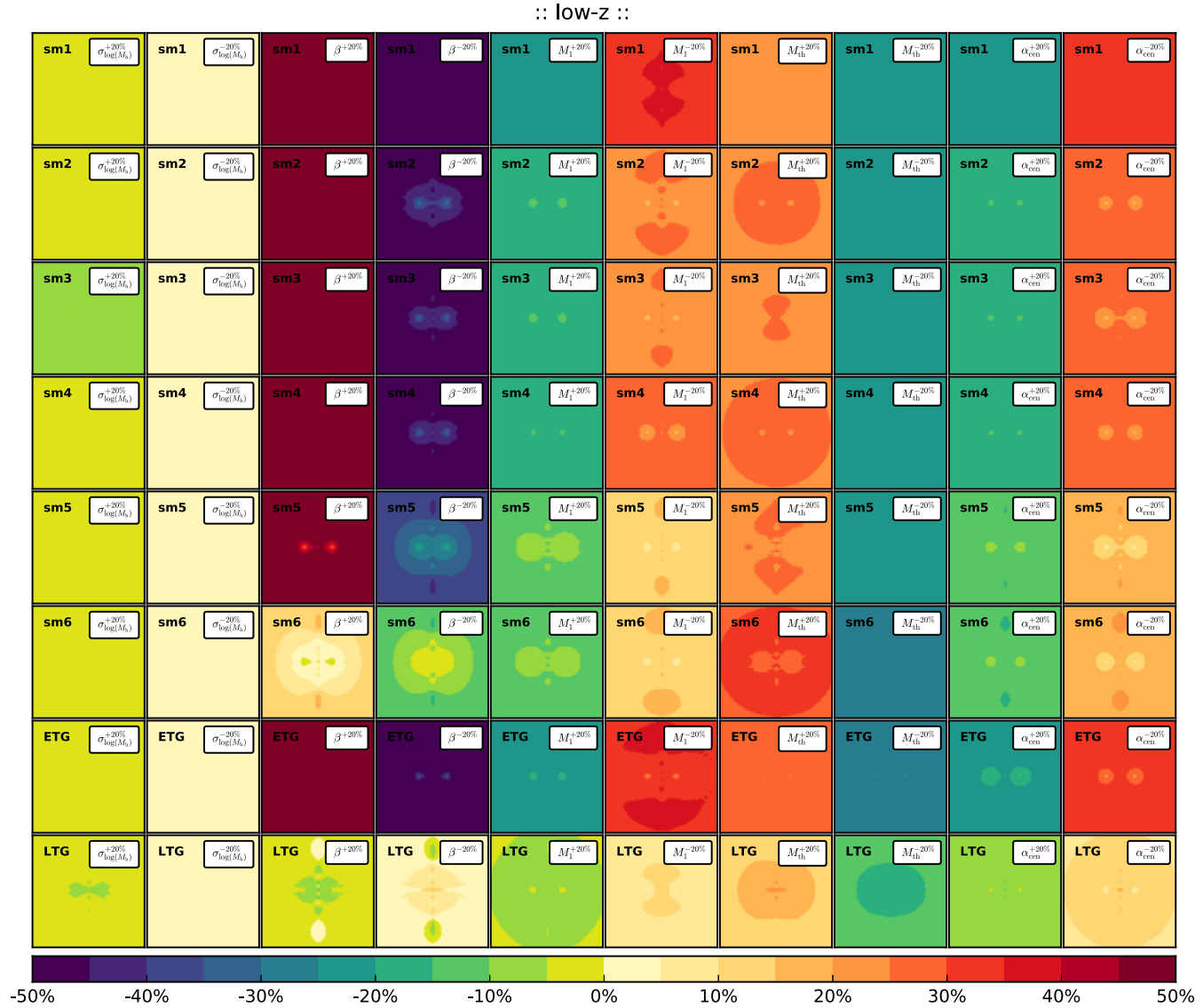
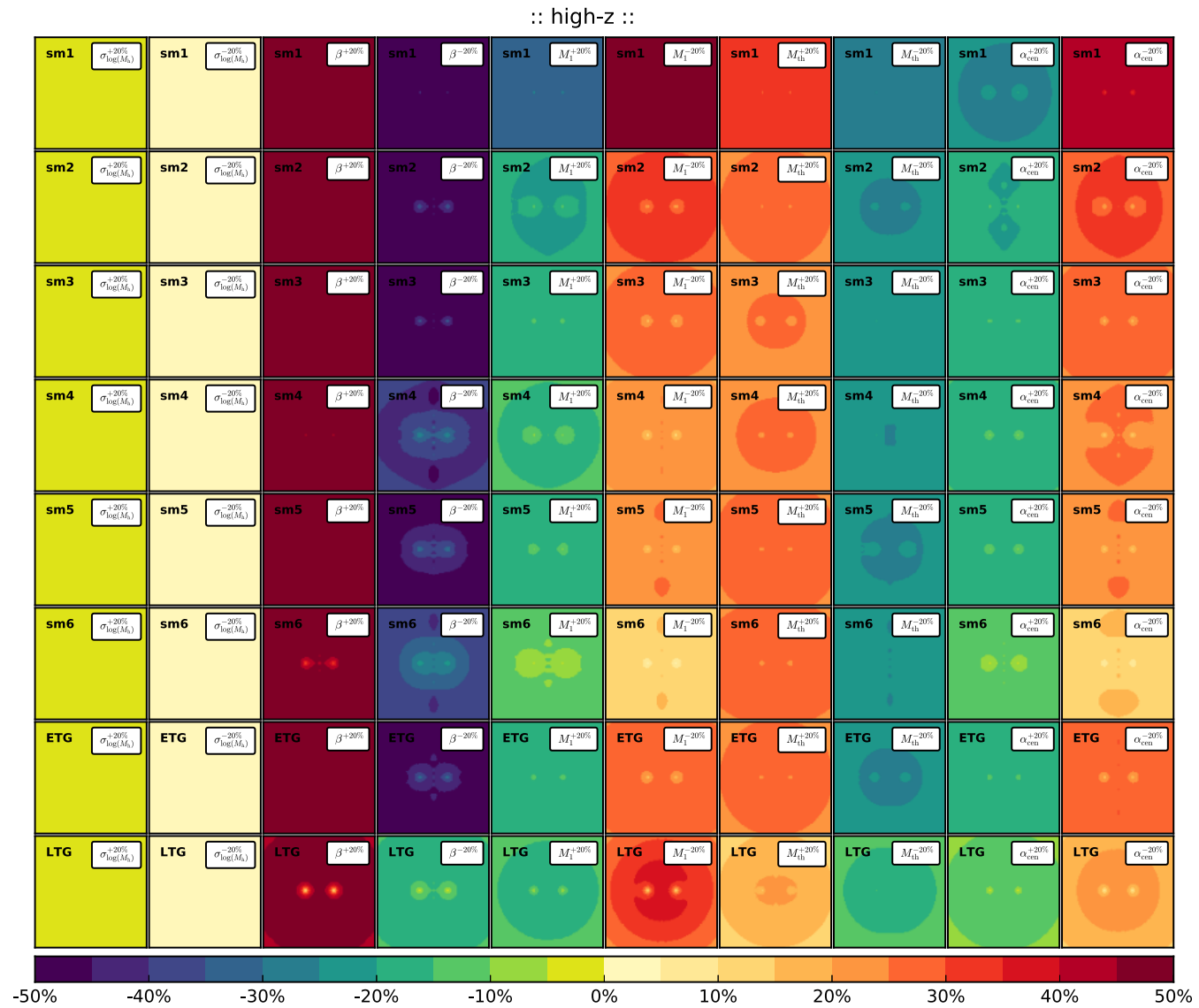


Figure 5.10: As Fig. 5.9 but for high- z .



Now the sensitivity of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ with respect to the HOD parameters is considered. The results for the stellar mass samples are presented in Fig. 5.11 for low- z and in Fig. G.8 for high- z . The results for the galaxy-type samples are shown in Fig. 5.12 for low- and high- z . The large panels show the aperture statistics prediction for the +20% and -20% variation of the individual model parameters, and the small panels the relative deviation of the predicted signal from the best-fit reference. These plots reveal several aspects that are not, or not easily, detectable in the excess mass results.

First of all, with the exception of the low- z late-type sample, the relative deviation from the best-fit reference model for all parameters is greatest in the angular range observed by CFHTLenS. This indicates that these measurements are well-suited to constrain the HOD parameters. Nevertheless, depending on the parameter and the galaxy sample, also scales smaller and larger than 1 – 10 arcmin are promising to determine the HOD parameters, e.g. changes of around 20% are apparent for M_{th} on scales smaller than 1 arcmin and for β up to scales of 30 arcmin.

Second, the changes for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ converge towards the reference model for small and large aperture radii. This depends in detail on the considered HOD parameter and sample, but the behaviour is generally expected for the following reasons. On small scales the one-halo term is dominated by correlations of centrals, and thus depends only weakly on $\langle N_{\text{sat}}(M_{\text{h}}) \rangle$. Thus the curves for β , M_1 and α_{cen} , affecting satellites, converge. On large scales the three-halo term is dominant, which depends only on the first-order moment of the HOD and all curves are converging.

Third, the relative deviation depends on stellar mass. For β , M_1 and α_{cen} the relative deviation clearly decreases from more than 30% to less than 20% with higher stellar mass. For β this is due to the trend the best-fit reference parameter show with changing stellar mass (Table 4.4). As a result the low stellar mass samples are better suited to constrain these three HOD parameters than the high stellar mass samples.

Fourth, for the parameter β there is a crossover at a certain aperture scale radius: whether the aperture signal is enhanced or lowered by changing β depends on the considered scale. This crossover scale shifts with increasing stellar mass towards larger aperture radii. The behaviour is most extreme for the late-type low- z sample, where the crossover happens at the upper end of the range tested with CFHTLenS, and is less a crossover and more a convergence. Note that for high- z the behaviour of the late-type sample resembles the general one of the other samples. This sheds light on why the excess mass prediction for the late-type low- z sample shows an opposite behaviour compared to the predictions for all other samples when varying β by $\pm 20\%$.

Last, the distinct behaviour found for the late-type sample makes the galaxy-type samples particularly suited for future attempts to constrain the HOD parameters by direct model fits to G3L. On one hand, Fig. 5.12 suggests that the sensitivity regarding changes of the model parameters is comparable to the one of a low stellar mass sample. On the other hand, the maximal deviation from the reference model happens for distinct aperture radii, for the early-types between 1 and 10 arcmin, for the late-types around 1 arcmin. Furthermore, the behaviour for the late-type sample changes with redshift regarding the shape of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ and the sensitivity to the model parameters.

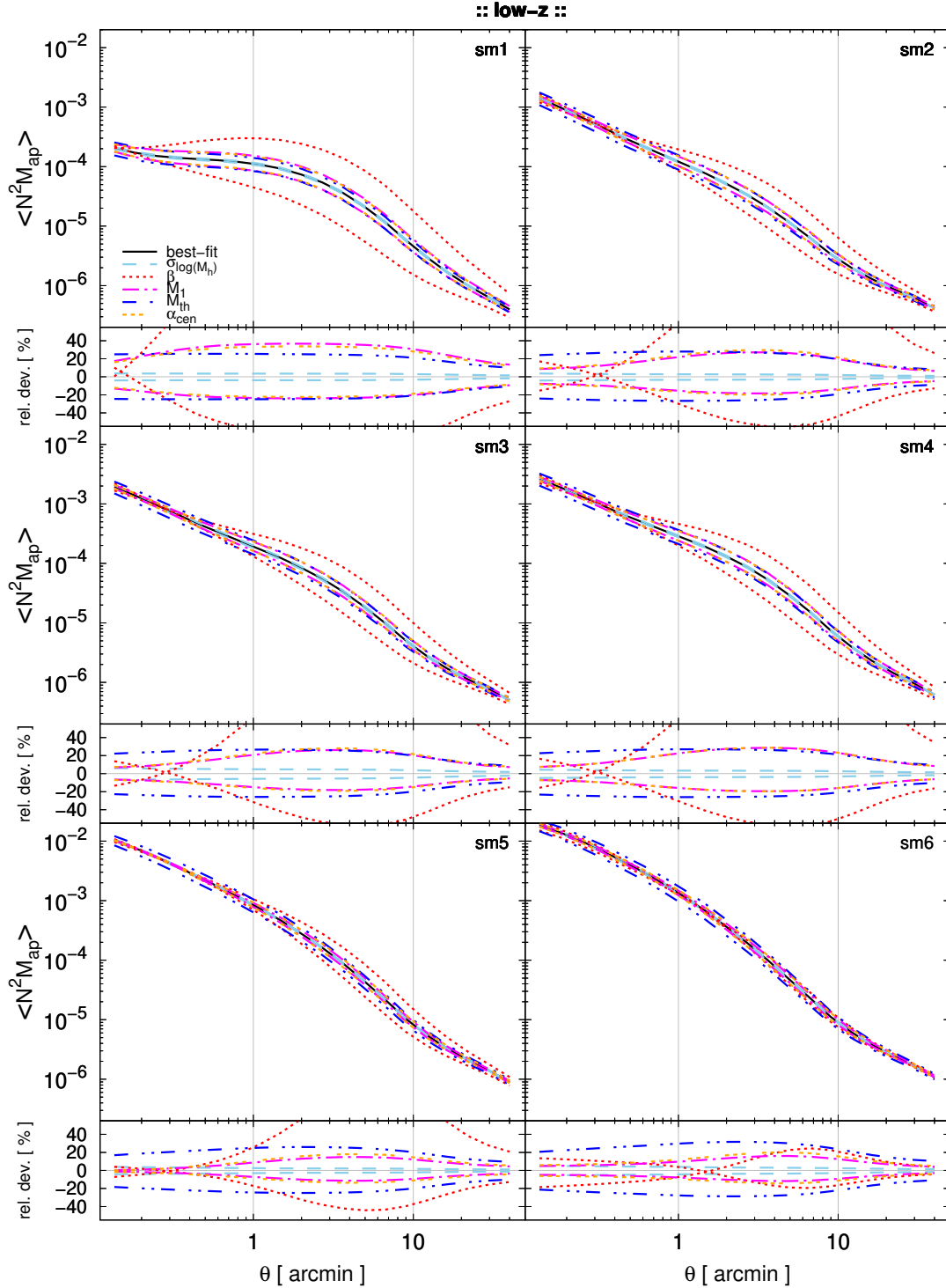


Figure 5.11: Sensitivity of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ on variations in the five HOD parameters by $\pm 20\%$ with respect to the best-fit reference model. The large panels show the predictions for the six stellar mass samples sm1 to sm6 as a function of aperture scale radius, the small panels the relative deviation from the reference model in percent. Shown are the predictions for low- z .

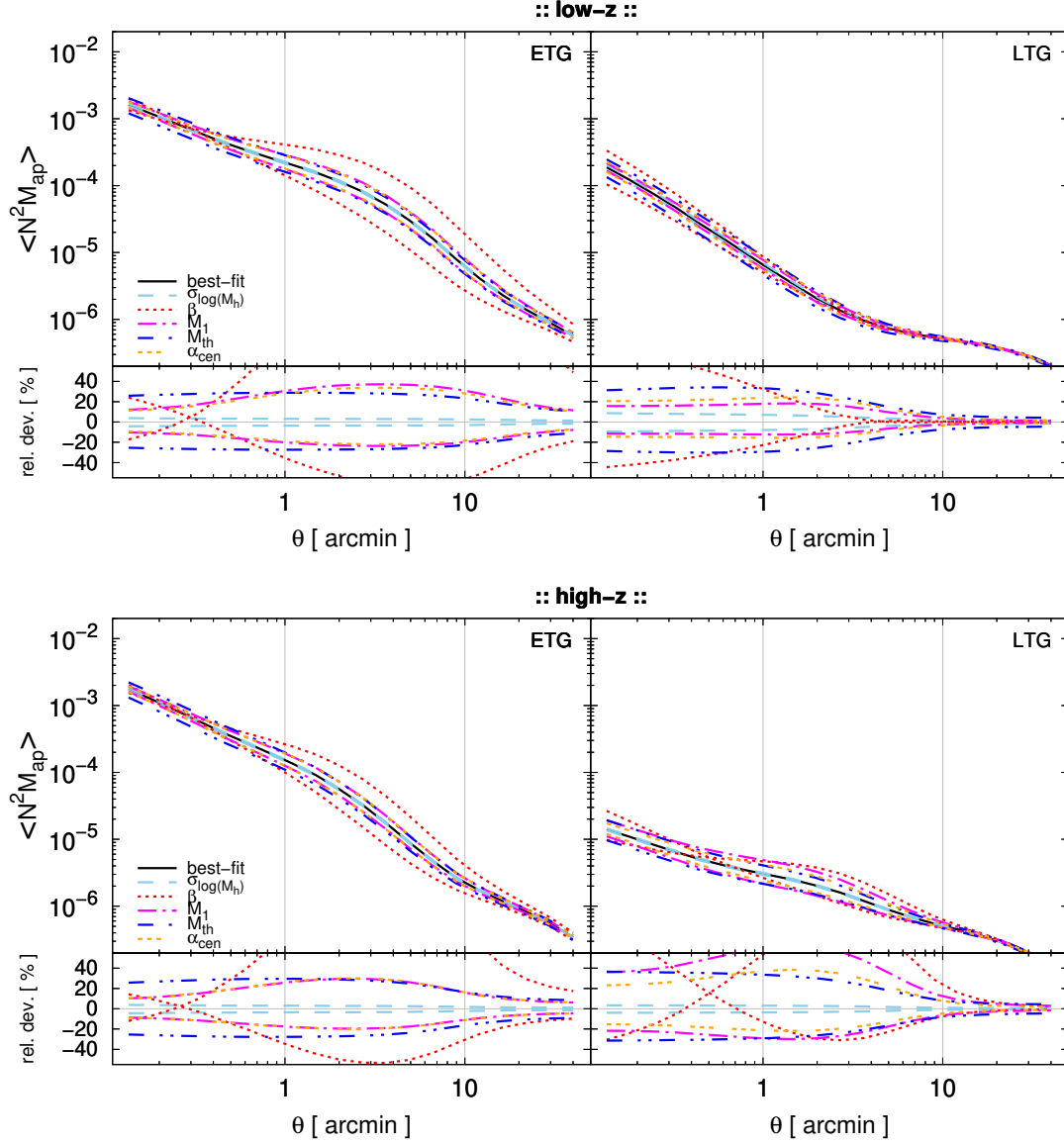


Figure 5.12: As Fig. 5.11 but for the early- and late-type samples for low- (*top*) and high- z (*bottom*).

5.5 Excess mass maps – model predictions versus observations

Observations of CFHTLenS indicate a vertical bulge-like distribution of excess mass between pairs of galaxies in the early-type or combined sm1-sm6 sample, which can not be reproduced with the recent SAM by Henriques+15 (Simon+13, Simon+19). In this section the halo model is used to specifically predict excess mass maps based on the best-fits to G2L from CFHTLenS for the combined sm1-sm6 sample, to see whether the halo model predicts the observed structures.

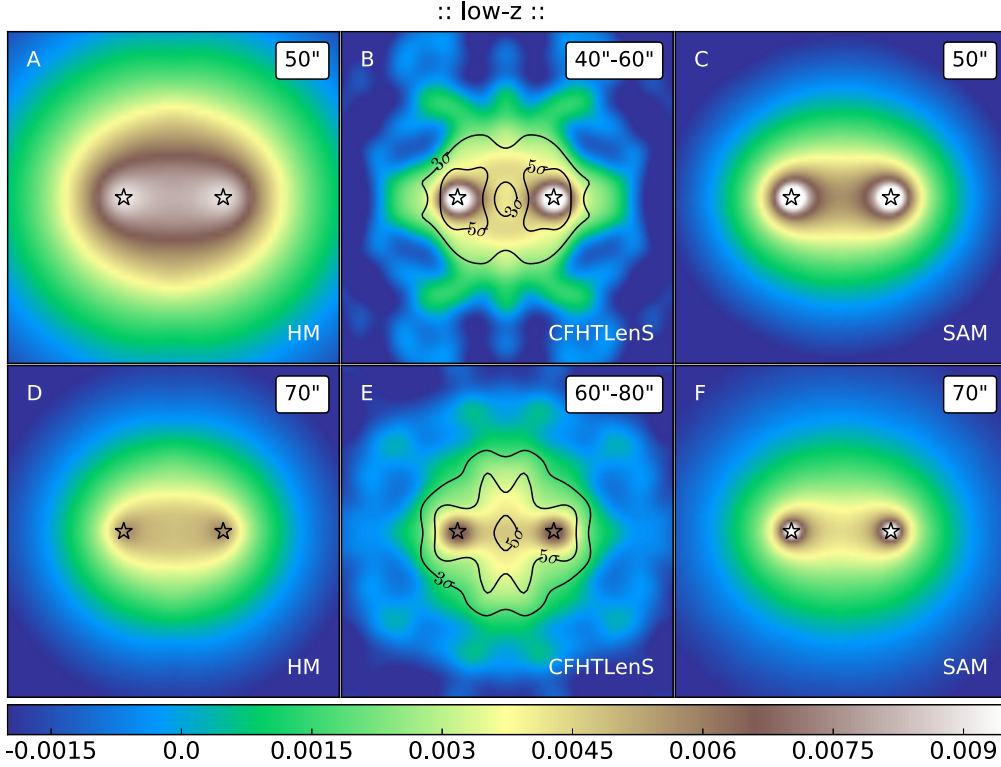


Figure 5.13: Comparison of halo model predictions for the excess mass (panels A, D) to the [CFHTLenS](#) measurements (B, E) and to predictions from a [SAM](#) (C, F). The results are presented for the combined low- z sm1-sm6 sample, considering a ‘close- θ ’ (*top*) and a ‘wide- θ ’ (*bottom*) lens-lens separation. The size of each ‘close- θ ’ map is $2.8 \times 2.8 \text{ arcmin}^2$, the size of each ‘wide- θ ’ map is $3.8 \times 3.8 \text{ arcmin}^2$. The contours give the 3σ - and 5σ -significance of the measurement. The lens positions are indicated by stars.

The [CFHTLenS](#) maps published in Simon+19 probe the excess mass about pairs of lenses with a ‘close- θ ’ angular separation of 40 – 60 arcsec with an average of ~ 50 arcsec, and a ‘wide- θ ’ separation of 60 – 80 arcsec with an average of ~ 70 arcsec. Therefore, the halo model maps are computed for angular lens-lens separations of 50 and 70 arcsec, similar to the [SAM](#) ones. All maps are smoothed and normalised, i.e. by applying a Gaussian smoothing with a kernel size of $\sigma_{\text{rms}} \approx 4 \text{ arcsec}$ and by offsetting the map such that the mean lensing convergence κ is zero over the map area.

The results are shown in Figs. 5.13 and 5.14 for low- and high- z , respectively. In agreement with the [CFHTLenS](#) measurements and [SAM](#) predictions the halo model predicts the excess mass to peak around the individual lens positions, and to decrease from the centre towards the outer regions of the map by about $\sim 4 \times 10^{-3}$, corresponding to a drop in the excess surface mass density of $\sim 17 h M_{\odot} \text{pc}^{-2}$ (Simon+19). In contrast to the two other methods, the halo model predicts, first, a slightly lower maximal amplitude at the positions of the lenses and, second, for the close- θ case the excess mass to extend to larger distances from the lens pair.

