Topics in Large-Scale Structure

Thesis by

Anna Elisabeth Krause

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2012
(Defended April 16, 2012)

Acknowledgments

Getting to and through grad school has been a quite a journey, and I would like to thank some of the people who made it possible and more enjoyable.

First and foremost, I would like to thank my adviser Chris Hirata, who is primarily responsible for my growth and development as a scientist during my time at Caltech. I am very thankful for the opportunity to learn from and work with him. Thank you, Chris, for suggesting such a broad range of interesting and educational problems to work on, giving me the freedom to explore different topics and to find my own solutions, always being available to students even in the busiest of times, and all your support and patience.

I thank my thesis committee, Andrew Benson, Nick Scoville, Tony Readhead, and Wal Sargent for their time and advice.

I thank Elsbeth Kukowski, Thomas Biedermann, Torsten Enßlin, Lars Hernquist, and Peter Schneider for directions and encouragement at various stages of my education path.

I am very grateful to my parents, Dorothea and Jürgen, my sister Katharina, and Tim for *everything*, especially, their constant support and believing in me and my dreams.

And thanks to Chester House, for keeping me sane through a lot of nonscientific insanities and fun.

Abstract

This thesis presents my personal survey of topics and methods in large-scale structure, covering a range of cosmological probes and analytical, numerical, and observational techniques.

Chapters 2–4 present analytic calculations of systematic effects relevant for the interpretation of data from upcoming large-scale structure surveys: In chapter 2 we derive the relation between measured galaxy ellipticities and the cosmic shear power spectrum up to fourth order in the matter density field, accounting for multiple deflections along the light path, reduced shear, and magnification bias. In chapter 3 we develop a new third-order cosmic shear statistics, which separates shear three point correlation functions exactly into E- and B-mode correlations on a finite interval. This is the first third-order shear statistics free of E/B-mode leakage that relies solely on information from the data. Chapter 4 considers the effect of tidal galaxy alignments on the projected galaxy bispectrum, which are found to bias the inferred galaxy bias parameters.

Chapter 5 focusses on the halo-occupation distribution formalism, which constrains the relation between galaxy luminosities and the masses of their host halos through clustering measurements. We extend this method to model the cross-correlation functions between a galaxy sample of interest and multiple tracer populations simultaneously. This technique improves the accuracy of clustering analyses for sparse galaxy populations, and we apply it to constrain the environment of $(NUV - r)$ selected green valley galaxy samples. These galaxy samples are constructed by matching the Sloan Digital Sky Survey Data Release 7 with the latest Galaxy Evolution Explorer source catalog which provides near ultraviolet photometry. We present cross-correlation function measurements and determine the halo occupation distribution of these transitional galaxies using the multiple tracer technique. At fixed luminosity we find the halo occupation distribution of green and blue central galaxies to be indistinguishable, and the halo masses of green satellite galaxies to be intermediate between those of blue and red satellite galaxies.

In chapter 6 we examine sources of scatter in scaling relations between galaxy cluster mass and thermal Sunyaev-Zeldovich (SZ) effect using cluster samples extracted from cosmological hydrodynamical simulations. This sample enables us to study for the first time the detailed evolution of merging clusters around the scaling relation for a cosmologically representative distribution of merger parameters. We find major mergers to cause an asymmetric scatter such that the inferred mass of merging systems is biased low. As the fraction of dynamically disturbed clusters increases with redshift, this analysis indicates that mergers cause a redshift-dependent bias in cluster mass scaling relations.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Summary

1.1 Cosmological Background

The CfA Redshift Survey (Davis et al. 1982) mapped the distribution on 2400 galaxies with $m < 14.5$. It was the first wide-angle survey to reach beyond the Local Supercluster and to provide evidence of the filamentary distribution of large-scale structure. Since then, cosmology has developed into a precise, data-rich science. The latest generation of large-scale structure surveys consist of order one million (spectroscopic) to several million (photometric) galaxies. The results from these surveys are in excellent agreement with other cosmological probes. Together with current measurements of the Cosmic Microwave Background (CMB) temperature anisotropies, the supernova luminosity–redshift relation, and the local expansion rate of the Universe, the distribution and growth of large-scale structure is one of the foundations of our current standard cosmology picture of a spatially flat Universe whose energy density consists to about 73% of dark energy, 22% dark matter, and 4.5% baryonic matter (Komatsu et al. 2010).

Dark energy is needed in this picture to counteract gravity and explain the observed accelerated expansion of the Universe. Its physical nature is a mystery to fundamental physics, but its abundance and equation of state can be constrained by measurements of the expansion history of the Universe and the abundance and growth of large-scale structure. Several upcoming and proposed galaxy redshift surveys are designed to measure the spatial distribution and shapes of millions of galaxies and aim to determine the composition and initial conditions of the Universe and the nature of dark energy through galaxy clustering, weak gravitational lensing, and other cosmological probes at unprecedented accuracy.

The large volume probed by these surveys will enable us to measure not only the power spectrum, but also higher-order statistics with percent-level precision. As the evolved density field is non-Gaussian, the three-point correlation function, and its Fourier space equivalent the bispectrum, contain significant cosmological information complementary to the more commonly used two-point statistics (e.g., Takada and Jain 2004, Sefusatti et al. 2006, for cosmic shear/galaxy clustering respectively).

This thesis describes a number systematic effects and methods relevant for interpreting data from upcoming large-scale structure surveys, and I summarize the different cosmological probes and specific projects in the following sections.

1.2 Part I: Cosmic Shear

Cosmic shear, the distortion of images of distant galaxies by the tidal field of the intervening large-scale structure, is one of the most promising methods to probe the matter distribution in the universe. Light emitted from distant galaxies travels through the Universe and is continuously deflected by the gravitational field of the inhomogeneous matter distribution. As a consequence the shapes of galaxy images are distorted and the statistical properties of these distortions are related to the statistical properties of the large-scale matter distribution and the geometry of the universe, and can thereby be used to constrain cosmology. Current results already demonstrate the power of cosmic shear observations at constraining the clustering amplitude σ_8 and the matter density $\Omega_{\rm m}$ (e.g., Fu et al. 2008, Schrabback et al. 2007, Huff et al. 2011). Furthermore, cosmic shear provides an ideal tool to study dark energy through measuring the growth of structure with large large-scale structure surveys. These experiments will limit the statistical uncertainties in weak lensing measurements to the percent level. In order to extract the cosmological information, the increased data quality needs to be accompanied by a thorough analysis and treatment of a wide range of systematic errors, from photometric redshifts and galaxy shape measurements to the removal of astrophysical contaminants. The prediction of lensing observables also requires precise models of the nonlinear matter power spectrum and models for the relation between lensing distortion and large-scale matter distribution that go beyond linear theory. We address two of these problems in the following chapters.

1.2.1 Chapter 2: Weak Lensing Power Spectra for Precision Cosmology

It is usually assumed that the ellipticity power spectrum measured in weak lensing observations can be expressed as an integral over the underlying matter power spectrum. This is true at order $O(\Phi^2)$ in the gravitational potential. We extend the standard calculation, constructing all corrections to order $O(\Phi^4)$. There are four types of corrections: corrections to the lensing shear due to multiple-deflections; corrections due to the fact that shape distortions probe the reduced shear $\gamma/(1-\kappa)$ rather than the shear itself; corrections associated with the nonlinear conversion of reduced shear to mean ellipticity; and corrections due to the fact that observational galaxy selection and shear measurement is based on galaxy brightnesses and sizes which have been (de)magnified by lensing. We show how the previously considered corrections to the shear power spectrum correspond to terms in our analysis, and highlight new terms that were not previously identified. All correction terms are given explicitly as integrals over the matter power spectrum, bispectrum, and trispectrum, and are numerically evaluated for the case of sources at $z = 1$. We find agreement with previous works for the $O(\Phi^3)$ terms. We find that for ambitious future surveys, the $O(\Phi^4)$ terms affect the power spectrum at the $\sim 1\sigma - 5\sigma$ level; they will thus need to be accounted for, but are unlikely to represent a serious difficulty for weak lensing as a cosmological probe.