The vertical bulge of excess mass found for the [CFHTLenS](#) maps (Fig. 5.13 panel

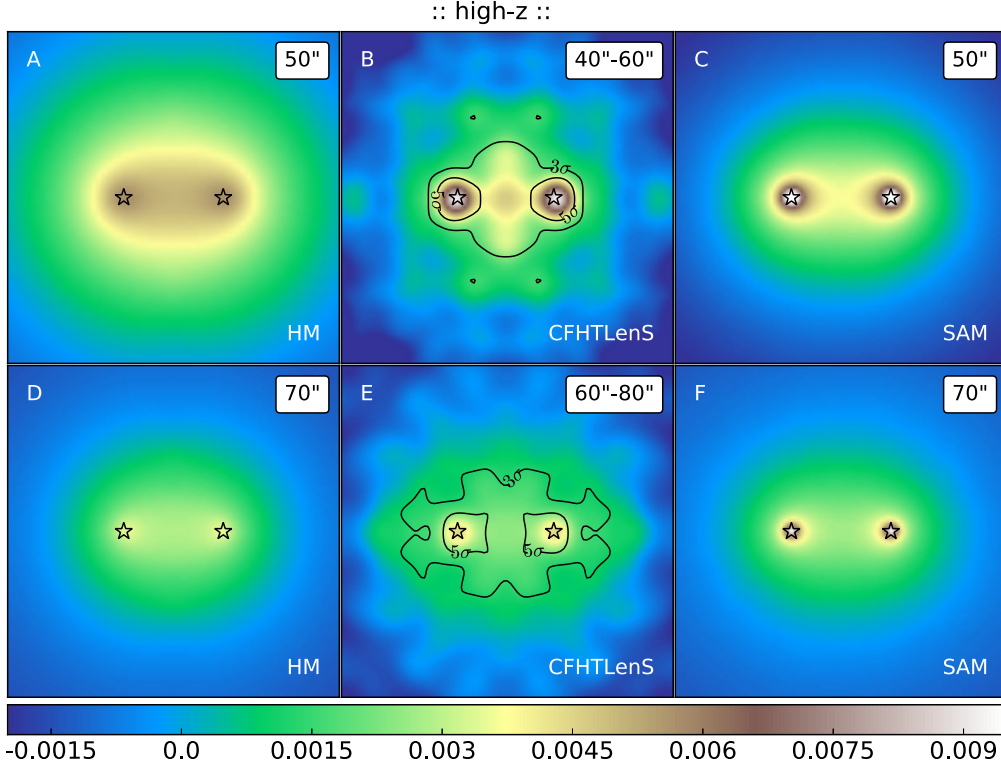


Figure 5.14: As Fig. 5.13, but for high- z .

E, Fig. 5.14 panel B) is not reproduced by the halo model. This is not surprising, as the halo model predictions are expected to resemble the SAM predictions. Simon+19 discuss the possibility that the vertical structure found in the observations originates from a misalignment between the distribution of the intra-cluster medium and of the dark matter, or between the orientation of lens pairs and of the parent halos. Both hypotheses could be tested using the halo model by including misaligned mass distributions and non-spherical halos.

Future surveys such as the Kilo-Degree (KIDS)¹ and the Dark Energy Survey (DES)² will allow the measurement of excess mass maps with a higher signal-to-noise, possibly even for distinct stellar mass samples. In that case, the halo model predictions presented in Sect. 5.2 will help interpret the measurements.

5.6 Conclusions

In this chapter first-time predictions of excess mass maps using the halo model are presented. Trends of excess mass and $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ with angular separation, galaxy properties, and redshift are explored by predicting G3L for 29 galaxy samples based on the model

¹<http://kids.strw.leidenuniv.nl/index.php>

²<https://www.darkenergysurvey.org>

fits to the **CFHTLenS G2L**. The halo model reproduces the trends found in observations (Simon et al. 2008, Simon+13, Simon+19), i.e. the amplitude of excess mass increases with decreasing lens-lens separation and the signal around pairs of early-type galaxies is more than a factor of ten higher than around late-types. Furthermore, a dependence of **G3L** on stellar mass and luminosity is found: the amplitude is higher for pairs of more massive and luminous galaxies, indicating that brighter and more massive galaxies inhabit more massive halos. The results indicate that this may also be the case for the least massive and luminous galaxies considered (sm1, L1 and L2), but to establish this finding one would need to account for the effect of different redshift distributions causing different lensing efficiencies. The same applies to the increase of **G3L** with time predicted by the halo model, which may reflect the accretion of matter expected in a Λ CDM universe.

Concerning the contributions of the one-, two- and three-halo terms to the full signal, the model predicts that the one-halo term dominates from the smallest aperture scales up to 10 arcmin, which covers the range tested with **CFHTLenS**. This is the case for all samples other than the late-type sample, for which the two- and three-halo terms dominate already starting from ~ 4 arcmin. The accuracy of approximating $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ by the one-halo term depends on the considered galaxy sample. The deviation is less than 10% up to scales of 3 arcmin for all but the late-type sample. The excess mass map representation reveals that for all samples but the late-type sample the dominance of the one-halo term at a lens-lens separation of 1 arcmin results in the distribution of excess mass in a common halo around the lens pair. For the late-type sample the dominance of the two-halo term results in two peaks of excess mass around the individual lens positions. The one-halo term is suppressed for late-types, which is a consequence of late-types being typically field galaxies, and inhabiting less massive halos for which the probability to find two lenses in one halo is comparatively low.

The sensitivity of **G3L** relative to $\pm 20\%$ variations of the **HOD** parameters is explored. For the satellite parameter β deviations of more than 50% are predicted. For the threshold masses M_1 and M_{th} , and the parameter α_{cen} changes of around 10 – 50% occur depending on the galaxy sample. The changes decrease with stellar mass, which suggests that the low stellar mass samples are better suited to constrain the **HOD**. Furthermore, the change is largest in the angular range tested by **CFHTLenS**, except for the late-type sample where the change is largest on scales smaller than 1 arcmin. This could help break parameter degeneracies if future measurements of **G3L** become reliable on scales smaller than 1 arcmin. Recent work by Laila Linke (priv. comm.) suggests that this will soon be possible. Unfortunately, no identifiable changes in the map features when varying the individual **HOD** parameters are found.

Excess mass map predictions for a combined sm1-sm6 sample are compared to their observational counterpart to see whether the halo model can explain the map features found with **CFHTLenS** that the recent **SAM** by Henriques+15 cannot. A generally good agreement is found: the excess mass peaks around the individual lens positions, and decreases from the centre towards the outer regions of the map by a value of $\sim 4 \times 10^{-3}$, corresponding to a drop in the excess surface mass density of $\sim 17 h M_{\odot} \text{pc}^{-2}$. Like the **SAM**, the halo model can not reproduce the vertical bulge found in the observations. However, the halo model could help test the genuineness of this feature by predicting **G3L**

for the scenarios Simon+19 name as possible physical sources: a misalignment between the distribution of the intra-cluster medium and of the dark matter, or between the orientation of lens pairs and of the parent halos.

As a final note, the most interesting results in terms of studying galaxy evolution can be expected from the galaxy-type samples. For all questions pursued in this chapter, the signal for the late-type sample shows a unique behaviour compared to the other samples at low redshifts, which becomes more similar to the ones of the other samples when studying high redshifts.

Chapter 6

Testing G3L halo model predictions against CFHTLenS observations

The central question to be answered by this doctoral thesis is whether a ‘standard’ halo model can consistently describe second- and third-order galaxy-dark matter correlations as probed by G2L and G3L. To answer this question, the best-fit models to the CFHTLenS G2L (Chapter 4) are used to predict G3L in terms of the aperture statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ for 28 galaxy samples of stellar mass, luminosity, and galaxy-type at two redshifts. The halo model predictions for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are then confronted with their observational counterpart from CFHTLenS (Simon+13). The results of this quantitative comparison are presented and discussed in this chapter. This work constitutes the next level test for the halo model and thereby puts to test our current understanding of the distribution of galaxies and dark matter in the Universe.

In Chapter 4 the halo model is also fitted to the G2L signal measured by Saghiha+17 in the Millennium Simulation into which galaxies have been included using the SAM by Henriques+15. In addition to predictions for G2L, Saghiha+17 also published predictions of the G3L aperture statistics for the same stellar mass samples used in this work and compared them to the CFHTLenS G3L observations. The halo model predictions for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are compared to these SAM predictions, and it is discussed whether the accuracy of the halo model predictions is comparable to the one of the SAM.

The content of this chapter is being prepared for submission to *Astronomy & Astrophysics*.

6.1 Data

The model predictions are confronted with measurements of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ from CFHTLenS for the galaxy samples described in Sect. 4.1.1. However, because of very faint limits not only the L1, but also the L2 high- z sample contains too few lenses for a G3L analysis (Simon+13). This reduces the total number of samples from 29 to 28. For the analysis $b = 8$ data points of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are used for a range of aperture radii θ between 1 – 10 arcmin. Within this range the signal is not dominated by the transformation bias, which stems from galaxy blending and the finite size of the observed field, and which limits the number

of triplets of galaxies at particularly small and large separations (Simon et al. 2008). The same data is also used in the G3L study by Saghiha+17.

6.2 Method

6.2.1 Predicting the G3L aperture statistics

As explained in Sect. 4.2, the halo model with five free HOD parameters is in a first step fitted to the measured CFHTLenS G2L signal, $\gamma_t(\theta)$, and the total number of galaxies per sample, N_{gal} , for 29 different galaxy samples. These 29 best-fit parameter sets are then in a second step used as inputs to predict G3L in terms of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ for each sample using the analytical model introduced in Sect. 3.2. A key strength of this approach is that the G3L predictions are directly comparable to the CFHTLenS G3L observations because they are based on the best-fits to the CFHTLenS G2L signal for the same galaxy samples.

6.2.2 Uncertainties of the model predictions

Given noisy CFHTLenS G2L data and inherent limitations to the precision of the fitting routine, the best-fit parameters are estimates for the ‘true’ model parameters at the maximum likelihood point. This uncertainty transforms into an uncertainty of the G3L predictions, which is quantified using the same error analysis as applied to all quantities that are derived from the best-fit parameter sets as, e.g. the mean halo mass and G2L. In short (see Sect. 4.2 for details), the covariance of the model parameters at the best-fit is obtained for each galaxy sample from a Fisher analysis (Eq. 4.11). Based on this knowledge, 129 random realisations of the model parameters around their maximum likelihood points are generated, accounting for a non-Gaussian parameter posterior by weighting the parameter realisations with importance sampling. Then, for each realisation the ‘derived quantities’, as e.g. the excess mass maps and aperture statistics, are computed. The errors on the best-fit parameter estimates and the derived quantities are then given by the sample variance of these 129 realisations. This way also the covariance matrix $\mathbf{C}_{\langle \mathcal{N}^2 M_{\text{ap}} \rangle}^{\text{model}}$ of the model prediction for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ is obtained, which contains on the diagonal the error variance in the $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ prediction for the eight aperture radii. This error estimate is here adopted as the uncertainty of the best-fit $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ prediction, although formally it is the uncertainty of the mean of the random realisations.

The deviation of the model prediction from the measured CFHTLenS signal is quantified using a χ^2 -test, where χ^2 is defined in analogy to Eq. (4.1) as

$$\chi^2 = \left(\mathbf{d}^{\text{model}} - \mathbf{d} \right)^{\text{T}} \mathbf{C}^{-1} \left(\mathbf{d}^{\text{model}} - \mathbf{d} \right). \quad (6.1)$$

The vector $\mathbf{d}^{\text{model}}$ contains the best-fit model prediction for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ for the eight considered aperture scale radii θ , and the vector \mathbf{d} the observed CFHTLenS G3L signal for the same aperture radii. The covariance matrix \mathbf{C} is assumed to be the sum of the error covariances of the G3L measurement and the model prediction,

$$\mathbf{C} = \mathbf{C}_{\langle \mathcal{N}^2 M_{\text{ap}} \rangle}^{\text{obs}} + \mathbf{C}_{\langle \mathcal{N}^2 M_{\text{ap}} \rangle}^{\text{model}}, \quad (6.2)$$

which is valid if the measurement and model errors are uncorrelated. As the uncertainties in the model are partly due to the noise in the G2L signal the model is fitted to, the assumption of uncorrelated errors is not exactly true but still made. Here $C_{\langle \mathcal{N}^2 M_{\text{ap}} \rangle}^{\text{obs}}$ is obtained by Jackknife resampling the 129 CFHTLenS pointings (Simon+13). Because both, $C_{\langle \mathcal{N}^2 M_{\text{ap}} \rangle}^{\text{obs}}$ and $C_{\langle \mathcal{N}^2 M_{\text{ap}} \rangle}^{\text{model}}$, are computed from 129 pointings/realisations, their sum C can be inverted in Eq. (6.1) using the estimator by Hartlap et al. (2007).

6.3 $\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ – model predictions versus observations

The model predictions for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ are shown and compared to the CFHTLenS measurements in Figs. 6.1 and 6.2 for the stellar mass samples, in Figs. 6.3 and 6.4 for the luminosity samples, and in Fig. 6.5 for the galaxy-type samples. In each figure the dashed line displays the best-fit prediction with the gray shaded error band giving the $1\text{-}\sigma$ uncertainty of the prediction. Since the best-fit prediction itself does not contain importance sampling, also the mean over the 129 random realisations for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ per bin in aperture scale radius θ is shown as dotted line. The predictions are compared to the observed CFHTLenS signal, which is given by the data points with the error bars representing the $1\text{-}\sigma$ uncertainty. A log-log version of the figures is presented in the appendix (Figs. H.1 - H.5).

The resulting reduced χ^2 -values of the model test are presented in Table 6.1, where boldface values indicate a tension between model prediction and measurement at a 95% confidence level for eight degrees of freedom ($\chi^2/8 > 1.94$). In general a good agreement between model predictions and CFHTLenS data can be stated as the reduced χ^2 exceeds 1.94 only for one sample, which is what one expects as false positive rate.

However, for this particular sample, the sm2 low- z sample, the deviation of the best-fit curve from the observed signal is a result of the best-fit prediction for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ not accounting for a non-Gaussian parameter posterior, in contrast to the mean over the importance-sampled random realisations, which agrees reasonably well with the observations. That the mean is outside the confidence region of the best-fit (also in case of L6 low- z) is also a consequence of the skewed parameter posterior distribution for this sample.

For some samples significant discrepancies between the best-fit prediction and the measurement data appear to be evident in the figures (e.g. for sm7 low- z , sm4 high- z , as well as for L3, L4 and L5 high- z), albeit the χ^2 -test indicates no tension. It seems that generally the data points for larger θ are better fit than those for smaller θ . However, the log-log versions of the figures reveal that the CFHTLenS data point for the largest considered θ is for many samples (e.g. for sm2 low- z , sm3 high- z , sm5 high- z) above and

Table 6.1: χ^2 -test for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ best-fit predictions and CFHTLenS measurements for all galaxy and redshift samples. Given are the reduced $\chi^2/\text{d.o.f.}$ values with eight degrees of freedom. Bold values indicate a tension between prediction and measurement at a 95% confidence level.

Sample	sm1	sm2	sm3	sm4	sm5	sm6	sm7	L1	L2	L3	L4	L5	L6	ETG	LTG
low- z	0.79	3.78	0.30	0.50	0.96	0.75	0.86	1.12	0.18	0.61	1.67	0.79	0.82	0.39	1.91
high- z	1.09	1.10	1.12	1.50	1.41	1.56	0.75	-	-	0.95	0.90	1.15	0.96	0.84	0.59

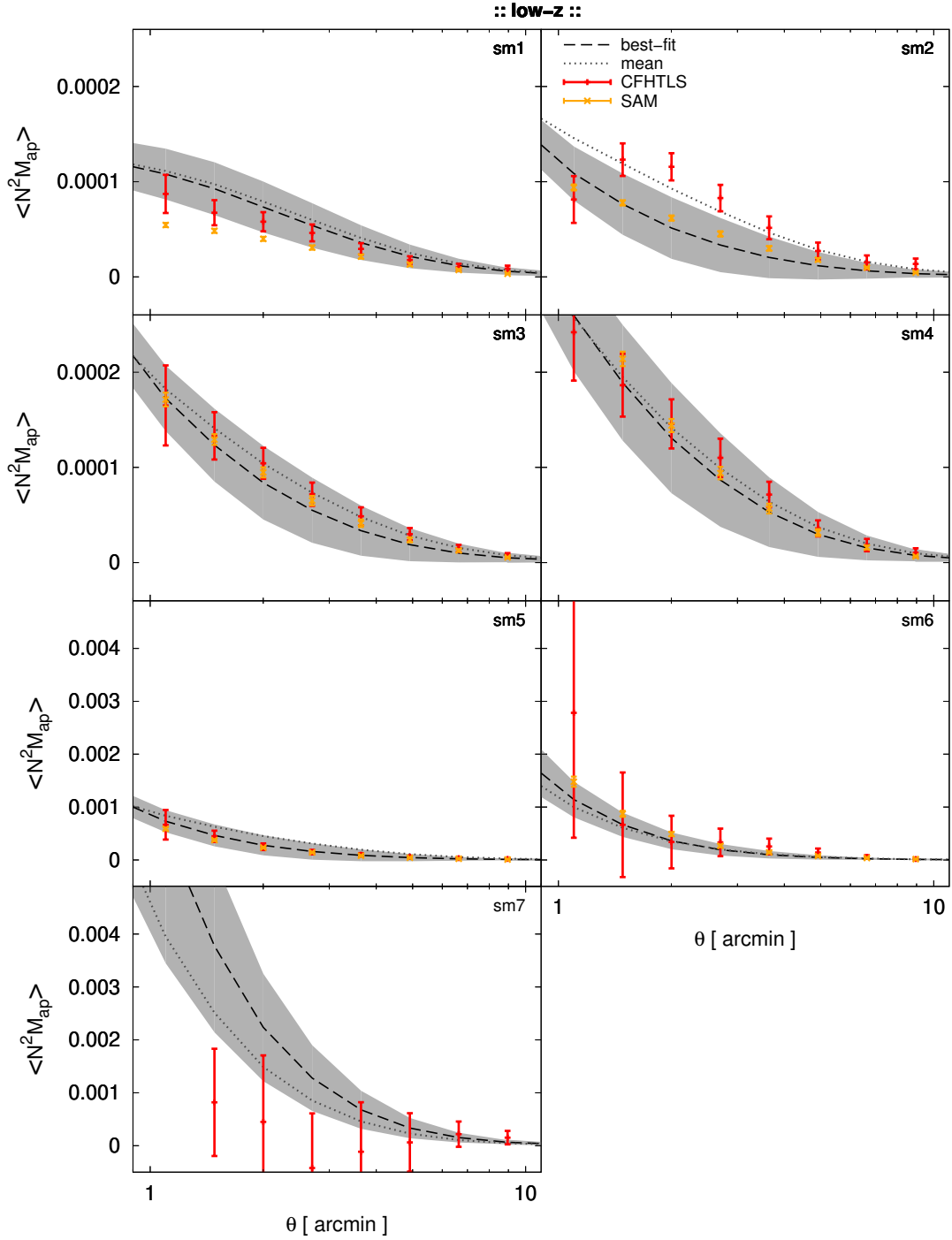


Figure 6.1: Comparison of the $\langle N^2 M_{\text{ap}} \rangle(\theta)$ halo model predictions (dashed line) to the CFHTLenS measurements (bar symbols) as a function of aperture scale radius θ . Shown are the results for the stellar mass samples and low- z . The error bars and the error bands give the $1\text{-}\sigma$ uncertainty of the measurement and model, respectively. The mean over the random realisations is given by the dotted line. In addition, the SAM predictions by Saghiha+17 are shown (cross symbols). Note the different scales across the panels.

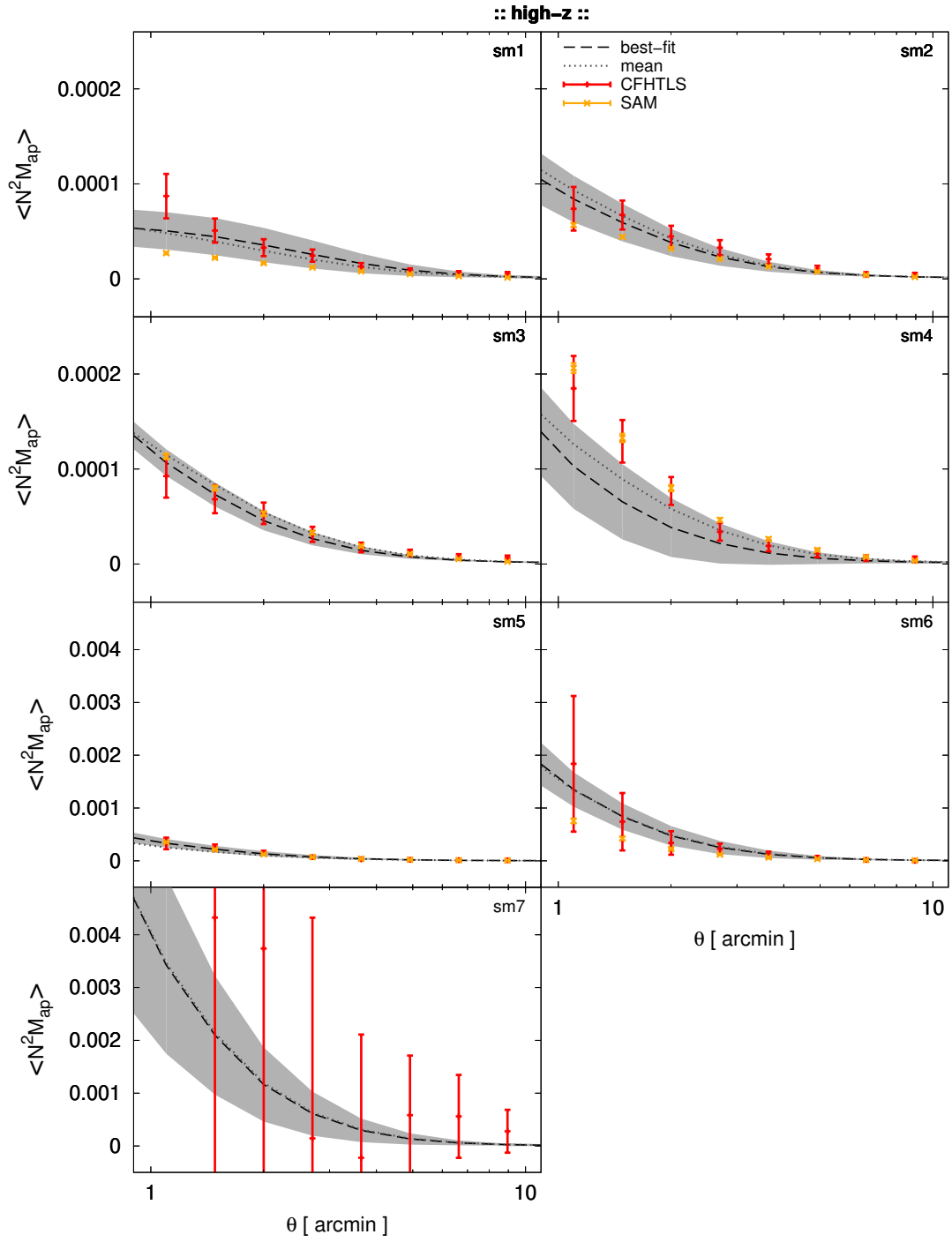


Figure 6.2: As Fig. 6.1 but for high- z .

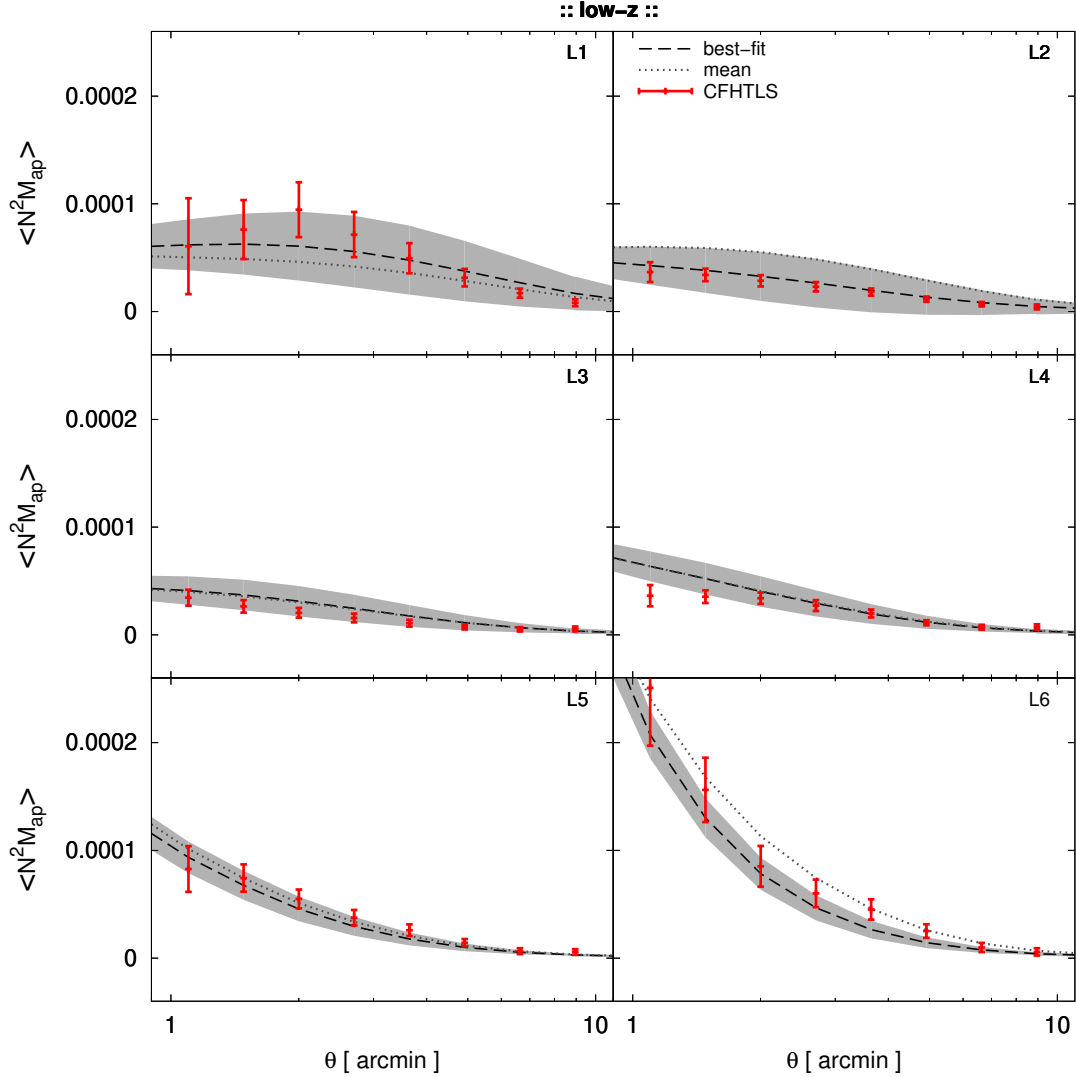


Figure 6.3: Comparison of the $\langle N^2 M_{\text{ap}} \rangle(\theta)$ halo model predictions (dashed line) to the CFHTLenS measurements (bar symbols) as a function of aperture scale radius θ . Shown are the results for the luminosity samples and low- z . The error bars and the error bands give the $1\text{-}\sigma$ uncertainty of the measurement and the model, respectively. The mean over the random realisations is given by the dotted line.

outside the confidence region of the halo model predictions.

The reduced χ^2 values presented in Table 6.1 are relatively small compared to the ones of the G2L model fits ($0.5 \lesssim \chi^2/9 \lesssim 3.8$, Table 4.4), although the G3L predictions are based on the G2L model fits. However, for the model fits only uncertainties in the measurements have been taken into account (see Eq. 4.1). By contrast, for the calculation of the χ^2 values for the G3L analysis, errors stemming from both the measurements and the model are taken into account (Eq. 6.1).

Figures 6.1 and 6.2 also show the SAM predictions of $\langle N^2 M_{\text{ap}} \rangle(\theta)$ by Saghiha+17 for

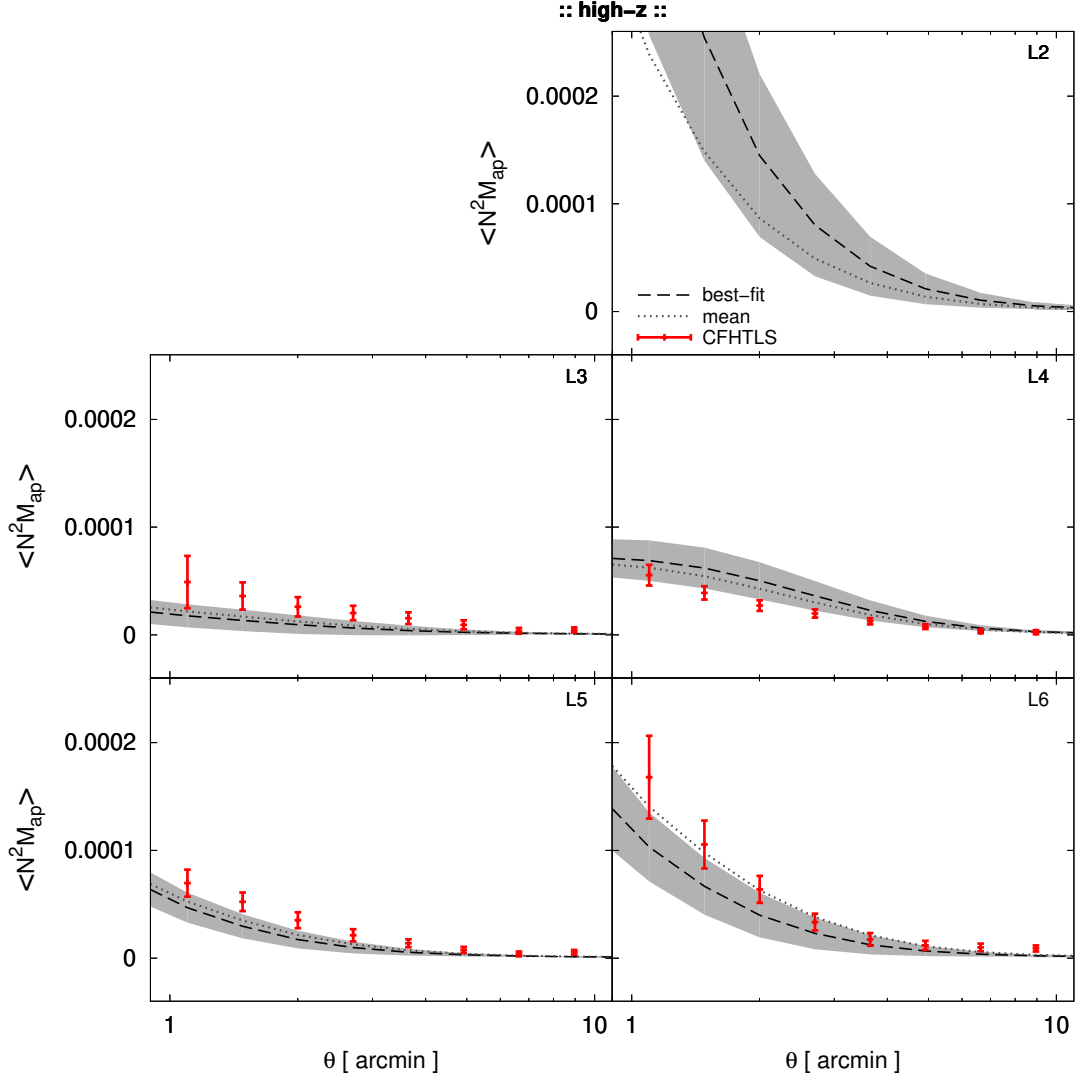


Figure 6.4: As Fig. 6.3 but for high- z . No CFHTLenS measurement for the L2 sample is available.

the stellar mass samples. Whereas for low- z a good agreement between halo model and SAM predictions is seen for all samples other than the sm1 sample, discrepancies between the two models occur for high- z (sm1, sm4, and sm6). There does not seem to be any relationship between the two predictions, as e.g. that one model generally underpredicts the other, so the reason for these discrepancies is unclear.

Regarding the accuracy of predicting the CFHTLenS $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$, Saghiha+17 quantify the deviation between the SAM predictions and the CFHTLenS observations in the same way as this work, using the χ^2 -test defined in Eq. (6.1). Their reduced χ^2 values are in the range $0.29 \lesssim \chi^2/\text{d.o.f.} \lesssim 4.13$, which is comparable to the range found in this work (Table 6.1). However, note that the SAM predictions have a much smaller uncertainty compared to the halo model predictions. To conclude, the combination of the information

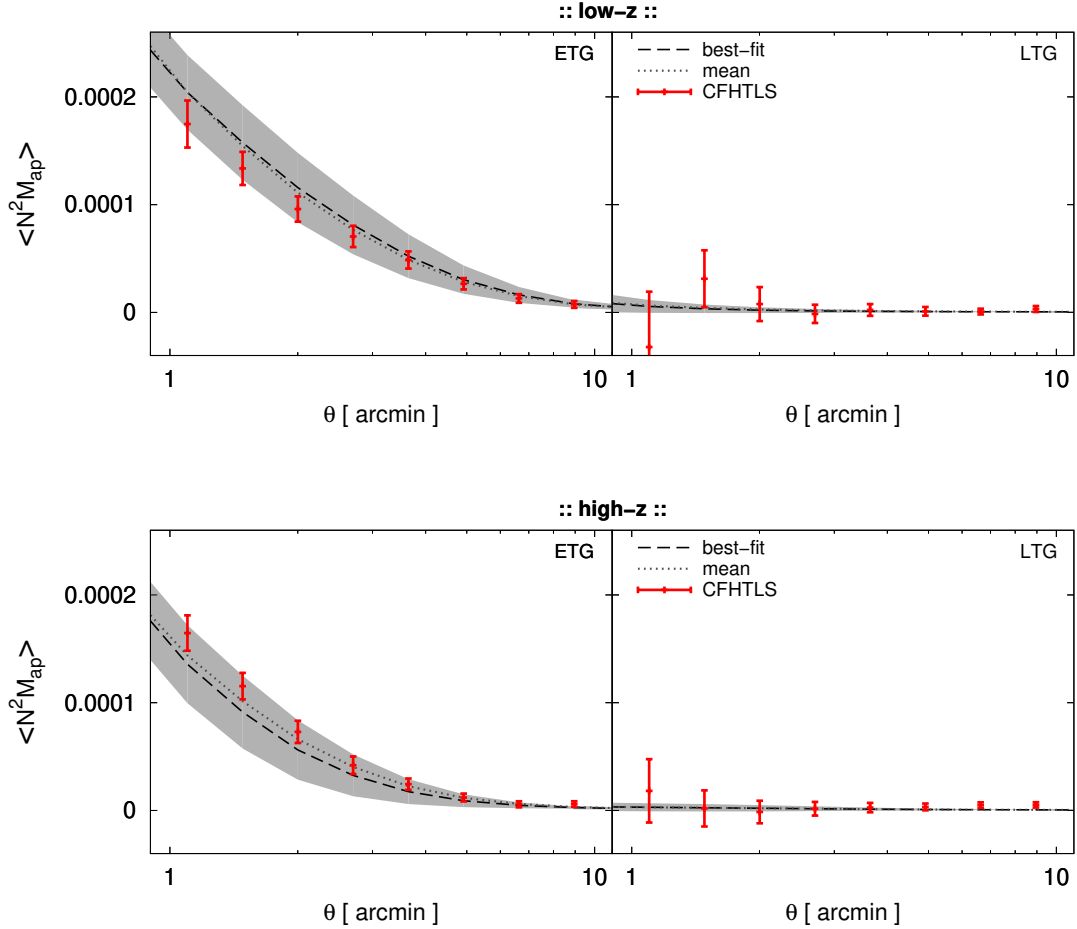


Figure 6.5: Comparison of the $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ halo model predictions (dashed line) to the [CFHTLenS](#) measurements (bar symbols) as a function of aperture scale radius θ . Shown are the results for the galaxy-type samples, for low- z (*top*) and for high- z (*bottom*). The error bars and the error bands give the $1\text{-}\sigma$ uncertainty of the measurement and the model, respectively. The mean over the random realisations is given by the dotted line.

contained in Figs. 6.1 and 6.2, and the χ^2 -test indicates that the halo model employed in this work can compete with [SAMs](#) with respect to the accuracy of predicting third-order galaxy-dark matter clustering in terms of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$.

6.4 Conclusions

This chapter is concerned with the question whether a ‘standard’ halo model can consistently describe second- and third-order galaxy-matter correlations in the Universe as probed by [G2L](#) and [G3L](#). The comparison of the halo model prediction for the [G3L](#) aperture statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ to its observational counterpart from [CFHTLenS](#) in this chapter shows that the employed halo model can successfully predict [G3L](#), based on the

model fits to the **CFHTLenS G2L** presented in Chapter 4. Only for one out of 28 galaxy samples a tension between the best-fit model prediction and the measurement at the 95% confidence level is found, which is expected as false positive rate. This can be seen as a confirmation of the model itself and its assumptions on the distribution of dark matter and galaxies. But at the same time this also suggests that fitting the halo model to **G2L** and **G3L** simultaneously would only add constraining power because of the availability of more data points, but not because **G3L** has more discriminating power than **G2L**.

This is in contradiction to the findings by Saghiha+17, who obtain for their comparison of **SAM** predictions to **CFHTLenS** measurements larger reduced χ^2 -values for **G3L** compared to **G2L**, which they interpret as **G3L** having more discriminating power than **G2L**. However, Saghiha+17 consider for the χ^2 -tests solely the error covariance of the measurements and not of the model. In the approach used in this work – to predict **G3L** based on model fits to **G2L** – uncertainties in the predictions, however, do arise due to uncertainties in the **CFHTLenS G2L** measurements, and need to be accounted for in the χ^2 -test for **G3L**.

Although it depends on the application whether the halo model or a **SAM** is better suited for describing the distribution of matter in the Universe, it is a major success for the halo model to largely reproduce the predictions of the **SAM** for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$, and to achieve the same level of accuracy in predicting $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ as measured from **CFHTLenS**. Yet, it should be noted that despite these good agreements, the accuracy tests using mock data (Sect. 4.3.1) showed that the halo model fails to reproduce the mean halo mass $\langle M_{\text{h}} \rangle$, and fraction of central galaxies f_{cen} , with an inaccuracy of up to 45% and 20%, respectively. That the **CFHTLenS G3L** observations are nevertheless well matched, thus indicates that galaxy-galaxy-(galaxy) lensing is largely independent of the identified inadequate description of satellite galaxies in the halo model. On the other hand, the sensitivity analysis (Sect. 5.4) revealed that **G3L** shows the strongest dependence on those **HOD** parameters that describe the behaviour of satellite galaxies in the halo. A sensitivity analysis that accounts for correlations between the model parameters, as well as investigations using a halo model that includes a prescription for subhalo physics, are the next crucial steps to resolve this contradiction.

Conclusions & Outlook

This doctoral thesis answers the question whether a ‘standard’ halo model can consistently describe second- and third-order cross-correlations of the galaxy density and matter density field as probed by applications of the weak gravitational lensing effect: galaxy-galaxy lensing (**G2L**) and galaxy-galaxy-galaxy lensing (**G3L**). Moreover, this work presents first-time halo model predictions of excess mass maps, and quantifies their dependence on galaxy and halo model properties, providing key insights in the functional form of the predicted and measured **G3L** signal. This chapter summarises the performed research, singling out the main results and highlighting their significance in the field of research.

This work builds up on a complex mathematical framework which connects the standard model of Cosmology, the Λ CDM model, with a model for the galaxy-matter connection, the halo model, and with the description of the physical phenomenon that illuminates this galaxy-matter connection, weak gravitational lensing. All aspects of modern cosmology, the halo model and weak gravitational lensing crucial for this doctoral thesis are introduced in **Chapters 1 - 3**. This mathematical framework is implemented into a numerical code that was written as part of this work to, first, fit the halo model to measurements of **G2L** and, second, to generate halo model predictions for **G3L** based on the obtained best-fit model parameters.

The foundation for this work is laid in **Chapter 4**, where the results from the halo model fits to **G2L** measured from the Canada-France Hawaii Telescope Lensing Survey (**CFHTLenS**) for 29 galaxy samples are presented. These best-fit models are necessary for generating halo model predictions of **G3L**. Possible limitations of the halo model and the employed fitting routine are investigated independently of the **CFHTLenS** data by performing halo model fits to mock **G2L** data from the Millennium Simulation, in which galaxies are implemented using a semi-analytical model (**SAM**). The main conclusions of this chapter are:

- The halo model test using mock data reveals that the model has difficulties correctly predicting the mean halo mass $\langle M_{\text{h}}^{\bar{z}} \rangle$ and the fraction of central galaxies $f_{\text{cen}}^{\text{V}}$, with the inaccuracies reaching up to 45% and 20%, respectively. The discrepancy is connected to the treatment of satellite galaxies in the employed halo model and appears to be related to the absence of subhalos in the model. Follow-up research using a model that includes subhalos promises insights into the physics of subhalos and satellites.
- The best-fit parameters to the **CFHTLenS** data are generally in good agreement

with the literature. The ratio of stellar-to-halo mass $M_*(M_h)/M_h$ derived from the best-fit parameters is also in good overall agreement with the literature, particularly around the peak of $M_*(M_h)/M_h$ at $M_h \approx 10^{12} h^{-1} M_\odot$. For lower halo masses, the halo model somewhat overestimates the fraction of stellar mass compared to [Coupon et al. \(2015\)](#), whereas for larger halo masses the model underestimates the fraction of stellar mass due to an incompleteness of the galaxy selection in [CFHTLenS](#). These results confirm that the used halo model can keep up with state of the art methods to describe the galaxy-halo connection.

Chapter 5 presents first-time predictions of excess mass maps from the halo model. The predictions are based on the halo model fits to the [CFHTLenS G2L](#) for 29 galaxy samples, which allows to study not only the dependence of excess mass on halo model but also on galaxy properties. The map predictions are discussed together with predictions for the aperture statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$. The main conclusions of this chapter are:

- The halo model can reproduce observed trends ([Simon et al. 2008](#), [Simon et al. 2013](#), [Simon et al. 2019](#)), i.e. the amplitude of excess mass increases with decreasing lens-lens separation and the signal around pairs of early-type galaxies is more than a factor of ten higher than around late-types.
- The halo model predicts an increase of [G3L](#) for pairs of more massive and luminous galaxies, indicating that these galaxies inhabit more massive halos. Due to the low signal-to-noise of the [CFHTLenS](#) excess mass map measurements, these trends can currently only be measured for $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$. By the time of their completion, surveys such as the Kilo-Degree ([KIDS](#))¹ and Dark Energy Survey ([DES](#))² will cover approximately 40 times the area of the sky probed by [CFHTLenS](#) (154 deg^2) and have improved redshift and shear estimates. This data should allow the measurement of excess mass maps for distinct stellar mass and luminosity samples.
- A recent measurement of excess mass maps for a combined sample of [CFHTLenS](#) galaxies reveals a vertical bulge-like feature ([Simon et al. 2019](#)). The comparison of halo model predictions of excess mass maps to these measurements shows that the mode cannot explain this feature. However, a generally good agreement is found regarding the central amplitude of excess mass and its drop-off towards the outer regions of the map as well as with the predictions of the [SAM](#) by [Henriques et al. \(2015\)](#). The halo model could be used in a follow-up study to test whether, e.g. a misalignment between the distribution of the intra-cluster medium and of the dark matter, or between the orientation of lens pairs and of the parent halos, could cause the observed feature.
- To be able to interpret measurements of [G3L](#), it is crucial to know the relative importance of the one-, two- and three-halo terms, i.e. the contributions from correlations stemming from a single halo, from two halos, and from three halos, respectively. According to the halo model, the one-halo term dominates below aperture scales of

¹<http://kids.strw.leidenuniv.nl/index.php>

²<https://www.darkenergysurvey.org>

10 arcmin, which corresponds the range probed by **CFHTLenS** (1 – 10 arcmin). This is the case for all samples other than the late-type sample, for which the one-halo term is suppressed as a consequence of late-types typically being field galaxies.

- Predicting **G3L** with the halo model is a computationally costly endeavour, and accurate approximations are desirable. For samples dominated by the one-halo term (here all other than the late-type sample), $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ can be approximated by the one-halo term with an accuracy better than 10% up to scales of 3 arcmin.
- To explore the sensitivity of **G3L** to the **HOD** parameters, the parameters are varied individually by $\pm 20\%$ around their best-fit values. For four out of five parameters, changes in the predicted excess mass of more than 20% occur. The strongest changes are found for the lowest stellar mass samples, suggesting that these samples are better suited to constrain the **HOD**. The changes are largest in the angular range tested by **CFHTLenS**. Again the exception is the late-type sample, for which the change is largest on scales smaller than 1 arcmin. With future measurements of **G3L** becoming reliable on scales smaller than 1 arcmin (Laila Linke; priv. comm.), this could help break parameter degeneracies.
- The most interesting results with regard to galaxy evolution are to be expected from the galaxy-type samples: the late-type sample shows a unique behaviour compared to all other samples at low redshifts, but adapts to the one of the other samples at high redshifts.

This work takes only the first steps in exploring **G3L** with a halo model. Among the possible avenues for future research are accounting for model parameter correlations in the sensitivity analysis, considering non-equilateral triangle configurations for the aperture statistics $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta_1, \theta_2, \theta_3)$, and probing the cross-correlation for galaxy pairs made up of distinct galaxy types.

The ability of the halo model to describe **G2L** and **G3L** consistently is tested in **Chapter 6** by confronting the halo model predictions of **G3L** in terms of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ with measurements from **CFHTLenS**. The results are also compared to findings by **Saghiha et al. (2017)**, who measure $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ from the Millennium Simulation where galaxies have been implanted using a **SAM**. The main conclusions of Chapter 6 are:

- With a tension at the 95% confidence level for only one out of 28 galaxy samples, an excellent agreement between halo model predictions and **CFHTLenS** measurements is found. The question whether the employed halo model can successfully predict **G3L**, based on the best-fit models to the **CFHTLenS G2L**, can therefore be answered with a resounding ‘yes’. This can be seen as a confirmation of the halo model itself and its assumptions on the distribution of galaxies and dark matter.
- The halo model is able to describe the **CFHTLenS** measurement at the same level of accuracy as the **SAM** by **Henriques et al. (2015)**, which is a major success for the halo model.

- That the [CFHTLenS G3L](#) observations are well matched by the halo model predictions, despite the model's inability to reproduce the mean halo mass $\langle M_h \rangle$ and fraction of central galaxies f_{cen} in mock data, indicates that galaxy-galaxy-(galaxy) lensing is largely independent of the description of satellite galaxies in the halo model. On the other hand, the sensitivity analysis shows that [G3L](#) depends most strongly on the [HOD](#) parameters describing the behaviour of satellite galaxies. To resolve this contradiction, follow-up research should, first, redo the sensitivity analysis accounting for correlations between the model parameters. And second, extend the employed halo model to include subhalos.

Besides the possible improvements mentioned throughout, the most straightforward extension of this work would be to optimise the halo model code in order to make simultaneous model fits to [G2L](#) and [G3L](#) computationally feasible. With the upcoming measurements of [G3L](#) with unprecedented accuracy from [KIDS](#) and [DES](#), it will be possible to constrain the halo model parameters better than ever before.

Appendix A

Natural constants & model parameters

Natural and astrophysical constants	Symbol	Value	Unit
Pi	π	3.1415926535897	rad
Speed of light (in vacuum)	c	299792.458	km/s
Gravitational constant	G_N	6.67384×10^{-11}	$\text{m}^3/(\text{kg s}^2)$
Solar mass	M_\odot	1.98892×10^{30}	kg
Parsec	pc	3.0856775×10^{16}	m
Cosmological parameters CFHTLenS (WMAP9; Hinshaw et al. 2013)			
Dark energy density	Ω_Λ	0.721	
Baryonic matter density	Ω_b	0.0463	
Dark matter density	Ω_{dm}	0.233	
Spectral index primordial spectrum	n_s	0.972	
Hubble constant	H_0	100	$h \text{ km/s/Mpc}$
Fluctuation amplitude at $8 h^{-1} \text{ Mpc}$	σ_8	0.821	
Cosmological parameters mock data (first-year WMAP; Spergel et al. 2003)			
Dark energy density	Ω_Λ	0.75	
Baryonic matter density	Ω_b	0.045	
Dark matter density	Ω_m	0.25	
Spectral index primordial spectrum	n_s	1	
Hubble constant	H_0	73	km/s/Mpc
Fluctuation amplitude at $8 h^{-1} \text{ Mpc}$	σ_8	0.9	
Halo model parameters			
Critical density	$\rho_{\text{crit}}(0)$	2.775×10^{11}	$h^2 M_\odot/\text{Mpc}^3$
Linear density contrast	$\delta_c^{\text{EdS}}(0)$	1.686	
Density ratio of a virialised halo	$\Delta_{\text{vir}}^{\text{EdS}}(0)$	178	
1st parameter Sheth & Tormen	p	0.707	
2nd parameter Sheth & Tormen	q	0.3	
1st parameter NFW density profile	c_0	9	
2nd parameter NFW density profile	α	0.13	
Halo occupation distribution parameters (default)			
Scatter central galaxy	$\sigma_{\log(M_h)}$	0.1	
Slope satellite power-law	β	1	
Threshold mass	M_{th}	10^{12}	M_\odot/h
Amplitude satellite power-law	M_1	10^{13}	M_\odot/h
Central fraction	α_{cen}	0.1	

Table A.1: Natural constants, astrophysical constants, cosmological parameters, halo model parameters and halo occupation distribution parameters used in this work.