1.2.2 Chapter 3: A new Third-Order Cosmic Shear Statistics

If the shear estimated from observed galaxy shapes is solely caused by gravitational lensing, then it should consist only of a "gradient component," the so-called E-mode shear. B-modes (or curl components) cannot be generated by gravitational light deflection to leading order, and higher-order corrections are expected to be very small. Hence observing any B-mode pattern indicates remaining systematics in the shear analysis. An E/B-mode decomposition is commonly performed using the aperture dispersion (Schneider et al. 1998) and related measures (e.g., Crittenden et al. 2002), which can be calculated from the measured shear two-point correlation function (2PCF) and is thus not affected by the masking geometry. However, these methods assume that the 2PCF is known either from 0 to some finite angular value (aperture dispersion) or to arbitrarily large separations. However, in reality the 2PCF can only be measured on a finite interval $[\theta_{\rm min}, \theta_{\rm max}]$, where the lower boundary is caused by inability to measure the shape of image pairs with very small angular separation. In Eifler, Schneider and Krause (2010) and Schneider, Eifler and Krause (2010) we develop statistical measures for an exact E/B-mode decomposition based on 2PCFs known only on a finite interval $[\theta_{\rm min}, \theta_{\rm max}]$.

Due to the rich mathematical structure of the shear three-point correlation functions (3PCFs), extending these concepts to the three-point level is far from trivial. Currently used methods to decompose shear 3PCFs into E- and B-correlations require knowledge of the 3PCF down to arbitrary small scales. This implies that the 3PCF needs to be modeled on scales smaller than the minimum separation between galaxies required to measure their shapes and subsequently will be biased toward the model, or, in the absence of a model, the statistics is affected by E/B-mode leakage (or mixing).

In this chapter we derive a new third-order E/B-mode statistics that performs the decomposition using the 3PCF only on a finite interval, and thereby is free of any E/B-mode leakage while at the same time relying solely on information from the data. In addition, we relate this third-order ring statistics to the convergence field, thereby enabling a fast and convenient calculation of this statistic from numerical simulations.

1.3 Part II: Galaxy Clustering

While the gravitational growths of dark matter perturbations in the current ΛCDM model is theoretically well understood, the relation between the galaxy distribution and the large-scale (dark) matter distribution is complicated by the detailed physics of galaxy formation and different models may lead to different clustering properties of galaxies.

Local theories of galaxy formation predict galaxy density fluctuations to trace the matter fluctuations on large scales, related by the *bias parameter b*, which is in general not known *a priori*. The unknown bias parameter represents a key problem (or nuisance parameter) for attempts to measure the growth of cosmological perturbations using galaxies. On the other hand, galaxy bias and the small-scale clustering of galaxies are also interesting in their own as contain an wealth of information about galaxy evolution. We explore these two aspects of galaxy clustering in the next two chapters.

1.3.1 Chapter 4: Tidal Alignments as a Contaminant of the Galaxy Bispectrum

In combination with the galaxy power spectrum, third-order galaxy clustering measures such as the bispectrum or (equivalently) the three-point correlation function can be used to measure nonlinear galaxy bias and break the degeneracy between the normalization of the matter power spectrum, σ_8, and the linear galaxy bias. This enables one to remove the effects of galaxy biasing and measure the cosmological growth of structure from the galaxy distribution (Fry 1994, Verde et al. 1998, Scoccimarro, Couchman and Frieman 1999), and thus constrain dark energy.