Appendix B

Eisenstein & Hu transfer function

In this work the widely used fitting function by Eisenstein & Hu (1998) is numerically implemented, a physically well-motivated expression to describe the growth of cold dark matter perturbations accounting for the presence of baryons and vice versa,

$$T(k) = \frac{\Omega_c}{\Omega_m} T_c(k) + \frac{\Omega_b}{\Omega_m} T_b(k). \quad (\text{B.1})$$

In this context Ω_m is the sum of the baryonic matter density Ω_b and the cold dark matter density Ω_c .

If all particle species would be free of interactions, the matter transfer function would only depend on the horizon size $k_{\text{eq}} = 7.46 \times 10^{-2} \Omega_m h^2 \Theta_{2.7}^{-2} \text{Mpc}^{-1}$ at radiation-matter equality $z_{\text{eq}} = 2.50 \times 10^4 \Omega_m h^2 \Theta_{2.7}^{-4}$. The CMB temperature enters these expressions by the term $2.7 \Theta_{2.7} \text{K}$, where $T_{\text{CMB}} = 2.728 \pm 0.004 \text{K}$ based on Cosmic Background Explorer (COBE) observations. The Hubble constant is defined by $H_0 = h 100 \text{km s}^{-1} \text{Mpc}^{-1}$ with $h = 0.72$. However, the wave vector k is assumed to have the dimension Mpc^{-1} , treating h as a unit.

The interaction of different species leads to further terms shaping the transfer function, where recombination differentiates between the era of a baryon-, electron- and photon-plasma with perturbations propagating as acoustic waves and the era of neutral atoms and free-streaming photons. Respectively, the size of the sound horizon s at recombination is needed,

$$s = \int_0^{t(z_d)} c_s (1+z) dt = \frac{2}{3k_{\text{eq}}} \sqrt{\frac{6}{R_{\text{eq}}}} \ln \frac{\sqrt{1+R_d} + \sqrt{R_d+R_{\text{eq}}}}{1 + \sqrt{R_{\text{eq}}}}, \quad (\text{B.2})$$

with the baryon to photon ratio at radiation-matter equality $R_{\text{eq}} = R(z_{\text{eq}})$ and the one at the drag epoch $R_d = R(z_d)$ given by

$$R(z) \equiv 3\rho_b/4\rho_\gamma = 31.5\Omega_b h^2 \Theta_{2.7}^{-4} (z/10^3)^{-1}. \quad (\text{B.3})$$

Here the redshift at the drag epoch is defined as

$$\begin{aligned} z_d &= 1291 \frac{(\Omega_m h^2)^{0.251}}{1 + 0.659(\Omega_m h^2)^{0.828}} [1 + b_1(\Omega_b h^2)^{b_2}], \\ b_1 &= 0.313(\Omega_m h^2)^{-0.419} [1 + 0.607(\Omega_m h^2)^{0.674}], \\ b_2 &= 0.238(\Omega_m h^2)^{0.223}. \end{aligned} \quad (\text{B.4})$$

Furthermore, the physical phenomena that the radiation pressure pushes the photons out of the potential wells, dragging the electrons and protons with them, leads to a wash out of the early density perturbations and, thus, defines the last length scale to be considered in this discussion. This *Silk damping* scale is given by,

$$k_{\text{silk}} = 1.6(\Omega_b h^2)^{0.52}(\Omega_m h^2)^{0.73} [1 + (10.4\Omega_m h^2)^{-0.95}] \text{ Mpc}^{-1}. \quad (\text{B.5})$$

The analytical equivalent to the foregoing qualitative discussion is given by the fitting function Eq. (B.1), where the first term is given by the following set of equations:

$$T_c(k) = f \tilde{T}_0(k, 1, \beta_c) + (1 - f) \tilde{T}_0(k, \alpha_c, \beta_c), \quad (\text{B.6})$$

with

$$f = \frac{1}{1 + (ks/5.4)^4}. \quad (\text{B.7})$$

Further,

$$\tilde{T}_0(k, \alpha_c, \beta_c) = \frac{\ln(e + 1.8\beta_c q)}{\ln(e + 1.8\beta_c q) + Cq^2}, \quad (\text{B.8})$$

with

$$C = \frac{14.2}{\alpha_c} + \frac{386}{1 + 69.9q^{1.08}}, \quad (\text{B.9})$$

and the dimensionless parameter

$$q = (k/\text{Mpc}^{-1})\Theta_{2.7}^2(\Omega_m h^2)^{-1} = \frac{k}{13.41k_{\text{eq}}}. \quad (\text{B.10})$$

The parameters α_c and β_c are fit by

$$\begin{aligned} \alpha_c &= a_1^{-\Omega_b/\Omega_m} a_2^{-(\Omega_b/\Omega_m)^3}, \\ a_1 &= (46.9\Omega_m h^2)^{0.670} [1 + (32.1\Omega_m h^2)^{-0.532}], \\ a_2 &= (12.0\Omega_m h^2)^{0.424} [1 + (45.0\Omega_m h^2)^{-0.582}], \\ \beta_c^{-1} &= 1 + b_1 [(\Omega_c/\Omega_m)^{b_2} - 1], \\ b_1 &= 0.944 [1 + (458\Omega_m h^2)^{-0.708}]^{-1}, \\ b_2 &= (0.395\Omega_m h^2)^{-0.0266}. \end{aligned} \quad (\text{B.11})$$

The second term in the fitting function can be expressed by,

$$T_b = \left[\frac{\tilde{T}_0(k; 1, 1)}{1 + (ks/5.2)^2} + \frac{\alpha_b}{1 + (\beta_b/ks)^3} e^{-(k/k_{\text{silk}})^{1.4}} \right] j_0(k\tilde{s}), \quad (\text{B.12})$$

where

$$\begin{aligned} \alpha_b &= 2.07k_{\text{eq}}s(1 + R_d)^{-3/4} G \left(\frac{1 + z_{\text{eq}}}{1 + z_d} \right), \\ G(y) &= y \left[-6\sqrt{1 + y} + (2 + 3y) \ln \left(\frac{\sqrt{1 + y} + 1}{\sqrt{1 + y} - 1} \right) \right], \end{aligned} \quad (\text{B.13})$$

and

$$\beta_b = 0.5 + \frac{\Omega_b}{\Omega_m} + \left(3 - 2\frac{\Omega_b}{\Omega_m}\right) \sqrt{(17.2\Omega_m h^2)^2 + 1}. \quad (\text{B.14})$$

Finally, the spherical Bessel function $j_0(x) \equiv \sin(x)/x$ depends on a correction term for small k ,

$$\tilde{s}(k) = \frac{s}{[1 + (\beta_{\text{node}}/ks)^3]^{1/3}}, \quad (\text{B.15})$$

with

$$\beta_{\text{node}} = 8.41(\Omega_m h^2)^{0.435}. \quad (\text{B.16})$$

Appendix C

Normalisation of the linear power spectrum

The normalisation factor A of the linear power spectrum can be derived by assuming a *linear galaxy dark matter bias* $\sigma_{8,\text{gal}}^2 \equiv b^2 \sigma_8^2$ between the variance of number counts of galaxies in a sphere of $R = 8 h^{-1} \text{Mpc}$, where $\sigma_{8,\text{gal}}^2 \approx 1$, and the variance in the dark matter distribution in the same sphere, $\sigma_8 \approx 0.8$ (see Table A.1). The variance in the dark matter distribution is related to the amplitude A of the linear dark matter power spectrum via the density field δ_R smoothed on the scale $R = 8 h^{-1} \text{Mpc}$,

$$\sigma_8^2 = \sigma^2(R = 8 h^{-1} \text{Mpc}) = \langle \delta_R^2(\mathbf{x}) \rangle = \int \frac{d^3 k}{(2\pi)^3} P_R(k), \quad (\text{C.1})$$

with

$$\delta_R(\mathbf{x}) = \int d^3 y \delta(\mathbf{y}) W_R(|\mathbf{x} - \mathbf{y}|), \quad (\text{C.2})$$

and W_R being a top-hat filter function, which is in Fourier space given by

$$W_R(k) = \frac{3}{(kR)^3} [\sin(kR) - kR \cos(kR)]. \quad (\text{C.3})$$

The power spectrum of the smoothed field is given by

$$P_R(k) = P_{\text{lin}}(k) |W_R(k)|^2. \quad (\text{C.4})$$

Applying the transformation from cartesian coordinates to spherical coordinates $\int_{\mathbb{R}^3} d^3 x = \int_0^\infty dr r^2 \int_0^\pi d\theta \sin(\theta) \int_0^{2\pi} d\phi$, the variance can be rewritten using Eq. (C.4) and Eq. (1.54),

$$\begin{aligned} \sigma_8^2 &= \int \frac{d^3 k}{(2\pi)^3} |W_8(k)|^2 P_{\text{lin}}(k, a = 1) \\ &= \int \frac{d^3 k}{(2\pi)^3} |W_8(k)|^2 A k^{n_s} T^2(k) D^2(a = 1) \\ &= \frac{A}{(2\pi)^3} \int_0^\infty dk k^2 W_8(k)^2 k^{n_s} T^2(k) \int_0^\pi d\theta \sin(\theta) \int_0^{2\pi} d\phi \\ &= \frac{A}{(2\pi)^3} \int_0^\infty dk k^2 W_8(k)^2 k^{n_s} T^2(k) \times [-\cos(\pi) + \cos(0)] \times [2\pi - 0] \\ &= \frac{A}{2\pi^2} \int_0^\infty dk k^2 W_8(k)^2 k^{n_s} T^2(k). \end{aligned} \quad (\text{C.5})$$

The amplitude A of the linear power spectrum from perturbation theory is finally given by

$$A = \sigma_8^2 \left[\frac{1}{2\pi^2} \int_0^\infty dk k^2 W_8^2(k) T^2(k) k^{n_s} \right]^{-1}. \quad (\text{C.6})$$

Appendix D

Derivative in the halo mass function

The derivative $\frac{d \ln \nu}{d \ln m}$ of the dimensionless variable ν , can be analytically calculated by

$$\frac{d \ln \nu}{d \ln m} = \frac{d}{d \ln m} \ln \left(\frac{\delta_c(z)}{D(z)\sigma(m)} \right) = -\frac{d}{d \ln m} \ln \sigma(m) = -\frac{m}{\sigma} \frac{d}{d m} \sigma(m), \quad (\text{D.1})$$

using that $\frac{d}{d x} \ln(x) = \frac{1}{x}$. In a next step the derivation of $\frac{d}{d m} \sigma(m)$ is executed, using

$$\begin{aligned} \frac{d}{d m} &= \frac{d R}{d m} \frac{d}{d R} = \frac{d}{d m} \left(\frac{3 m}{4 \pi \bar{\rho}} \right)^{1/3} \frac{d}{d R} = \left[\frac{1}{3} \left(\frac{3 m}{4 \pi \bar{\rho}} \right)^{-2/3} \frac{3}{4 \pi \bar{\rho}} \right] \frac{d}{d R} \\ &= \left[\frac{1}{3} \left(\frac{3 m}{4 \pi \bar{\rho}} \right)^{-2/3} \frac{3 m}{4 \pi \bar{\rho}} \frac{1}{m} \right] \frac{d}{d R} = \left[\frac{1}{3} R \frac{1}{m} \right] \frac{d}{d R}, \end{aligned} \quad (\text{D.2})$$

so that

$$\frac{d}{d m} \sigma(m) = \left[\frac{1}{3} R \frac{1}{m} \right] \frac{d}{d R} \sigma(m). \quad (\text{D.3})$$

In a next step the derivative $\frac{d}{d R} \sigma(m)$ is calculated using Eq. (C.5),

$$\begin{aligned} \frac{d}{d R} \sigma(m) &= \frac{d}{d R} \left[\int \frac{d k}{2 \pi^2} k^2 P_{\text{lin}}(k, z=0) |W_R^2(k)| \right]^{1/2} \\ &= \frac{1}{2 \sigma^2(m)} \int \frac{d k}{2 \pi^2} k^2 P_{\text{lin}}(k, z=0) \frac{d |W_R^2(k)|}{d R}. \end{aligned} \quad (\text{D.4})$$

Here the derivative of $\frac{d |W_R^2(k)|}{d R}$ is calculated separately,

$$\begin{aligned} \frac{d |W_R^2(k)|}{d R} &= \frac{d W_R^2(k)}{d R} = \frac{d}{d R} \left(\frac{3}{(k R)^3} [\sin(k R) - k R \cos(k R)] \right)^2 \\ &= 2 W_R(k) \frac{d}{d R} \left(\frac{3}{(k R)^3} [\sin(k R) - k R \cos(k R)] \right) \\ &= 6 W_R(k) \left[\frac{k^2 R \sin(k R)}{(k R)^3} - \frac{W_R(k)}{R} \right]. \end{aligned} \quad (\text{D.5})$$

The final result is then

$$\frac{d \ln \nu}{d \ln m} = -\frac{R}{\sigma^2(m)} \int \frac{dk}{2\pi^2} k^2 P_{\text{lin}}(k, 0) W_R(k) \left[\frac{k^2 R \sin(kR)}{(kR)^3} - \frac{W_R(k)}{R} \right], \quad (\text{D.6})$$

where the derivative of a dimensionless quantity is again dimensionless.

Appendix E

Numerical interfaces to the bispectrum

The numerical transformations of the projected galaxy-matter bispectrum $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ (defined in Eq. 2.36) to the excess mass maps and the aperture statistics are given by Eqs. (3.28) and (3.30), respectively. For these integral transformations an approach using a multipole expansion of the bispectrum is applied. This approach is based on notes by Patrick Simon (priv. comm.), which are reproduced with his permission in the following two sections. In these notes, the angular bispectrum $B^{\text{gg}\kappa}(\ell_1, \ell_2; \ell_3)$ is denoted by $b_{\text{gg}\kappa}(\ell_1, \ell_2, \phi)$.

E.1 Excess mass maps

To tackle the numerical computation of the integral transformation (Eq. 3.28) of the bispectrum, we found it useful to follow the recipe of Zheng (2004) and Szapudi (2004), which is summarised in the following. For this purpose, we express the angular bispectrum $b_{\text{gg}\kappa}(\ell_1, \ell_2, \phi)$ by the moduli of $\ell_i = |\ell_i|$ and the angle ϕ between the vectors ℓ_1 and ℓ_2 . We perform a Fourier expansion of $b_{\text{gg}\kappa}$ with respect to ϕ in terms of the multipoles $B_n(\ell_1, \ell_2)$:

$$b_{\text{gg}\kappa}(\ell_1, \ell_2, \phi) = B_0(\ell_1, \ell_2) + 2 \sum_{n=1}^{\infty} B_n(\ell_1, \ell_2) \cos n\phi \quad (\text{E.1})$$

with

$$B_n(\ell_1, \ell_2) = \frac{1}{2\pi} \int_0^{2\pi} d\phi b_{\text{gg}\kappa}(\ell_1, \ell_2, \phi) \cos n\phi, \quad (\text{E.2})$$

where the symmetries $b_{\text{gg}\kappa}(\ell_1, \ell_2, \phi) = b_{\text{gg}\kappa}(\ell_1, \ell_2, -\phi)$ and $B_n(\ell_1, \ell_2) = B_{-n}(\ell_1, \ell_2)$ are used (parity invariance and real-valued correlation functions). Using this expansion, we can recast Eq. (3.28) with the relation $F_n(\theta_{13}, \theta_{23}) = F_{-n}(\theta_{13}, \theta_{23})$ into

$$\langle \kappa_g \kappa_g \kappa \rangle(\theta_{13}, \theta_{23}, \psi) = F_0(\theta_{13}, \theta_{23}) + 2 \sum_{n=1}^{\infty} F_n(\theta_{13}, \theta_{23}) \cos n\psi, \quad (\text{E.3})$$

where the function $F_n(\theta_{13}, \theta_{23})$ of $\theta_{13} \equiv |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_3|$ and $\theta_{23} \equiv |\boldsymbol{\theta}_2 - \boldsymbol{\theta}_3|$ is defined as

$$F_n(\theta_{13}, \theta_{23}) = \int_0^{\infty} \frac{d\ell_1 d\ell_2 \ell_1 \ell_2}{(2\pi)^2} (-1)^n B_n(\ell_1, \ell_2) J_n(\ell_1 \theta_{13}) J_n(\ell_2 \theta_{23}). \quad (\text{E.4})$$

The angle ψ denotes the angle spanned by $\boldsymbol{\theta}_1 - \boldsymbol{\theta}_3$ and $\boldsymbol{\theta}_2 - \boldsymbol{\theta}_3$.

The series Eq. (E.3) can be truncated after some n once an acceptable accuracy has been achieved, usually of the order of $n \lesssim 10$ (here we choose $n = 6$). The $B_n(\ell_1, \ell_2)$ are computed once for all n in question on a fine logarithmic grid with $N_g \times N_g$ pixels, spanning all scales of interest for the $\langle \kappa_g \kappa_g \gamma \rangle$ map; we employ the limits $10^1 \leq \ell \leq 10^6$ and $N_g = 100$ pixels. Furthermore, assuming that $B_n(\ell_1, \ell_2)$ is constant within one grid pixel, Eq. (E.4) simplifies to

$$F_n(\theta_{13}, \theta_{23}) \approx \sum_{i,j=1}^{N_g} \frac{B_n(\ell_i, \ell_j)}{(\theta_{13}\theta_{23})^2} \left(\mathcal{J}_n([\ell_i + \Delta\ell_i]\theta_{13}) - \mathcal{J}_n(\ell_i\theta_{13}) \right) \left(\mathcal{J}_n([\ell_j + \Delta\ell_j]\theta_{23}) - \mathcal{J}_n(\ell_j\theta_{23}) \right), \quad (\text{E.5})$$

where $\Delta\ell_i\Delta\ell_j$ is the area of pixel (i, j) in the 2-D ℓ -grid of $B_n(\ell_1, \ell_2)$. The integrated Bessel function of first kind is defined by

$$\mathcal{J}_n(x) \equiv \int_0^x ds s J_n(s), \quad (\text{E.6})$$

which we compute once as lookup-function for the needed range of x .

E.2 Aperture statistics

Similar to the excess mass maps, also for the aperture statistics $\langle \mathcal{N}(\theta_1)\mathcal{N}(\theta_2)M_{\text{ap}}(\theta_3) \rangle$ it is convenient to compute Eq. (3.30) for a given bispectrum $b_{\text{gg}\kappa}$ by utilising the multipole expansion of the bispectrum, Eq. (E.1):

$$\langle \mathcal{N}(\theta_1)\mathcal{N}(\theta_2)M_{\text{ap}}(\theta_3) \rangle = \frac{1}{4\pi^2} \sum_{n=0}^{\infty} (2 - \delta_{n0}^{\text{K}}) \int_0^{\infty} d\ell_1 d\ell_2 \ell_1 \ell_2 B_n(\ell_1, \ell_2) \hat{u}(\ell_1\theta_1) \hat{u}(\ell_2\theta_2) \square_n(\ell_1\theta_3, \ell_2\theta_3), \quad (\text{E.7})$$

where the auxiliary function \square_n is defined as

$$\square_n(\ell_1, \ell_2) = \int_0^{2\pi} \frac{d\phi_1 d\phi_2}{(2\pi)^2} \hat{u} \left(\sqrt{\ell_1^2 + \ell_2^2 + 2\ell_1\ell_2 \cos(\phi_1 - \phi_2)} \right) \cos(n(\phi_1 - \phi_2)), \quad (\text{E.8})$$

and δ_{ij}^{K} is the Kronecker symbol. In this work $\square_n(\ell_1, \ell_2)$ is computed using a Riemann sum.

Since the \square_n are independent of the bispectrum, they can be computed once as lookup-table on a logarithmic ℓ -grid and be reused any time the aperture statistics are computed. Specifically, we compute $\square_n(\ell_1\theta_3, \ell_2\theta_3)$ for a fiducial value $\theta_3^{\text{f}} = 0.5$ arcmin on a logarithmic (ℓ_1, ℓ_2) -grid. If θ_3 is changed relative to its fiducial value, one has to offset the grid indices, both ℓ_1 and ℓ_2 direction, by $\log(\theta_3/\theta_3^{\text{f}})/\Delta \log \ell_{1/2}$, where $\Delta \log \ell$ is the size of a ℓ grid pixel; the pixels are constructed such to be equally sized on a log-scale.

For given coefficients $B_n(\ell_1, \ell_2)$, the multipole expansion of the aperture statistics is now approximated by (a) truncating the multipole series beyond some n , and (b) by

assuming that B_n and \square_n are essentially constant over the size of a ℓ -grid pixel:

$$\begin{aligned} \langle \mathcal{N}(\theta_1) \mathcal{N}(\theta_2) M_{\text{ap}}(\theta_3) \rangle &\approx \frac{1}{4\pi^2} \sum_{n=0}^{\infty} (2 - \delta_{n0}^{\text{K}}) \\ &\times \sum_{i,j=1}^{N_g} \frac{B_n(\ell_i, \ell_j)}{(\theta_1 \theta_2)^2} \square_n(\ell_i \theta_3, \ell_j \theta_3) (\mathcal{U}([\ell_i + \Delta \ell_i] \theta_1) - \mathcal{U}(\ell_i \theta_1)) (\mathcal{U}([\ell_j + \Delta \ell_j] \theta_2) - \mathcal{U}(\ell_j \theta_2)). \end{aligned} \quad (\text{E.9})$$

Here, we integrated the aperture kernel over the grid pixel area:

$$\mathcal{U}(\ell) = \int_0^\ell ds s \hat{u}(s). \quad (\text{E.10})$$

For the exponential aperture kernel employed in [Schneider & Watts \(2005\)](#),

$$\hat{u}(\ell) = \frac{\ell^2}{2} e^{-\ell^2/2}, \quad (\text{E.11})$$

we find

$$\mathcal{U}(\ell) = - (1 + \ell^2/2) e^{-\ell^2/2}. \quad (\text{E.12})$$

Appendix F

Model fits to the mock & CFHTLenS G2L

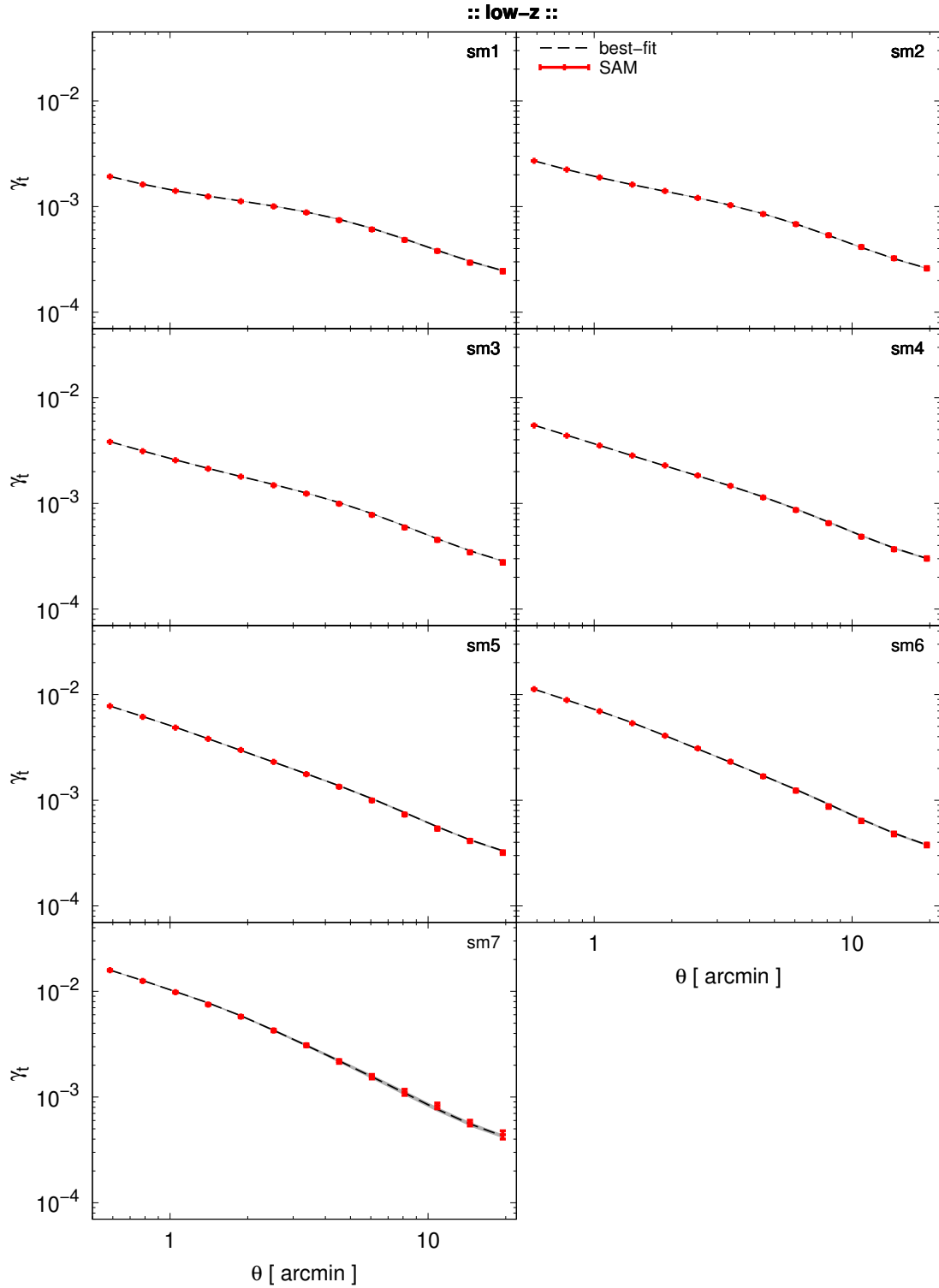


Figure F.1: Best-fit models (dashed line with shaded area representing the standard deviation) to the G2L signals from the Millennium Simulation using the SAM by Henriques+15 (symbols) for all stellar mass samples for low- z .