The most important systematic errors in interpreting the observed galaxy clustering arise in the nonlinear regime,where the behavior of galaxy biasing and models of the (redshift space) galaxy power spectrum and bispectrum are difficult to mode. Recently Hirata (2009) showed that the alignment of galaxies by large-scale tidal fields can cause a systematic error in the determination of the linear redshift space distortion parameter β (Kaiser 1987): the alignment of galaxies with the tidal field (along the stretching axis of the field for large elliptical galaxies) in combination with a viewing direction dependent galaxy selection effect, e.g., preferential selection of galaxies which are observed along their long axis, will lead to a selection probability for galaxies which is modulated by the tidal field along the line of sight. This results in an anisotropy in redshift-space clustering with the same scale and angular dependence as the linear redshift-space effect.

If the orientations of galaxies are correlated with large-scale structure, then anisotropic selection effects such as preferential selection of face-on disc galaxies can contaminate large-scale structure observables. In this chapter we consider the effect on the galaxy bispectrum, which has attracted interest as a way to break the degeneracy between galaxy bias and the amplitude of matter fluctuations σ_8. We consider two models of intrinsic galaxy alignments: one where the probability distribution for the galaxy's orientation contains a term linear in the local tidal field, appropriate for elliptical galaxies; and one with a term quadratic in the local tidal field, which may be applicable to disc galaxies. We compute the correction to the redshift-space bispectrum in the quasilinear regime, and then focus on its effects on parameter constraints from the transverse bispectrum, i.e., using triangles $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ in the plane of the sky. We show that in the linear alignment model, intrinsic alignments result in an error in the galaxy bias parameters, but do not affect the inferred value of σ_8. In contrast, the quadratic alignment model results in a systematic error in both the bias parameters and σ_8. However, the quadratic alignment effect has a unique configuration dependence that should enable it to be removed in upcoming surveys.

1.3.2 Chapter 5: Clustering of Green Valley Galaxies

In this chapter we use the halo-occupation formalism, which describes the relation between galaxies and halo mass in terms of the probability $P(N, M_\mathrm{h})$ that a halo of given mass M_h contains N galaxies, We extend this technique to model the cross-correlation function between a galaxy sample of interest and multiple tracer populations simultaneously. Our method can be applied to commonly used luminosity threshold samples as well as to color and luminosity bin selected galaxy samples, and improves the accuracy of clustering analyses for sparse galaxy populations.

We demonstrate this newly developed method in a clustering analysis of $(NUV - r)$ color selected luminosity bin samples of green valley galaxies. These galaxy samples are constructed by matching the Sloan Digital Sky Survey Data Release 7 with the latest Galaxy Evolution Explorer source catalog which provides near ultraviolet photometry. We present cross-correlation function measurements and determine the halo occupation distribution of these transitional galaxies using the multiple tracer technique. We confirm the previously observed trend that red galaxies reside in more massive halos and are more likely to be satellite galaxies than average galaxies of similar luminosity. At fixed luminosity we find the halo occupation distribution of green and blue central galaxies to in indistinguishable, and the halo masses of green satellite galaxies to be intermediate between those of blue and red satellite galaxies.

1.4 Part III: Galaxy Clusters

Clusters of galaxies are the most massive gravitationally bound objects in the universe, which makes them an important tool for cosmology: among other tests, their abundance provides information on the gravitational growth of structures and is regulated by the initial density field, gravity, and the expansion history of the universe, which critically depend on the underlying cosmology. Thus number counts of clusters, for which masses and redshifts are known, can be used to constrain cosmological parameters (see Allen, Evrard and Mantz 2011, for a recent review).

To relate observed number counts to theoretical predictions of the cluster mass function, these experiments need to infer cluster masses from observables. The thermal Sunyaev Zeldovich (SZ) effect, the signature of inverse Compton scattering of cosmic microwave background photons with hot cluster electron, is thought to provide an excellent mass proxy as the SZ signal is proportional to the total thermal energy of a cluster and is thus less affected by physical processes in the cluster core which can largely affect the X-ray luminosity. This is confirmed by simulations (e.g., Nagai 2006, Shaw, Holder and Bode 2008, Battaglia et al. 2010, Sehgal et al. 2010) finding the scatter in the mass–SZ scaling relation to be of order 5% - 10%. Furthermore, the SZ effect is not subject to surface brightness dimming and has a very weak redshift dependence, making it an ideal probe to study galaxy clusters at high redshift.