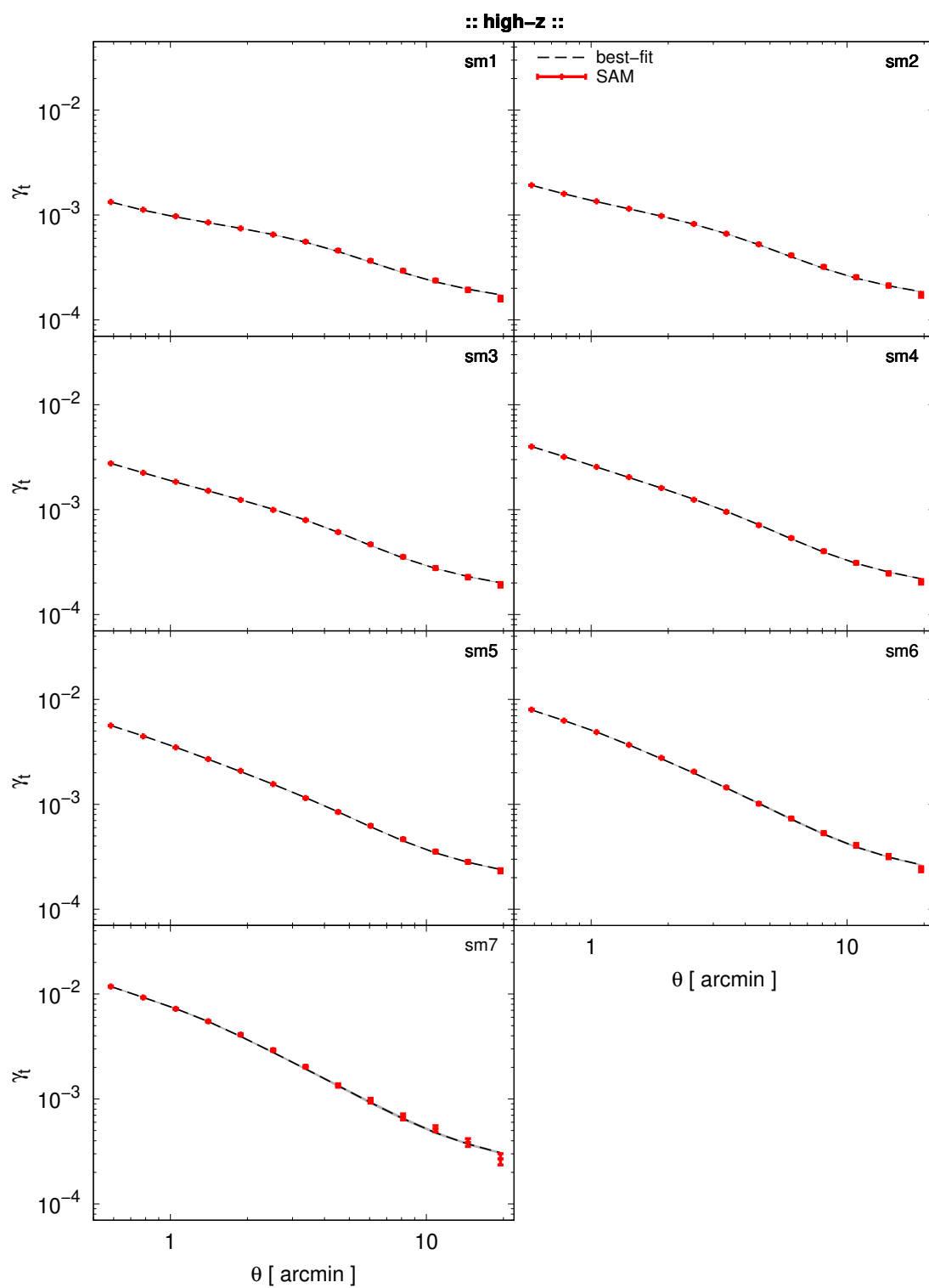


Figure F.2: As Fig. F.1 but for high- z .

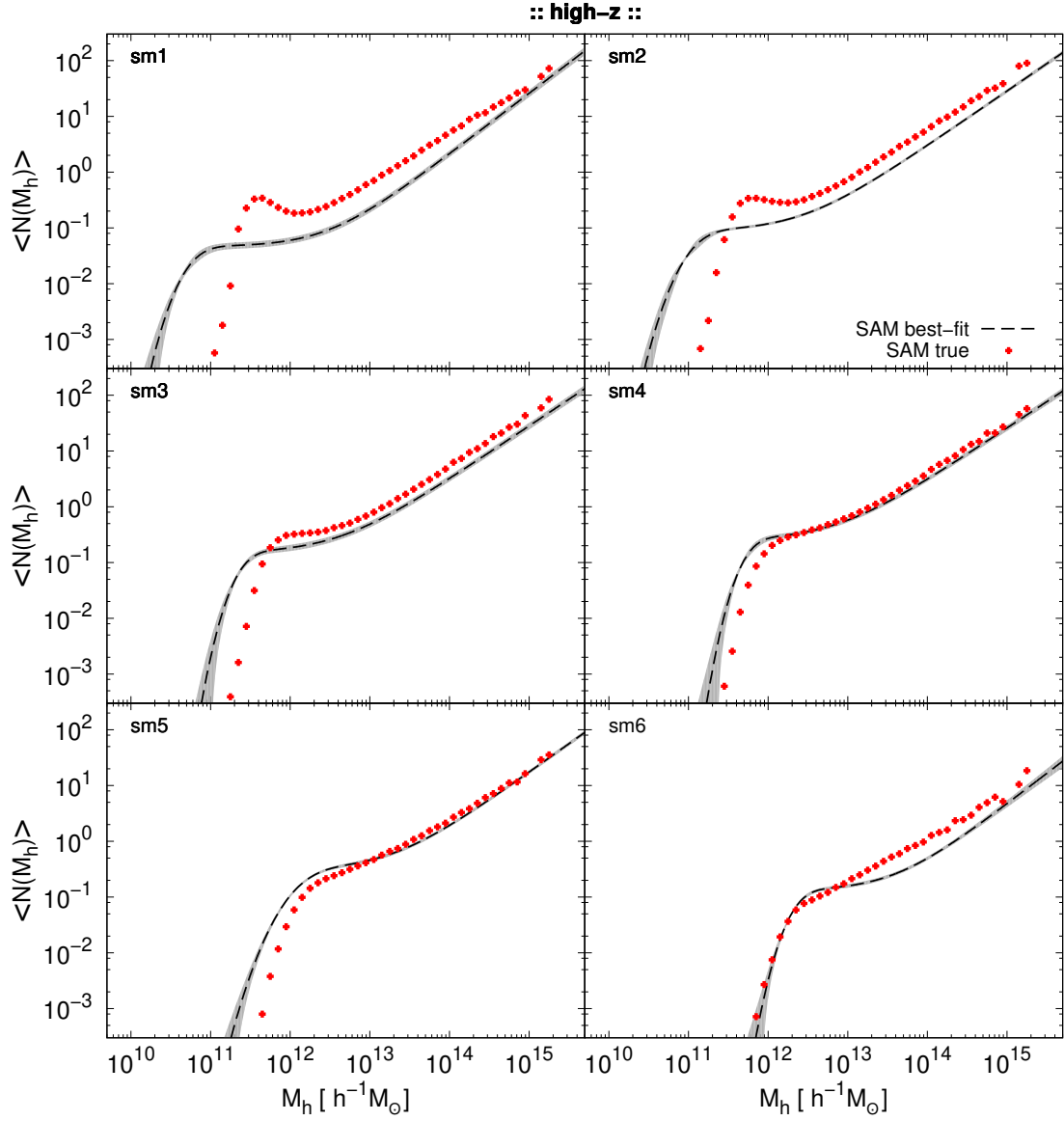


Figure F.3: Best-fit HODs obtained from the models fits to the G2L from the Millennium Simulation using the SAM by Henriques+15 (dashed line with shaded 1σ uncertainty region) in comparison to the ‘true’ HODs (symbols). The results are shown for the various stellar mass samples for high- z .

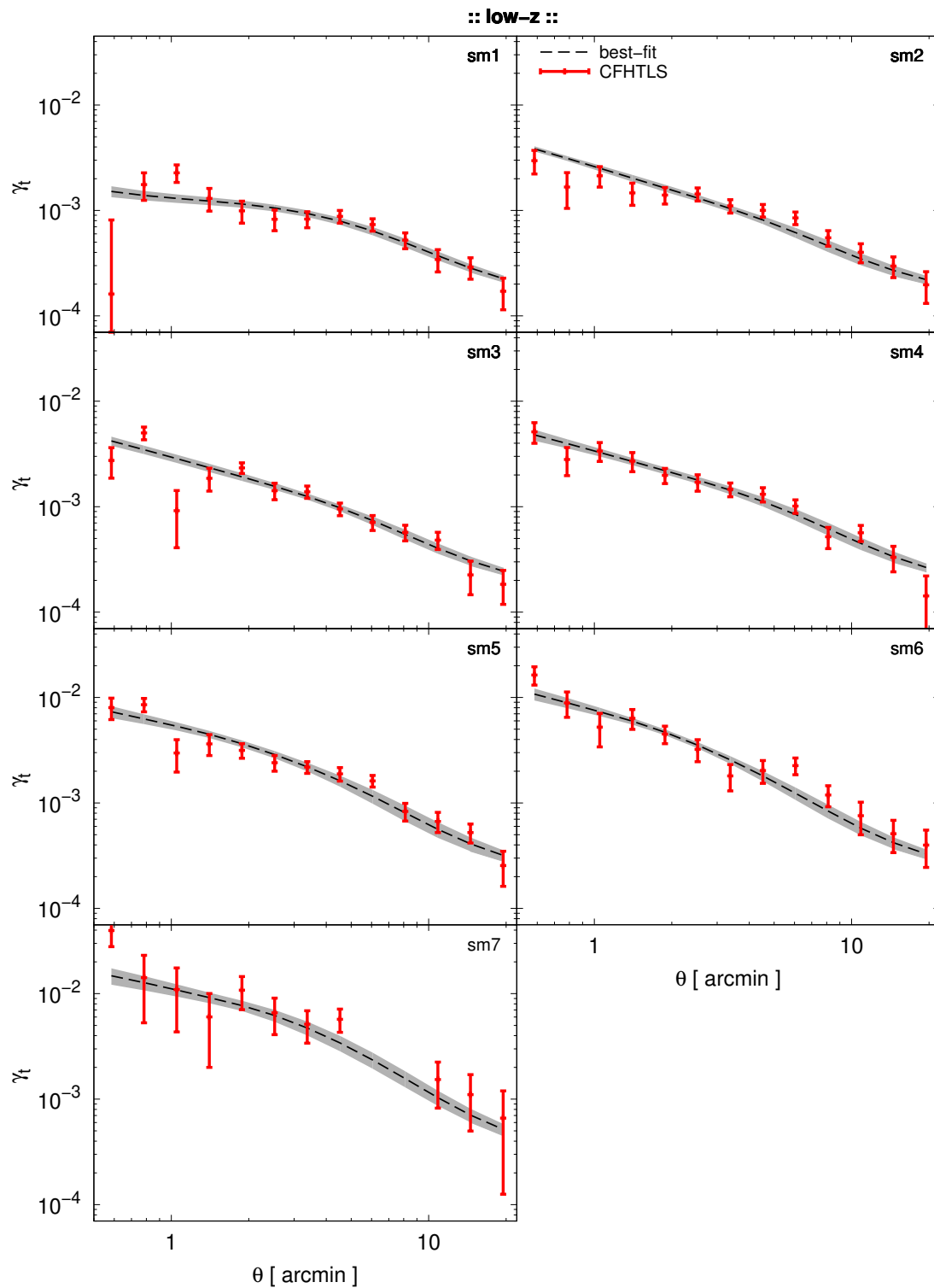
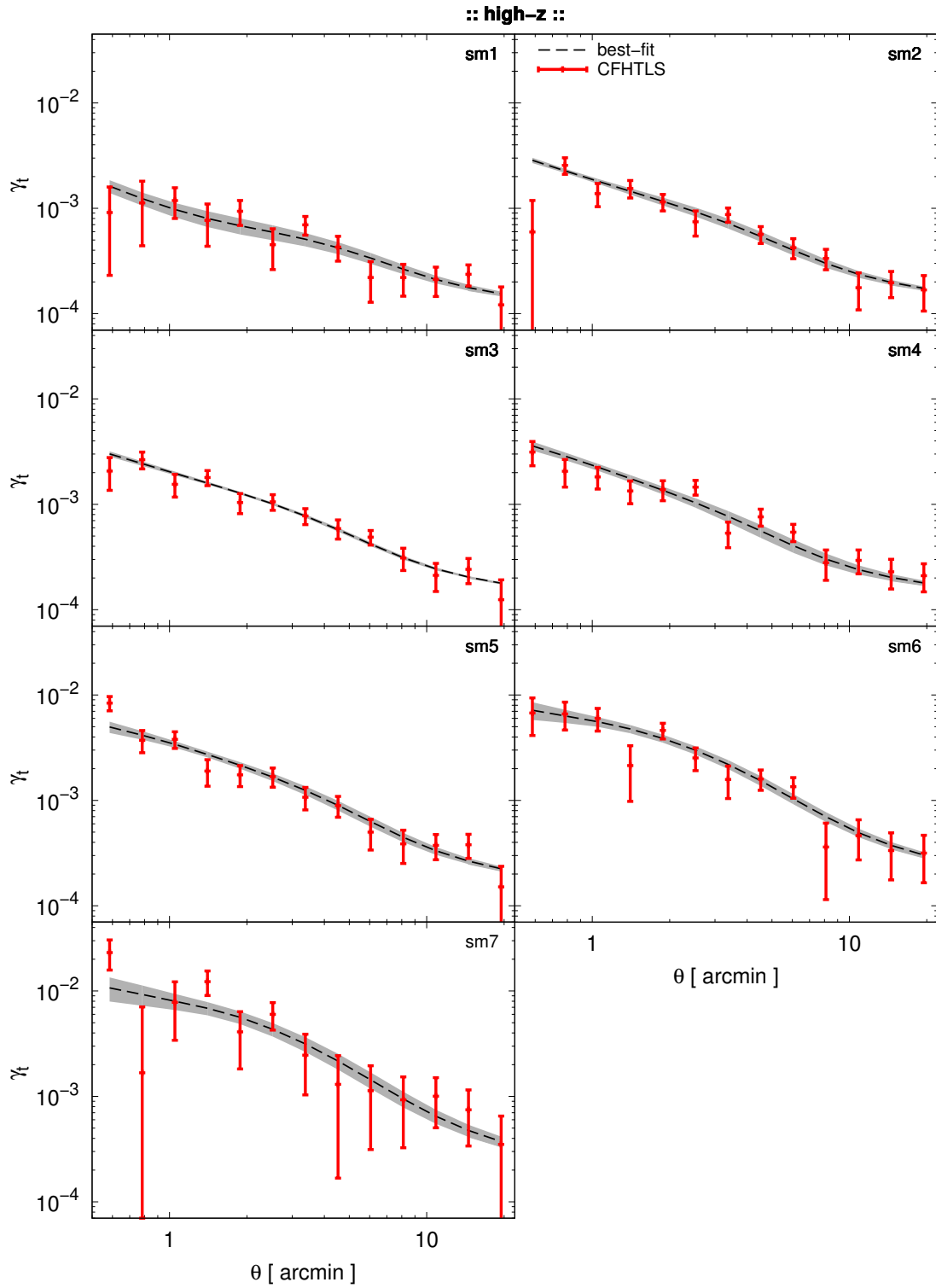


Figure F.4: Best-fit models (dashed line with shaded area representing the standard deviation) to the G2L from CFHTLenS (symbols) for all stellar mass samples for low- z .

Figure F.5: As Fig. F.4 but for high- z .

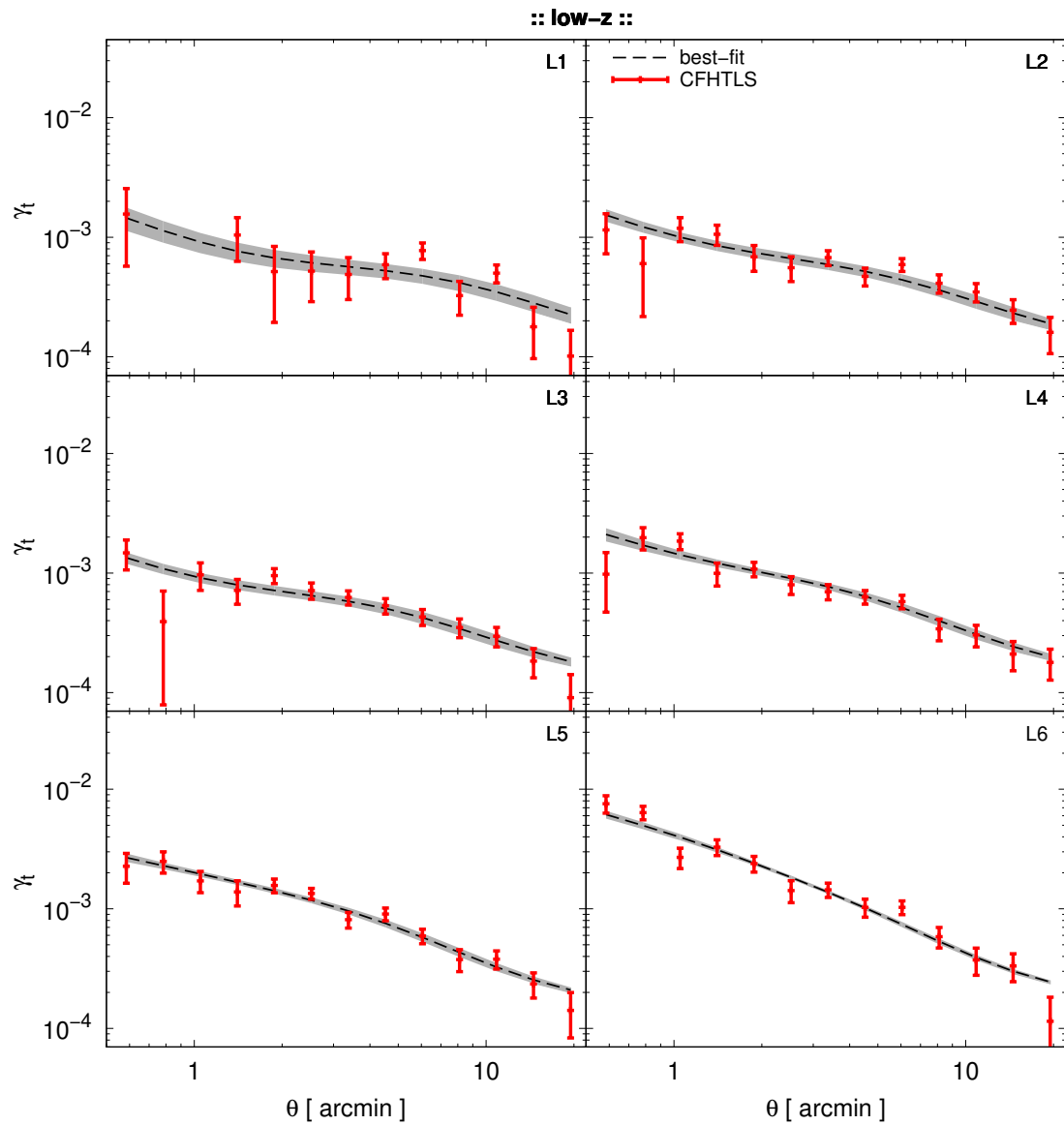


Figure F.6: Best-fit models (dashed line with shaded area representing the standard deviation) to the G2L from CFHTLenS (symbols) for all luminosity samples for low- z .

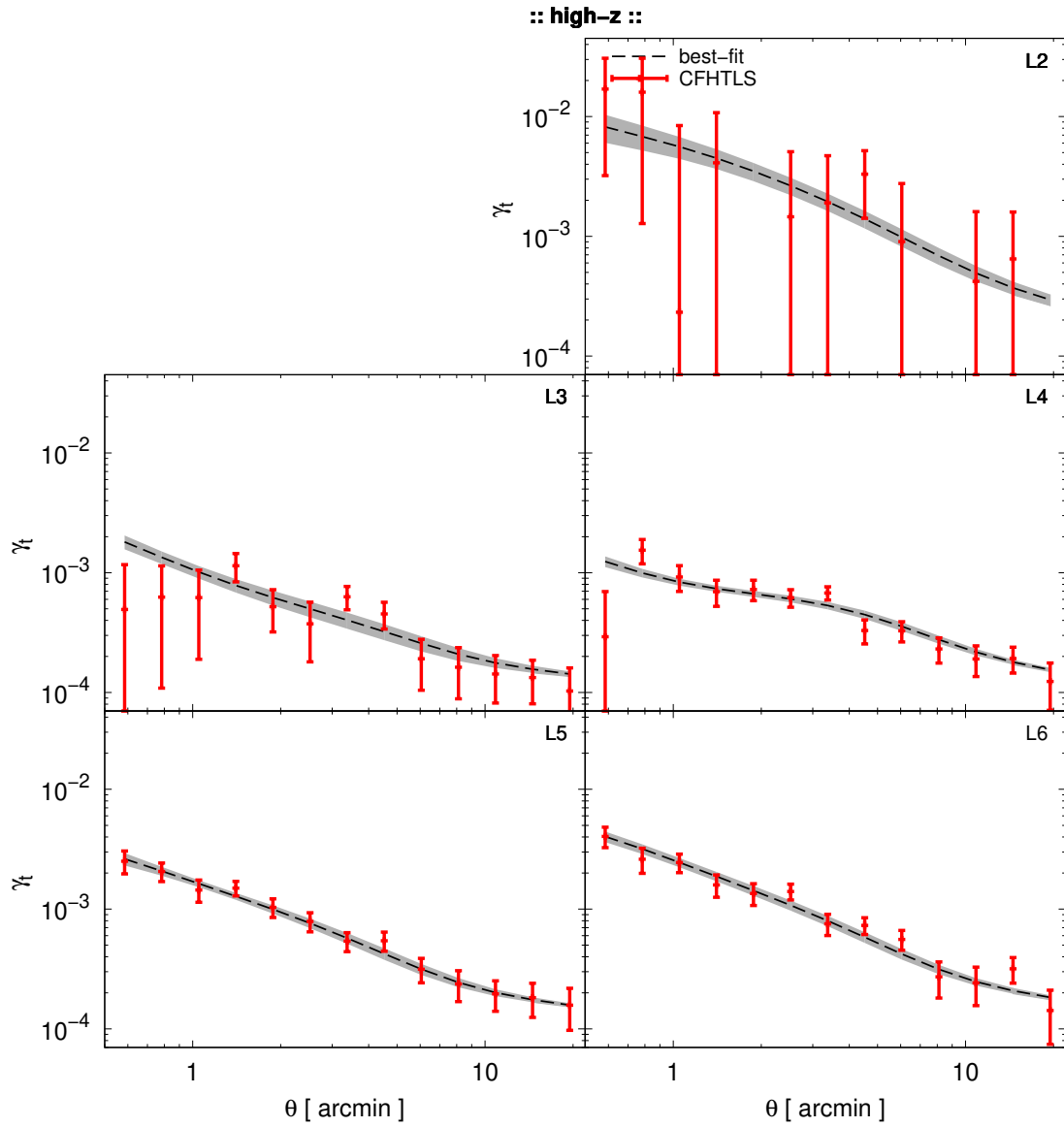


Figure F.7: As Fig. F.6 but for high- z .

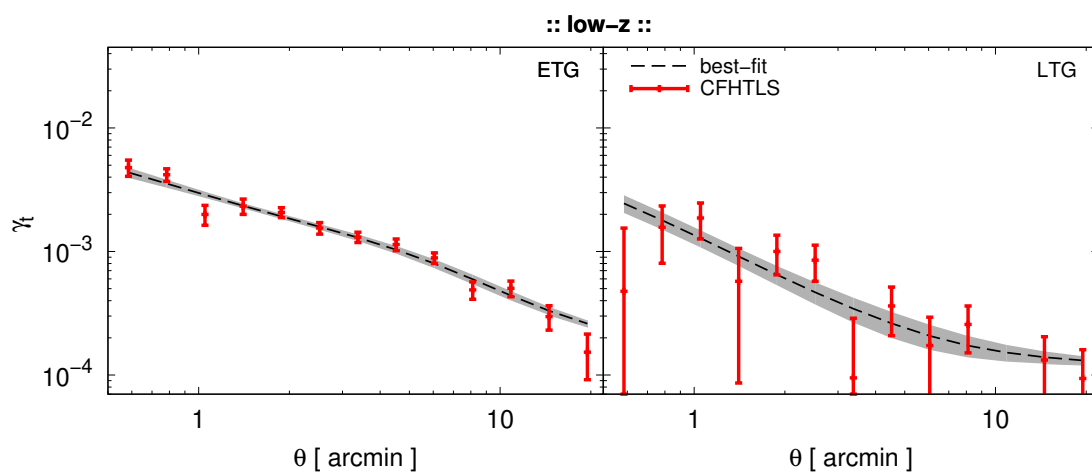


Figure F.8: Best-fit models (dashed line with shaded area representing the standard deviation) to the G2L from CFHTLenS (symbols) for the galaxy-type samples for low- z .

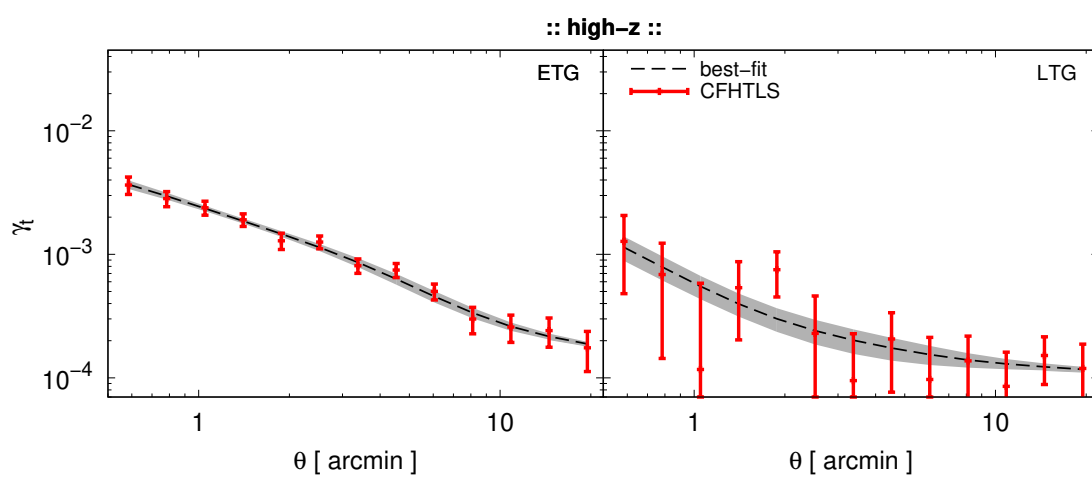


Figure F.9: As Fig. F.8 but for high- z .

Appendix G

First-time halo model prediction of excess mass

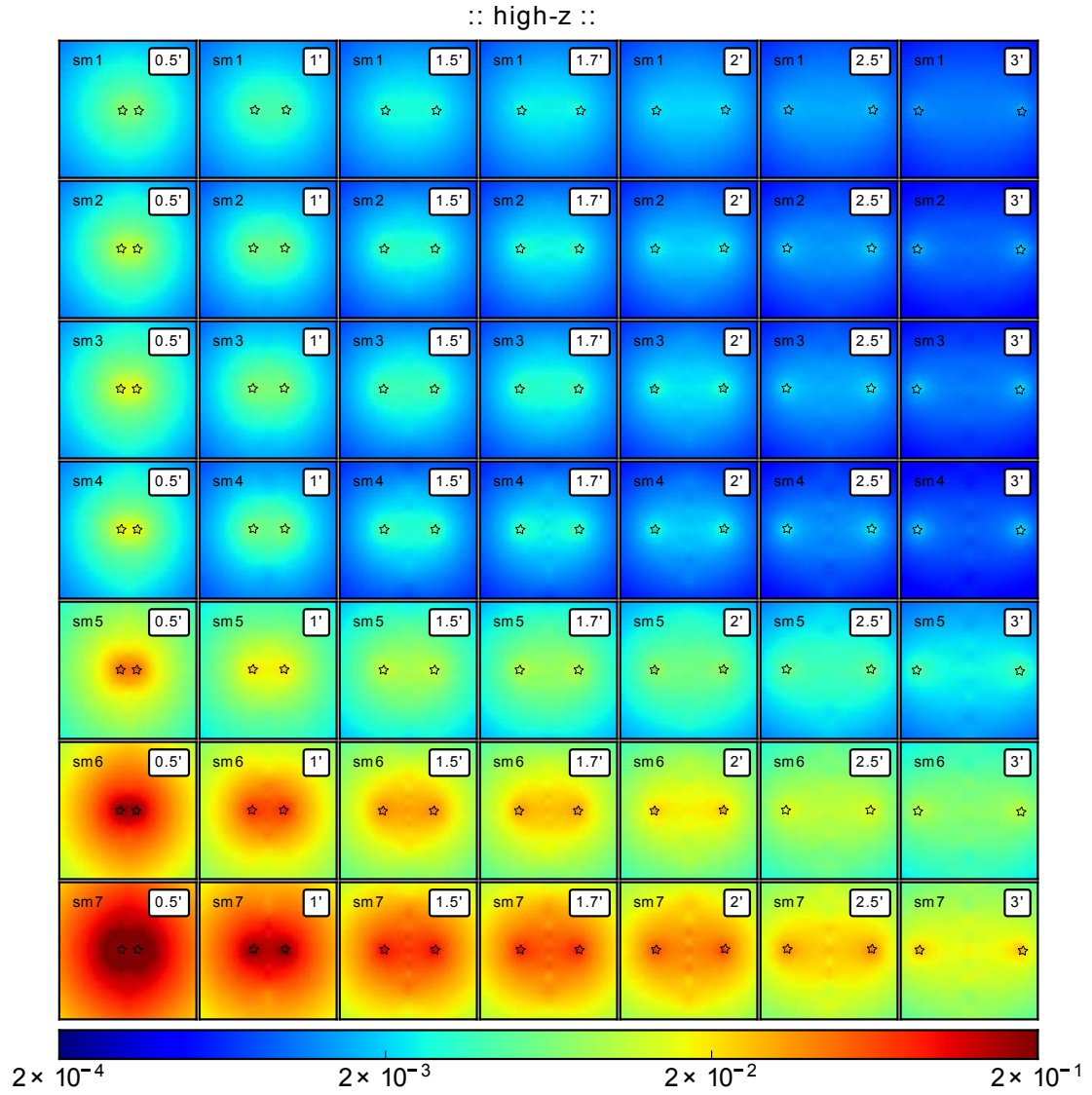


Figure G.1: Excess mass map predictions for seven stellar mass samples (*top to bottom*) for lens-lens separations ranging from 0.5 arcmin to 3 arcmin (*left to right*). Shown are the high- z results. The lens positions are indicated by stars. The size of each map is 4×4 arcmin².

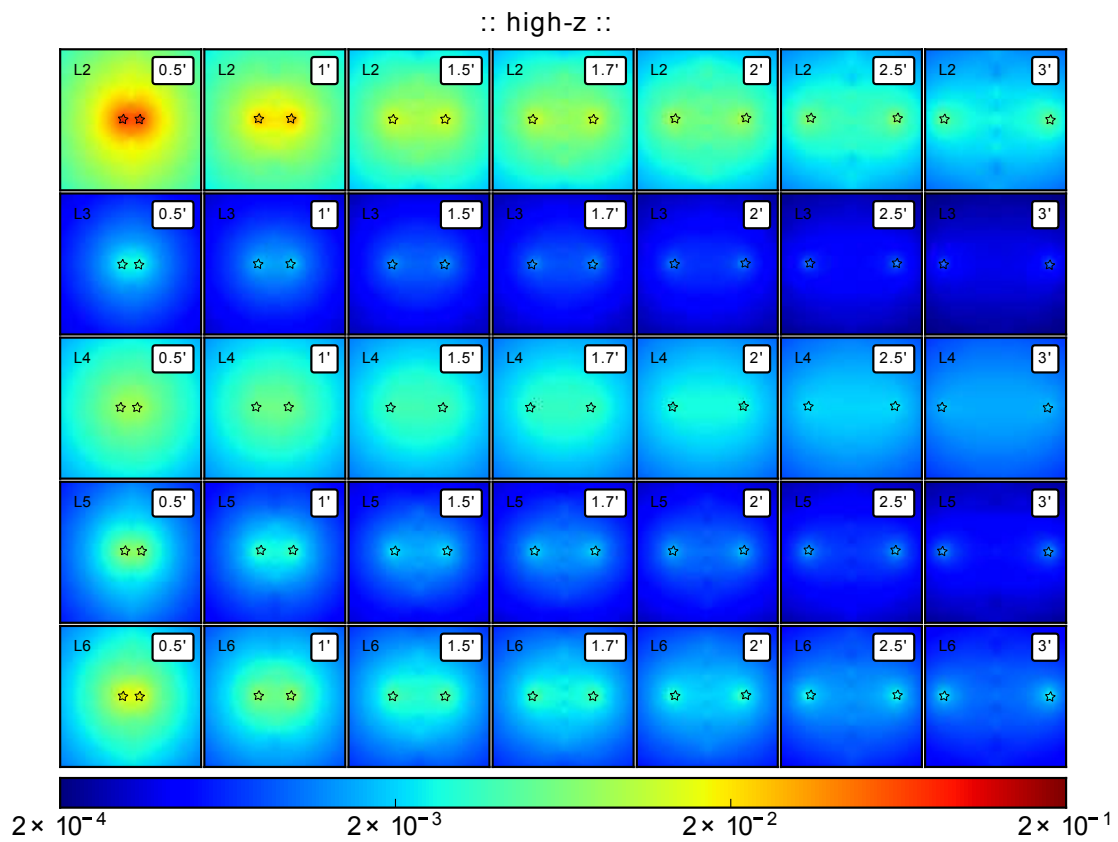


Figure G.2: Excess mass map predictions for five luminosity samples (*top to bottom*) for various lens-lens separations. Shown are the high- z results. The lens positions are indicated by stars. The size of each map is 4×4 arcmin².

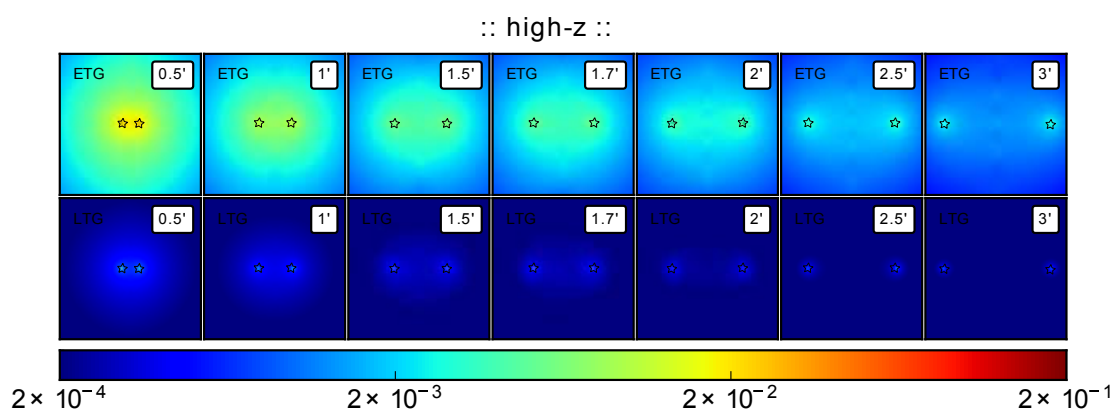


Figure G.3: Excess mass map predictions for the early-type (*top*) and the late-type galaxy (*bottom*) sample for various lens-lens separations. Shown are the high- z results. The lens positions are indicated by stars. The size of each map is 4×4 arcmin².

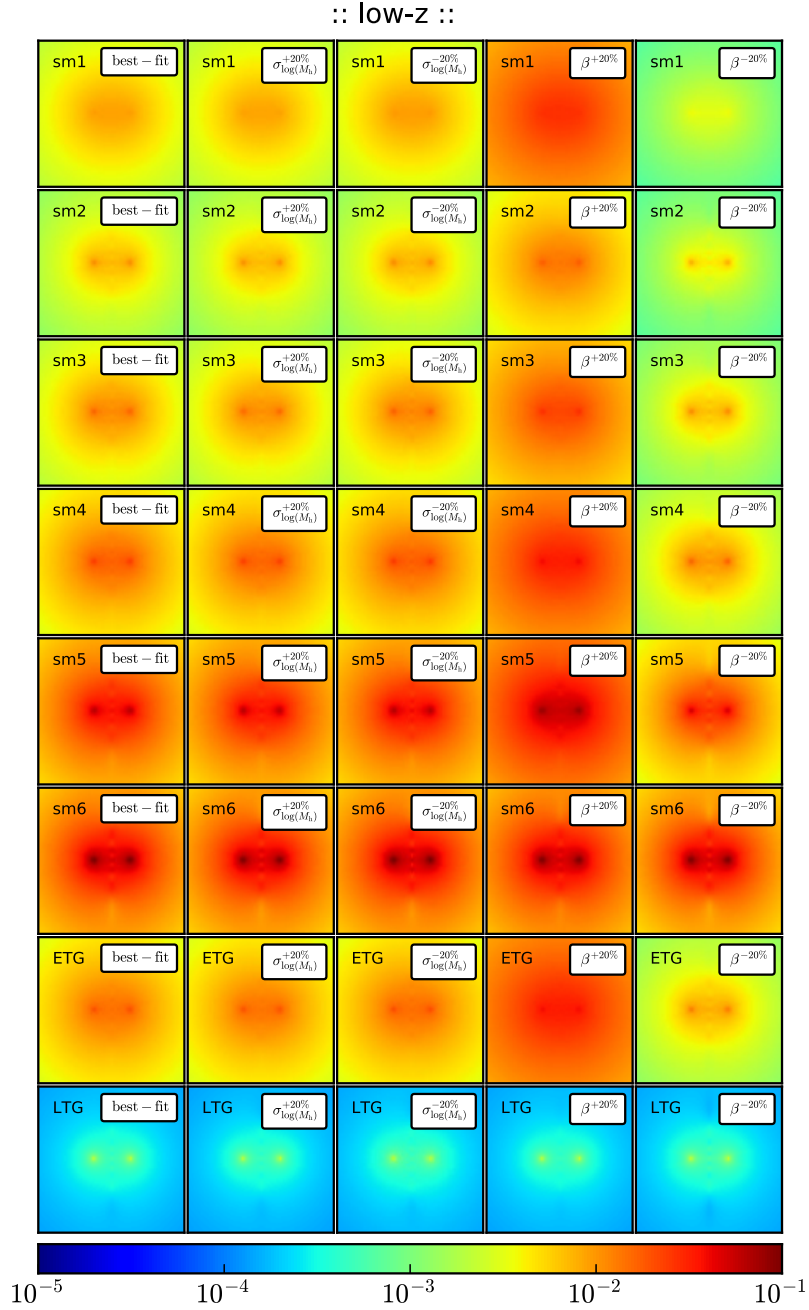


Figure G.4: Dependence of excess mass map predictions on varying the HOD parameters β and $\sigma_{\log(M_h)}$ individually by $\pm 20\%$ with respect to their best-fit values. Shown are the results for different stellar mass and galaxy-type samples (*top to bottom*) for a lens-lens separation of 1 arcmin and low- z . The size of each map is $4 \times 4 \text{ arcmin}^2$.

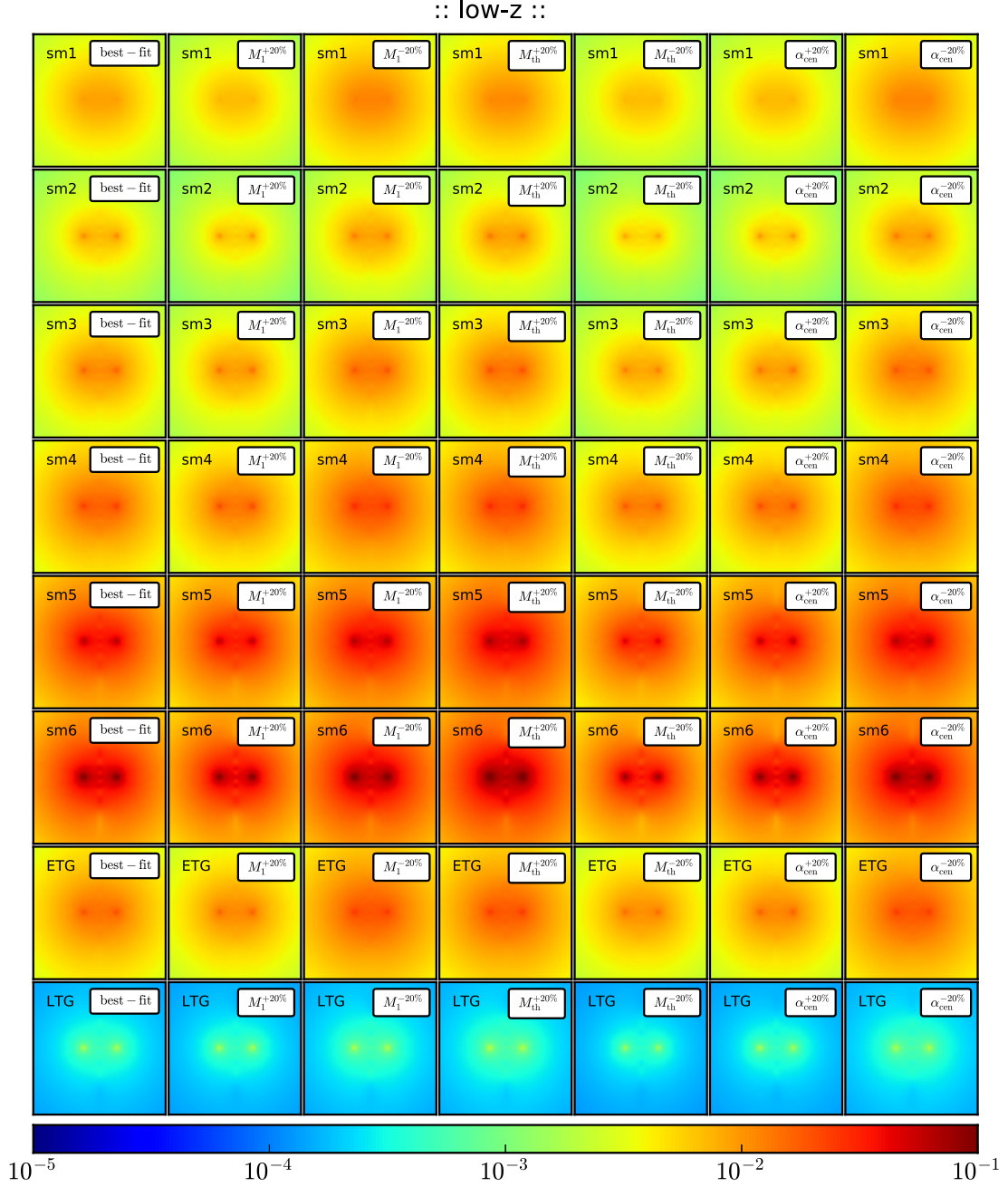


Figure G.5: Dependence of excess mass map predictions on varying the HOD parameters M_1 , M_{th} and α_{cen} individually by $\pm 20\%$ with respect to their best-fit values. Shown are the results for different stellar mass and galaxy-type samples (*top to bottom*) for a lens-lens separation of 1 arcmin and low- z . The size of each map is 4×4 arcmin².

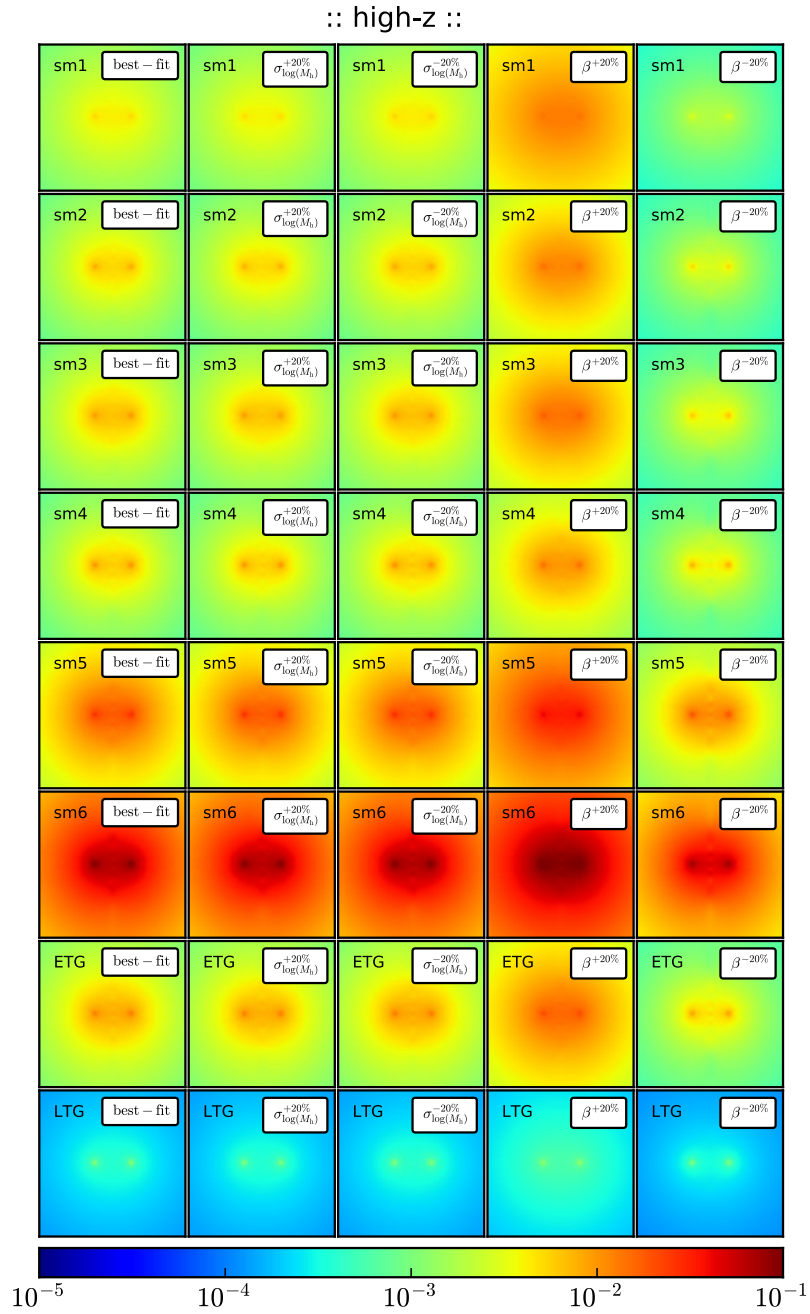


Figure G.6: As Fig. G.4 but for high- z .