Currently several large surveys are starting to detect hundreds of galaxy clusters through their SZ signal (Vanderlinde et al. 2010, Marriage et al. 2011, Planck Collaboration et al. 2011a) and derive cosmological constraint based on these samples (Andersson et al. 2011, Sehgal et al. 2011, Williamson et al. 2011). To exploit the statistical power of these upcoming cluster samples, the mapping between SZ signal and cluster mass needs to be well understood. Observations find normalization and slope of the scaling relations between SZ signal and lensing derived masses (Marrone et al. 2011), or between SZ signal and X-ray properties (Planck Collaboration et al. 2011b,c) to be consistent with self-similar scaling and predictions from simulations.

Due to the steep slope of the cluster mass function, competitive cosmological constraints from these experiments require information about the distribution and redshift evolution of scatter in the mass scaling relation (e.g., Majumdar and Mohr 2004, Lima and Hu 2005, Shaw, Holder and Dudley 2010). As the true cluster mass and other physical cluster properties which may bias the mass proxy are unobservable, and as the noise and biases in the different mass estimators may be correlated, characterizing the intrinsic scatter in any of these scaling relation is difficult to obtain from observations. Hence the sources and distribution of scatter in different mass estimators are mainly studied through simulations and mock observations (e.g., Rasia et al. 2006, Nagai, Vikhlinin and Kravtsov 2007, Shaw, Holder and Bode 2008, Becker and Kravtsov 2010, Yang, Bhattacharya and Ricker 2010, Fabjan et al. 2011).

1.4.1 Chapter 6: Merger Induced Scatter and Bias in the Cluster Mass–Sunyaev-Zeldovich effect scaling relation

We examine sources of scatter in scaling relations between galaxy cluster mass and thermal Sunyaev-Zeldovich (SZ) effect using cluster samples extracted from cosmological hydrodynamical simulations. Overall, the scatter of the mass-SZ scaling relation is well correlated with the scatter in the mass-halo concentration relation with more concentrated halos having stronger integrated SZ signals at fixed mass. Additional sources of intrinsic scatter are projection effects from correlated structures, which cause the distribution of scatter to deviate from log-normality and skew it toward higher inferred masses, and the dynamical state of clusters. We study the evolution of merging clusters based on simulations of 39 clusters and their cosmological environment with high time resolution. This sample enables us to study for the first time the detailed evolution of merging clusters around the scaling relation for a cosmologically representative distribution of merger parameters. Major mergers cause an asymmetric scatter such that the inferred mass of merging systems is biased low. We find mergers to be the dominant source of bias toward low inferred masses: over 50% of outliers on this side of the scaling relation underwent a major merger within the last Gigayear. As the fraction of dynamically disturbed clusters increases with redshift, our analysis indicates that mergers cause a redshift-dependent bias in scaling relations. Furthermore, we find the SZ morphology of massive clusters to be well correlated with the clusters' dynamical state, suggesting that morphology may be used to constrain merger fractions and identify merger-induced outliers of the scaling relation.

Chapter 2

Weak Lensing Power Spectra for Precision Cosmology*

2.1 Introduction

Cosmic shear, the distortion of light from distant galaxies by the tidal gravitational field of the intervening large-scale structure, is an excellent tool to probe the matter distribution in the universe. The statistics of the image distortions are related to the statistical properties of the large-scale matter distribution and can thereby be used to constrain cosmology. Current results already demonstrate the power of cosmic shear observations at constraining the clustering amplitude σ_8 and the matter density $\Omega_{\rm m}$ (e.g., Massey et al. 2007b, Schrabback et al. 2007, Benjamin et al. 2007, Fu et al. 2008). Furthermore, cosmic shear provides an ideal tool to study dark energy through measuring the evolution of nonlinear structure and upcoming large weak lensing experiments will limit the statistical uncertainties to the percent level.