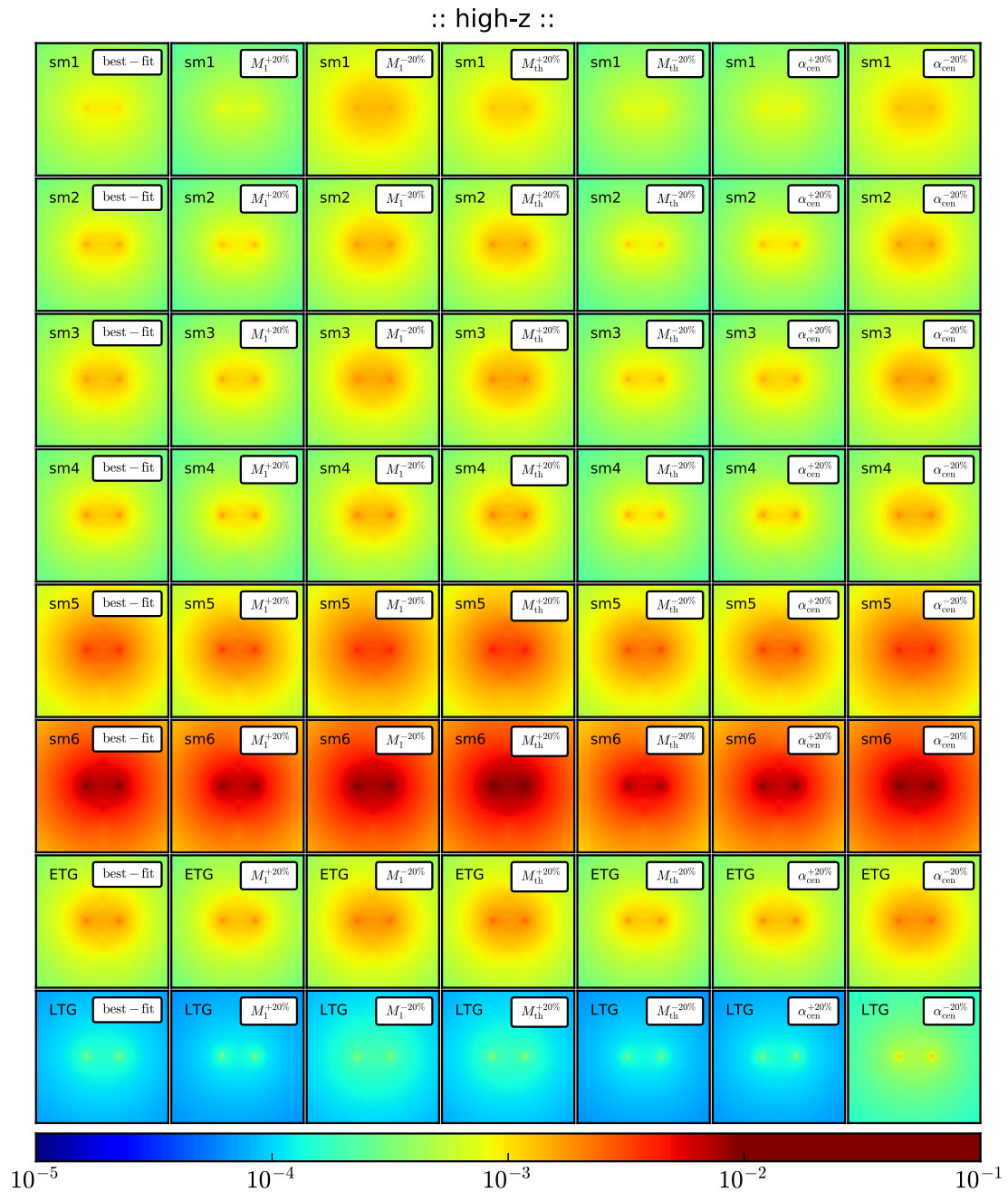


Figure G.7: As Fig. G.5 but for high- z .

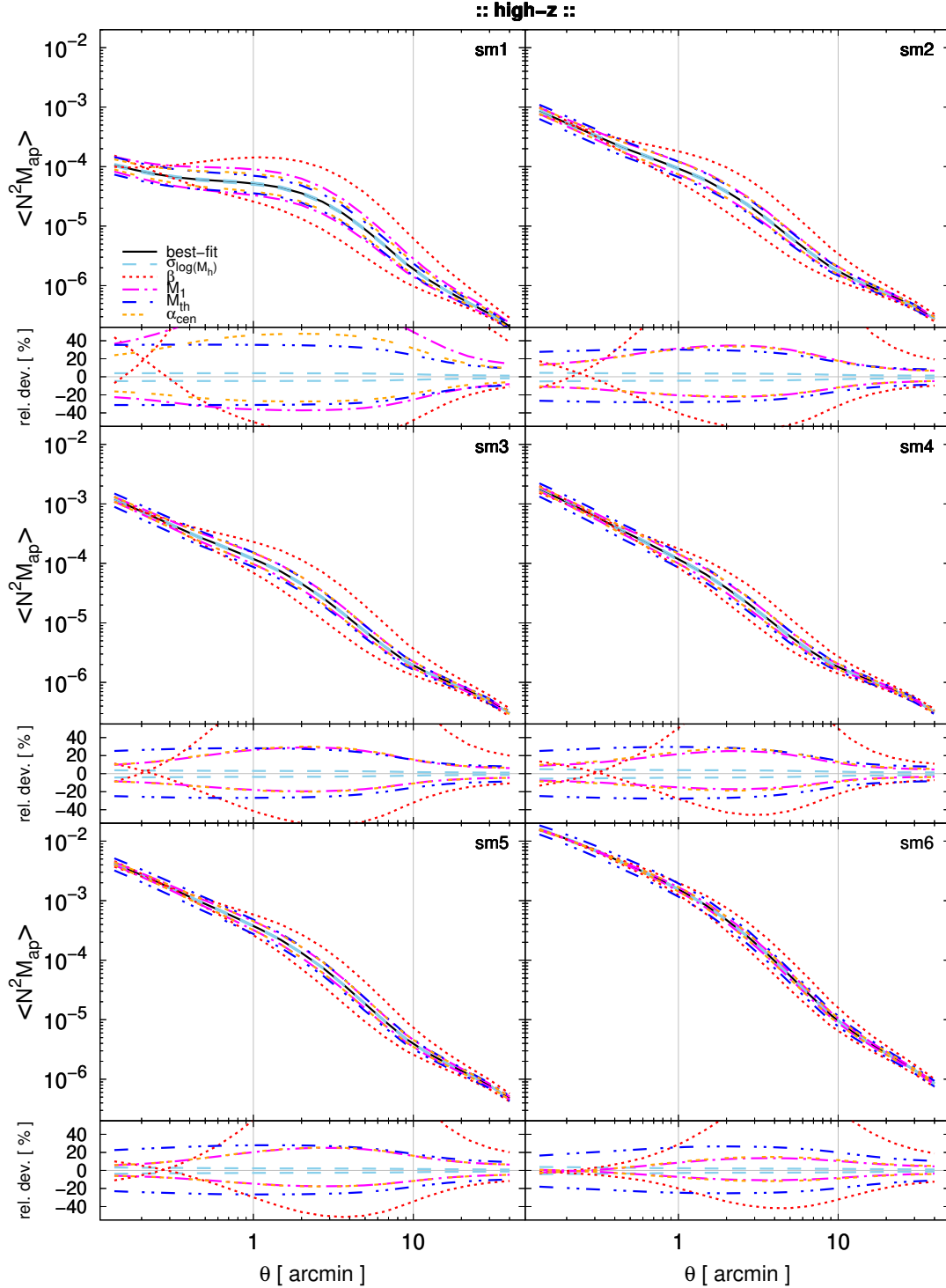


Figure G.8: Sensitivity of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle(\theta)$ on variations in the five HOD parameters by $\pm 20\%$ with respect to their respective best-fit values. The main plots show the absolute predictions for the six stellar mass samples as a function of aperture scale radius, the minor plots the relative deviation from the best-fit reference in percent. Shown are the predictions for high- z .

Appendix H

Testing G3L halo model predictions

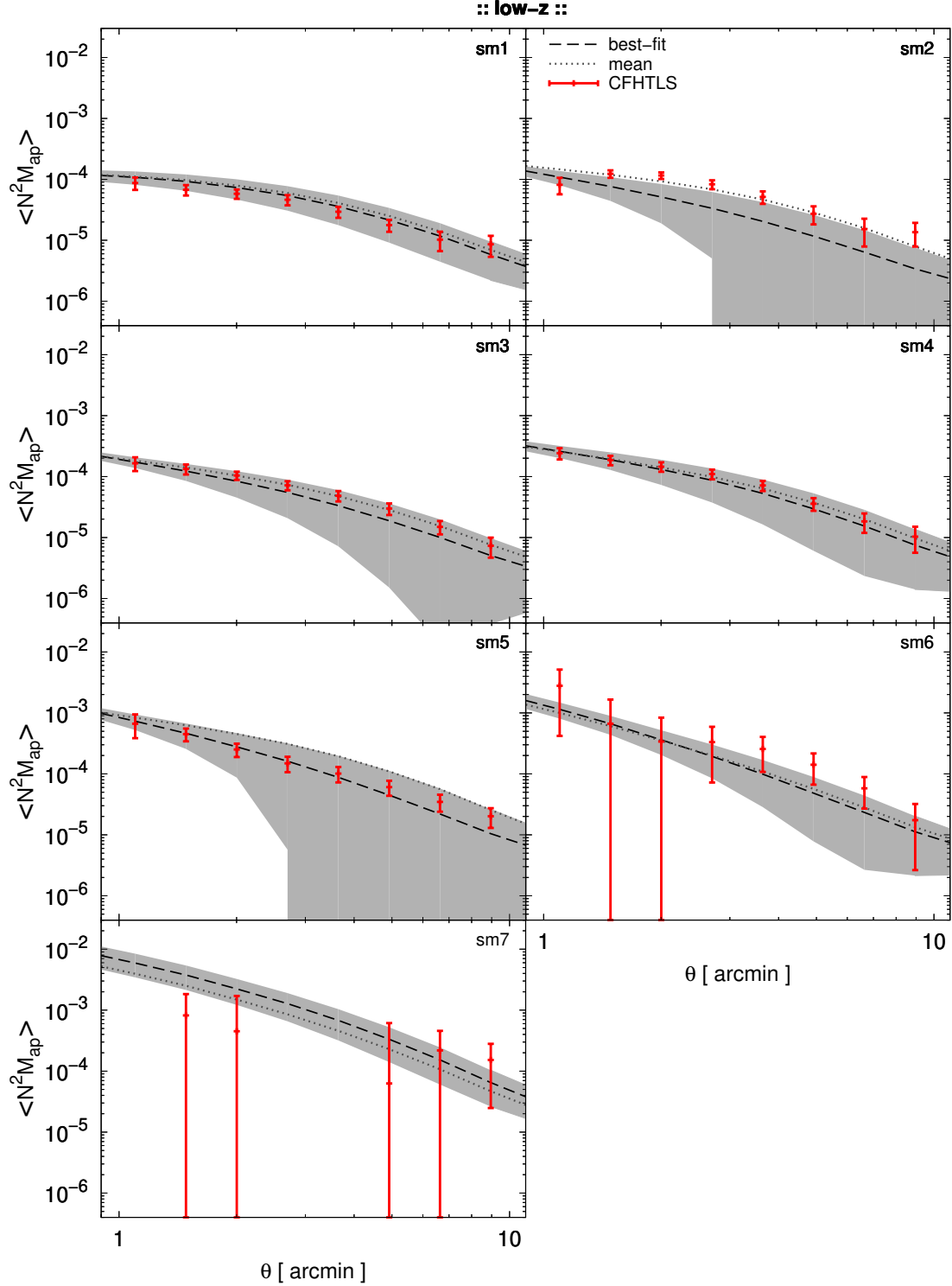


Figure H.1: Comparison of the $\langle N^2 M_{\text{ap}} \rangle(\theta)$ halo model predictions (dashed line) to the CFHTLenS measurements (symbols) as a function of aperture scale radius θ . Shown are the results for the stellar mass samples and low- z . The error bars and the error bands give the $1\text{-}\sigma$ uncertainty. The mean over the random realisations is given by the dotted line.

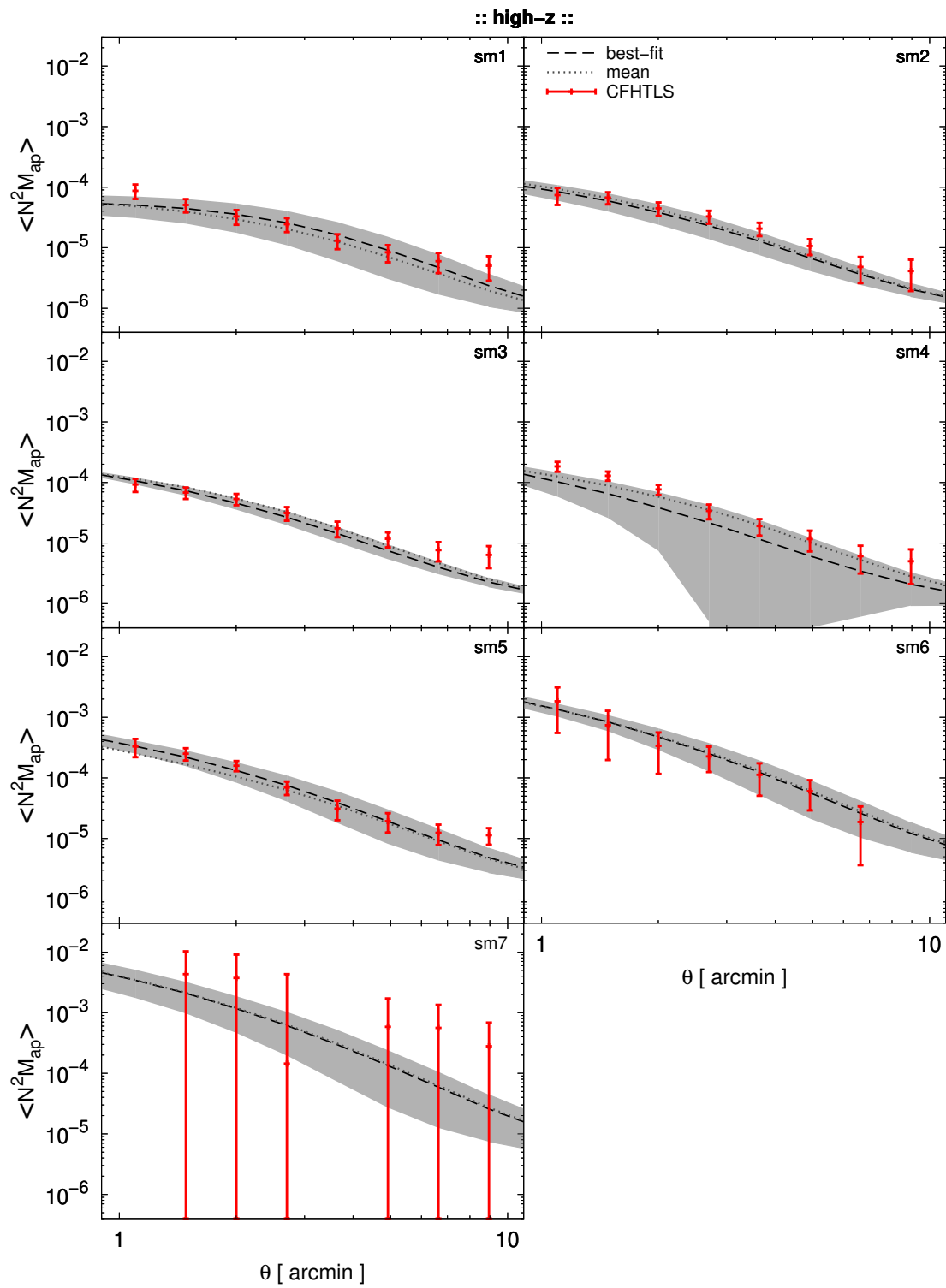


Figure H.2: As Fig. H.1 but for high- z .

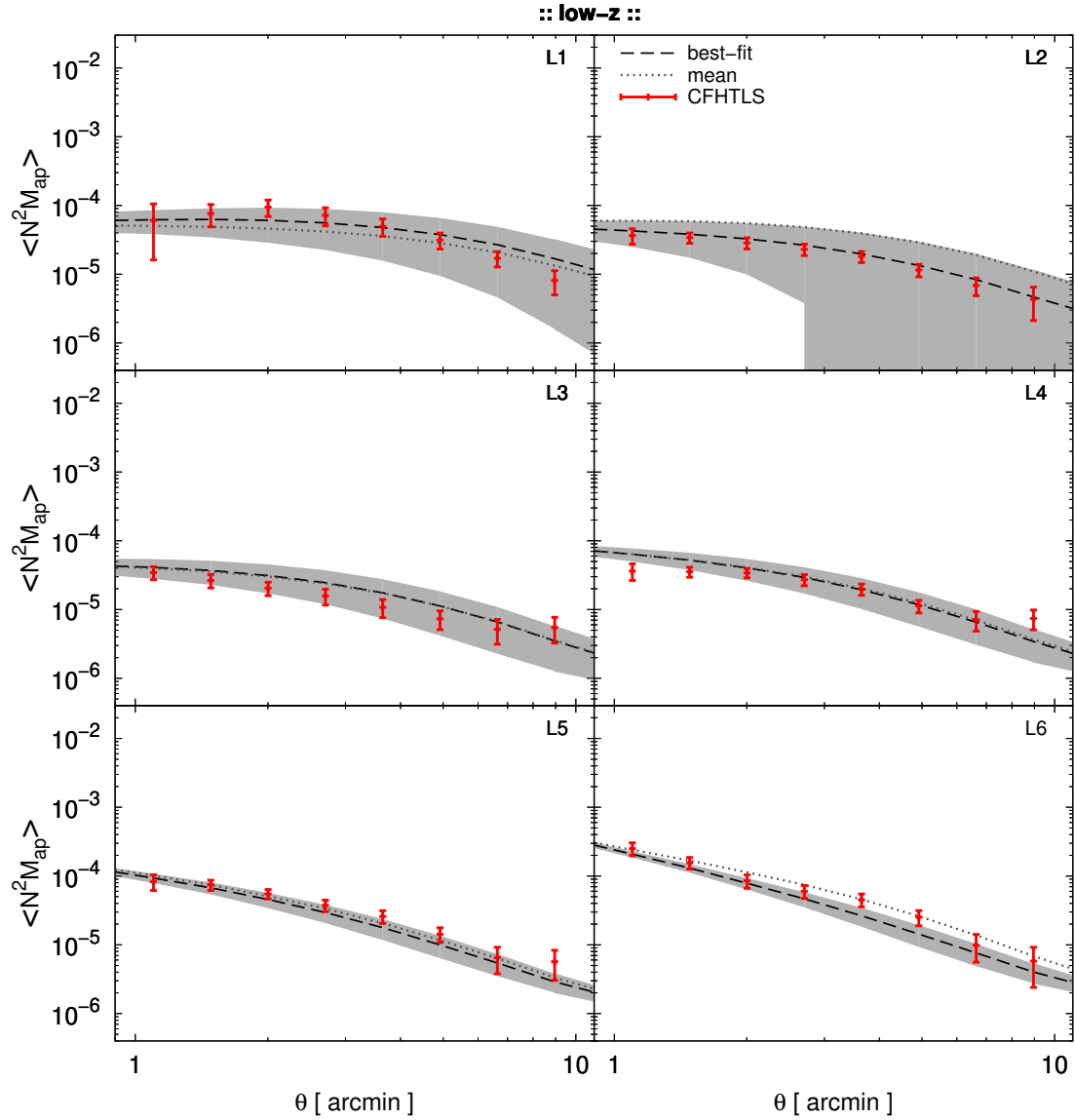


Figure H.3: Comparison of the $\langle N^2 M_{\text{ap}} \rangle(\theta)$ halo model predictions (dashed line) to the CFHTLenS measurements (symbols) as a function of aperture scale radius θ . Shown are the results for the luminosity samples and low- z . The error bars and the error bands give the 1- σ uncertainty. The mean over the random realisations is given by the dotted line.

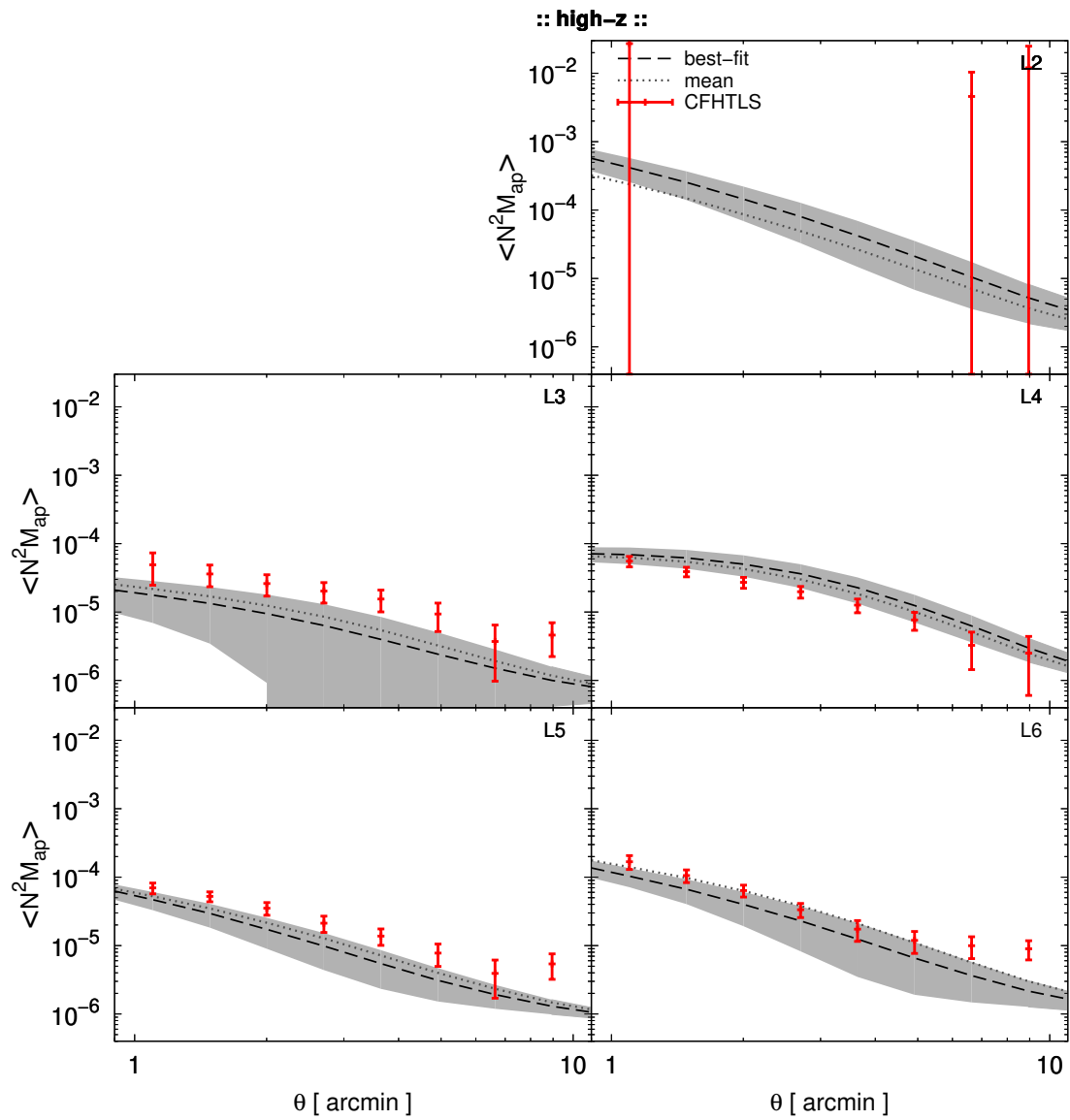


Figure H.4: As Fig. H.3 but for high- z .

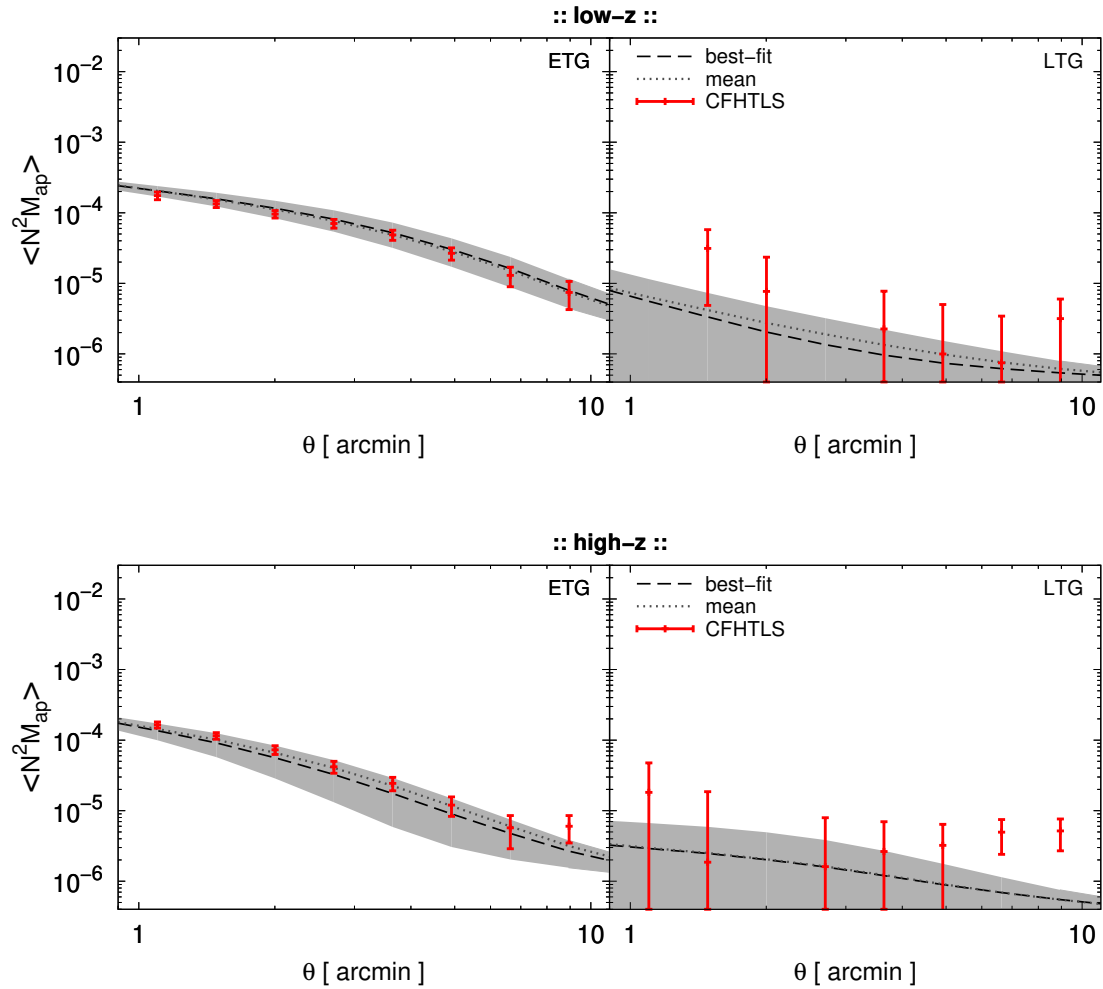


Figure H.5: Comparison of the $\langle N^2 M_{\text{ap}} \rangle(\theta)$ halo model predictions (dashed line) to the CFHTLenS measurements (symbols) as a function of aperture scale radius θ . Shown are the results for the galaxy-type samples, for low- z (*top*) and for high- z (*bottom*). The error bars and the error bands give the $1\text{-}\sigma$ uncertainty. The mean over the random realisations is given as dotted line.

Bibliography

- Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, *ARA&A*, 49, 409
- Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, *ApJ*, 304, 15
- Barlow, R. 1989, *Statistics. A guide to the use of statistical methods in the physical sciences* (John Wiley & Sons)
- Bartelmann, M. & Schneider, P. 2001, *Phys. Rep.*, 340, 291
- Benítez, N. 2000, *ApJ*, 536, 571
- Berlind, A. A. & Weinberg, D. H. 2002, *ApJ*, 575, 587
- Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *Phys. Rep.*, 367, 1
- Bertone, G., Hooper, D., & Silk, J. 2005, *Phys. Rep.*, 405, 279
- BICEP2/Keck and Planck Collaborations, Ade, P. A. R., Aghanim, N., et al. 2015, *Physical Review Letters*, 114, 101301
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, *ApJ*, 379, 440
- Bond, J. R. & Efstathiou, G. 1984, *ApJ*, 285, L45
- Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, *MNRAS*, 370, 645
- Buddendiek, A., Schrabback, T., Greer, C. H., et al. 2015, *MNRAS*, 450, 4248
- Budzynski, J. M., Kogosov, S. E., McCarthy, I. G., McGee, S. L., & Belokurov, V. 2012, *MNRAS*, 423, 104
- Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, *MNRAS*, 321, 559
- Carroll, S. M., Press, W. H., & Turner, E. L. 1992, *ARA&A*, 30, 499
- Clampitt, J., Sánchez, C., Kwan, J., et al. 2017, *MNRAS*, 465, 4204
- Colless, M. 1999, *Philosophical Transactions of the Royal Society of London Series A*, 357, 105
- Cooray, A. & Sheth, R. 2002, *Phys. Rep.*, 372, 1

- Coupon, J., Arnouts, S., van Waerbeke, L., et al. 2015, *MNRAS*, 449, 1352
- Coupon, J., Kilbinger, M., McCracken, H. J., et al. 2012, *A&A*, 542, A5
- Cramér, H. 1946, *Mathematical Methods of Statistics* (Princeton University Press)
- Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, *ApJ*, 568, 20
- Dodelson, S. 2003, *Modern cosmology* (Academic Press)
- Dressler, A. 1980, *ApJ*, 236, 351
- Efron, B. & Stein, C. 1981, *Ann. Statist.*, 9, 586
- Einstein, A. 1905, *Annalen der Physik*, 322, 891
- Einstein, A. 1915, *Feldgleichungen der Gravitation* (Preussische Akademie der Wissenschaften, Sitzungsberichte)
- Eisenstein, D. J. & Hu, W. 1998, *ApJ*, 496, 605
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560
- Erben, T., Hildebrandt, H., Miller, L., et al. 2013, *MNRAS*, 433, 2545
- Fixsen, D. J. 2009, *ApJ*, 707, 916
- Frieman, J. A., Turner, M. S., & Huterer, D. 2008, *ARA&A*, 46, 385
- Fu, L., Kilbinger, M., Erben, T., et al. 2014, *MNRAS*, 441, 2725
- Gao, L. & White, S. D. M. 2007, *MNRAS*, 377, L5
- Gunn, J. E. & Gott, III, J. R. 1972, *ApJ*, 176, 1
- Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, *MNRAS*, 413, 101
- Guth, A. H. 1981, *Phys. Rev. D*, 23, 347
- Harrison, E. R. 1970, *Phys. Rev. D*, 1, 2726
- Hartlap, J., Simon, P., & Schneider, P. 2007, *A&A*, 464, 399
- Heisenberg, W. 1927, *Zeitschrift für Physik*, 43, 172
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, *MNRAS*, 451, 2663
- Henry, J. P. 2000, *ApJS*, 534, 565
- Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, *MNRAS*, 427, 146
- Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, *A&A*, 499, 31

- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJ*, 208, 19
- Hoekstra, H., van Waerbeke, L., Gladders, M. D., Mellier, Y., & Yee, H. K. C. 2002, *ApJ*, 577, 604
- Hubble, E. 1929, *Proceedings of the National Academy of Science*, 15, 168
- Hudson, M. J., Gillis, B. R., Coupon, J., et al. 2015, *MNRAS*, 447, 298
- Jain, B. & Bertschinger, E. 1994, *ApJ*, 431, 495
- James, F. & Roos, M. 1975, *Computer Physics Communications*, 10, 343
- Jeong, D. & Komatsu, E. 2009, *ApJ*, 703, 1230
- Kaiser, N. 1995, *ApJ*, 439, L1
- Kaiser, N. 1998, *ApJ*, 498, 26
- Kaiser, N. & Squires, G. 1993, *ApJ*, 404, 441
- Knight, K. 1999, *Mathematical Statistics* (Chapman & Hall/CRC Texts in Statistical Science CRC Press)
- Kovac, J. 2018, in *COSPAR Meeting*, Vol. 42, 42nd COSPAR Scientific Assembly, E1.2–24–18
- Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, *ApJ*, 609, 35
- Lazanu, A., Giannantonio, T., Schmittfull, M., & Shellard, E. P. S. 2017, *Phys. Rev. D*, 95, 083511
- Leauthaud, A., Tinker, J., Behroozi, P. S., Busha, M. T., & Wechsler, R. H. 2011, *ApJ*, 738, 45
- Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, *ApJ*, 744, 159
- Limber, D. N. 1953, *ApJ*, 117, 134
- Mandelbaum, R., Seljak, U., Cool, R. J., et al. 2006a, *MNRAS*, 372, 758
- Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006b, *MNRAS*, 368, 715
- Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006c, *MNRAS*, 368, 715
- Mandelbaum, R., Tasitsiomi, A., Seljak, U., Kravtsov, A. V., & Wechsler, R. H. 2005, *MNRAS*, 362, 1451
- Mao, Y.-Y., Zentner, A. R., & Wechsler, R. H. 2018, *MNRAS*, 474, 5143

- Martin, J., Ringeval, C., & Vennin, V. 2014, *Physics of the Dark Universe*, 5, 75
- Mead, A. J., Peacock, J. A., Heymans, C., Joudaki, S., & Heavens, A. F. 2015, *MNRAS*, 454, 1958
- Mo, H., van den Bosch, F. C., & White, S. 2010, *Galaxy Formation and Evolution*
- Mo, H. J., Jing, Y. P., & White, S. D. M. 1997, *MNRAS*, 284, 189
- Mo, H. J. & White, S. D. M. 1996, *MNRAS*, 282, 347
- Montero-Dorta, A. D., Pérez, E., Prada, F., et al. 2017, *ApJ*, 848, L2
- Muzzin, A., Yee, H. K. C., Hall, P. B., Ellingson, E., & Lin, H. 2007, *ApJ*, 659, 1106
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, 462, 563
- Newton, I. 1687, *Philosophiae naturalis principia mathematica* (William Dawson and Sons Ltd., London)
- Neyman, J. & Scott, E. L. 1952, *ApJ*, 116, 144
- Neyman, J., Scott, E. L., & Shane, C. D. 1953, *ApJ*, 117, 92
- Peacock, J. 1999, *Cosmological Physics* (Cambridge University Press)
- Penzias, A. A. & Wilson, R. W. 1965, *ApJ*, 142, 419
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, 517, 565
- Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016a, *A&A*, 594, A1
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, *A&A*, 594, A16
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016c, *A&A*, 594, A17
- Press, W. H. & Schechter, P. 1974, *ApJ*, 187, 425
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, *Numerical Recipes 3rd Edition: The Art of Scientific Computing 3* (Cambridge University Press)
- Rao, C. 1965, *Linear Statistical Inference and Its Applications* (John Wiley & Sons)
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, 116, 1009
- Robertson, H. P. 1935, *ApJ*, 82, 284
- Robertson, H. P. 1936a, *ApJ*, 83, 187
- Robertson, H. P. 1936b, *ApJ*, 83, 257

- Rödiger, J. 2009, PhD thesis, Studying Galaxy-Galaxy Lensing and Higher-Order Galaxy-Mass Correlations Using the Halo Model, Rheinische Friedrich-Wilhelms Universität Bonn
- Roszkowski, L., Sessolo, E. M., & Trojanowski, S. 2018, Reports on Progress in Physics, 81, 066201
- Sachs, R. K. & Wolfe, A. M. 1967, ApJ, 147, 73
- Saghiha, H. 2017, PhD thesis, Comparing galaxy-galaxy(-galaxy) lensing in semi-analytic models and observations to study galaxy evolution, Rheinische Friedrich-Wilhelms Universität Bonn
- Saghiha, H., Hilbert, S., Schneider, P., & Simon, P. 2012, A&A, 547, A77
- Saghiha, H., Simon, P., Schneider, P., & Hilbert, S. 2017, A&A, 601, A98
- Saglia, R. P., Bertin, G., Bertola, F., et al. 1993, ApJ, 403, 567
- Sakharov, A. D. 1966, Soviet Journal of Experimental and Theoretical Physics, 22, 241
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
- Schneider, P. 2003, A&A, 408, 829
- Schneider, P. 2006, Saas-Fee Advanced Courses, Vol. 33, Gravitational Lensing: Strong, Weak and Micro (Springer Berlin Heidelberg)
- Schneider, P. 2014, Extragalactic Astronomy and Cosmology: An Introduction (Springer Berlin Heidelberg)
- Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. 1998, MNRAS, 296, 873
- Schneider, P. & Watts, P. 2005, A&A, 432, 783
- Soccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20
- Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119
- Shirasaki, M., Takada, M., Miyatake, H., et al. 2017, MNRAS, 470, 3476
- Simon, P., Erben, T., Schneider, P., et al. 2013, MNRAS, 430, 2476
- Simon, P., Saghiha, H., Hilbert, S., et al. 2019, A&A, 622, A104
- Simon, P., Schneider, P., & Kübler, D. 2012, A&A, 548, A102
- Simon, P., Watts, P., Schneider, P., et al. 2008, A&A, 479, 655
- Sofue, Y. & Rubin, V. 2001, ARA&A, 39, 137
- Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175

- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, 435, 629
- Szapudi, I. 2004, *ApJ*, 605, L89
- Takada, M. & Jain, B. 2003, *MNRAS*, 344, 857
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, 480, 22
- Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, *ApJ*, 688, 709
- Tisserand, P., Le Guillou, L., Afonso, C., et al. 2007, *A&A*, 469, 387
- van der Burg, R. F. J., Muzzin, A., Hoekstra, H., et al. 2014, *A&A*, 561, A79
- van Uitert, E., Hoekstra, H., Velandier, M., et al. 2011, *A&A*, 534, A14
- Velandier, M., van Uitert, E., Hoekstra, H., et al. 2014, *MNRAS*, 437, 2111
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *MNRAS*, 444, 1518
- Walker, A. G. 1937, *Proceedings of the London Mathematical Society (Series 2)*, 42, 90
- Wang, Y., Yang, X., Mo, H. J., van den Bosch, F. C., & Chu, Y. 2004, *MNRAS*, 353, 287
- Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, *ApJ*, 646, 881
- Wechsler, R. H. & Tinker, J. L. 2018, *ARA&A*, 56, 435
- White, S. D. M. & Rees, M. J. 1978, *MNRAS*, 183, 341
- Wilson, M. L. 1983, *ApJ*, 273, 2
- Wilson, M. L. & Silk, J. 1981, *ApJ*, 243, 14
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, *ApJ*, 736, 59
- Zeldovich, Y. B. 1972, *MNRAS*, 160, 1P
- Zentner, A. R., Hearin, A. P., & van den Bosch, F. C. 2014, *MNRAS*, 443, 3044
- Zheng, Z. 2004, *ApJ*, 614, 527
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, *ApJ*, 633, 791
- Zheng, Z., Coil, A. L., & Zehavi, I. 2007, *ApJ*, 667, 760
- Zwicky, F. 1933, *Helvetica Physica Acta*, 6, 110

List of Figures

1	Galaxy field	1
1.1	Fate of the Universe for different constellations of Ω_m and Ω_Λ	10
2.1	Halo model approach to describe complex distribution of dark matter	21
3.1	Illustration of G2L and G3L parametrisation	36
3.2	Illustration of excess mass map parametrisation	37
4.1	Best-fit models to the G2L mock data	49
4.2	Best-fit HOD prediction for the mock data	50
4.3	Accuracy test for the mean halo mass using mock data	52
4.4	Accuracy test for the fraction of centrals using mock data	52
4.5	Best-fit models for the satellites-only scenarios using mock data	54
4.6	Accuracy test for the HOD using mock data, low- z	55
4.7	Best-fit models to the CFHTLenS G2L	57
4.8	Best-fit HOD prediction for CFHTLenS	60
4.9	Halo masses for CFHTLenS as a function of stellar mass.	62
4.10	Halo masses for CFHTLenS as a function of luminosity	62
4.11	Fraction of centrals for CFHTLenS as a function of stellar mass	63
4.12	Fraction of centrals for CFHTLenS as a function of luminosity	63
4.13	Central M_*/M_h relationship as a function of halo mass M_h	64
5.1	Gauss test G3L model predictions	69
5.2	Uncertainty on G3L predictions using the bispectrum interpolation	70
5.3	Excess mass maps for various lens-lens separations, stellar mass samples, low- z	71
5.4	Excess mass maps for various lens-lens separations, luminosity samples, low- z	72
5.5	Excess mass maps for various lens-lens separations, galaxy-type samples, low- z	73
5.6	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ for all stellar mass, luminosity and galaxy-type samples, low- z and high- z	74
5.7	1h-, 2h- and 3h-term contribution to the excess mass maps, stellar mass and galaxy-type samples, low- z and high- z	76
5.8	1h-, 2h- and 3h-term contributions to $\langle \mathcal{N}^2 M_{\text{ap}} \rangle$, galaxy-type samples, low- z and high- z	78

5.9	Excess mass dependence on HOD parameters, stellar mass and galaxy-type samples, low- z	80
5.10	Excess mass dependence on HOD parameters, stellar mass and galaxy-type samples, high- z	81
5.11	Sensitivity of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ on HOD parameters, stellar mass samples, low- z . . .	83
5.12	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ dependence on HOD parameters, galaxy-type samples, low- z and high- z	84
5.13	Excess mass map prediction vs. CFHTLenS, combined sm1-sm6 sample, low- z	85
5.14	Excess mass map prediction vs. CFHTLenS, combined sm1-sm6 sample, high- z	86
6.1	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, stellar mass samples, low- z , linear . . .	92
6.2	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, stellar mass samples, high- z , linear . . .	93
6.3	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, luminosity samples, low- z , linear . . .	94
6.4	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, luminosity samples, high- z , linear . . .	95
6.5	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, galaxy-type samples, low- z and high- z , linear	96
F.1	Best-fit models to the G2L mock data, low- z	120
F.2	Best-fit models to the G2L mock data, high- z	121
F.3	Accuracy test for the HOD using mock data, high- z	122
F.4	Model fits to CFHTLenS G2L, stellar mass samples, low- z	123
F.5	Model fits to CFHTLenS G2L, stellar mass samples, high- z	124
F.6	Model fits to CFHTLenS G2L, luminosity samples, low- z	125
F.7	Model fits to CFHTLenS G2L, luminosity samples, high- z	126
F.8	Model fits to CFHTLenS G2L, galaxy-type samples, low- z	127
F.9	Model fits to CFHTLenS G2L, galaxy-type samples, high- z	127
G.1	Excess mass maps for various lens-lens separations, stellar mass samples, high- z	130
G.2	Excess mass maps for various lens-lens separations, luminosity samples, high- z	131
G.3	Excess mass maps for various lens-lens separations, galaxy-type samples, high- z	131
G.4	Excess mass dependence on HOD parameters, low- z , I	132
G.5	Excess mass dependence on HOD parameters, low- z , II	133
G.6	Excess mass dependence on HOD parameters, high- z , I	134
G.7	Excess mass dependence on HOD parameters, high- z , II	135
G.8	Sensitivity of $\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ on HOD parameters, stellar mass samples, high- z . . .	136
H.1	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, stellar mass samples, low- z	138
H.2	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, stellar mass samples, high- z	139
H.3	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, luminosity samples, low- z	140
H.4	$\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, luminosity samples, high- z	141

H.5 $\langle \mathcal{N}^2 M_{\text{ap}} \rangle$ prediction vs. CFHTLenS, galaxy-type samples, low- z and high- z 142

List of Tables

1.1	Growth behaviour of density perturbations within linear perturbation theory	14
4.1	Selection criteria for the lens galaxy samples for the G2L and G3L analysis	44
4.2	Best-fit HOD parameters for the mock data	50
4.3	Derived quantities from best-fit parameter sets for Millennium Simulation .	51
4.4	Best-fit HOD parameters for CFHTLenS	58
4.5	Derived quantities from best-fit parameter sets for CFHTLenS	61
6.1	χ^2 -test G3L model predictions vs. measurements	91
A.1	Natural and astrophysical constants; cosmological, halo model and HOD parameters	105

Acronyms

CMB	Cosmic Microwave Background
NFW	Navarro-Frenk-White
GR	General Relativity
ΛCDM	Lambda Cold Dark Matter
COBE	Cosmic Background Explorer
WMAP	Wilkinson Microwave Anisotropy Probe
HDM	Hot Dark Matter
CDM	Cold Dark Matter
BAO	Baryonic Acoustic Oscillations
MACHOs	MAssive Compact Halo Objects
WIMP	Weakly Interacting Massive Particle
HOD	halo occupation distribution
CFHTLenS	Canada-France Hawaii Telescope Lensing Survey
RCS	Red-Sequence Cluster Survey
G2L	galaxy-galaxy lensing
G3L	galaxy-galaxy-galaxy lensing
SAM	semi-analytical model
ML	maximum likelihood
KIDS	Kilo-Degree Survey
DES	Dark Energy Survey

Acknowledgements

First of all, I would like to thank Peter Schneider for the opportunity to do my PhD project under his supervision in the field of research that fascinates me ever since I can remember. I am grateful for his inputs, ideas and constructive support. I would also like to thank Cristiano Porciani for being my second referee, and for not hesitated to ask the painful questions which to answer were crucial for the success of this thesis. Many thanks also to Jochen Dingfelder and Anke Schiedel for showing interest in my work by agreeing to be part of my Promotionskommission.

Particularly as to mathematical and coding question, Patrick Simon as the G3L expert was the person to approach and also the person who spent countless hours to clarify questions and goals. Thank you for always having time for me, for providing your code to compute the excess mass maps and aperture statistics and for your effort to improve this thesis manuscript by giving many valuable comments. Many thanks also to the other two G3L team members Hananeh and Reiko. Moreover, I would like to thank Dandan, Stefan, David, Ellen, and Sabine.

Not able to escape, my office mate Sophia took an important role as a true all-rounder: sparring partner, debugging expert, interpolation tamer, Friday crisis consultant, Thursday crisis consultant, .., and invaluable friend. Thank you Sophia. For everything.

I am incredibly grateful that this PhD made me meet people that will enrich my life on the long term: Axel, Marcelo and Jan. And certainly, this PhD time would not have been as exciting, colourful and lovely without Geli, Johanna and the Fiddlers Friday people: Abel, Nathan, David and Bruce.

I am deeply grateful to my family, particularly to my father for making me wonder about why the world is as it is and thereby planting this seed of curiosity for the Universe as a whole, and to my mother for equipping me with a strong will to tackle the problems the Universe brings as a whole. And from the bottom of my heart I am utmost grateful to Elvijs for supporting me in every way possible, thank you.