In order to extract cosmological information from these cosmic shear experiments, the increased data quality needs to be accompanied by a thorough treatment of systematic errors. On the observational side, this requires accurate information on the redshift distribution of source galaxies (Ma, Hu and Huterer 2006) and precise measurements of galaxy shapes which correct for observational systematics such as pixelization, noise, blurring by seeing and a spatially variable point spread function (see Massey et al. 2007a, Bridle et al. 2009). On the theoretical side, astrophysical contaminants, like source lens clustering (Bernardeau, van Waerbeke and Mellier 1997, Schneider, van Waerbeke and Mellier 2002), intrinsic alignment (King and Schneider 2003) and the correlation between

---

*This chapter was adapted from Weak lensing power spectra for precision cosmology. Multiple-deflection, reduced shear, and lensing bias corrections, Elisabeth Krause and Christopher M. Hirata; *A&A*, **523**, A28 (2010). Reproduced here with permission, copyright (2010) by ESO.

the gravitational shear and intrinsic ellipticities of galaxies (Hirata and Seljak 2004, King 2005, Joachimi and Schneider 2008, Zhang 2008, Joachimi and Schneider 2009), need to be understood and removed. The prediction of lensing observables also requires precise models of the nonlinear matter power spectrum and models for the relation between lensing distortion and large-scale matter distribution which go beyond linear theory. While N-body simulations may predict the nonlinear dark matter power spectrum with percent level accuracy in the near future (Heitmann et al. 2008, 2009), the effect of baryons, which is a significant contamination to the weak lensing signal above $l \sim 2000$ (Jing et al. 2006, Rudd, Zentner and Kravtsov 2008), is more difficult to account for and is the subject of ongoing work.

In this chapter, we consider corrections to the relation between the observed lensing power spectra and the nonlinear matter density field. In the regime of weak lensing, the observed galaxy ellipticities (e_I) are an estimator of the reduced shear $g_I = \gamma_I/(1-\kappa)$,

$$\langle e_I \rangle = C\frac{\gamma_I}{1-\kappa}\ , \tag{2.1}$$

where C is a constant which depends on the type of ellipticity estimator (e.g., Schneider and Seitz 1995, Seitz and Schneider 1997) and the properties of the galaxy population under consideration, γ_I is a component of the shear, κ is the convergence, and the subscript I refers to the two components of the ellipticity/shear (see e.g., Bartelmann and Schneider 2001, for more details). The two-point statistics of the measured ellipticities are simply related to the reduced shear power spectrum. Cooray and Hu (2002) have calculated the shear power spectrum to fourth order in the gravitational potential. For the reduced shear power spectrum there exists an approximation to third-order in the gravitational potential (Dodelson, Shapiro and White 2006). Shapiro (2009) has demonstrated that on angular scales relevant for dark energy parameter estimates the difference between shear and reduced shear power spectra is at the percent level and ignoring these corrections will noticeably bias dark energy parameters inferred from future weak lensing surveys.

Schmidt et al. (2009a) introduced another type of corrections, termed *lensing bias*, which has a comparable effect on the shear power spectrum as the reduced shear correction: Observationally, shear is only estimated from those galaxies which are bright enough and large enough to be identified and to measure their shape. This introduces cuts based on observed brightness and observed size, both of which are (de)magnified by lensing (e.g., Broadhurst, Taylor and Peacock 1995, Jain 2002), and will thus bias the sampling of the cosmic shear field.

In the following we complete the calculation of the reduced shear power spectrum to fourth order in the gravitational potential to include multiple deflections and to account for the effects of lensing bias and the nonlinear conversion between ellipticity and reduced shear. We consider all lensing-related effects through $O(\Phi^4)$, but do *not* include effects associated with the sources (source clustering and intrinsic alignment corrections).

This chapter is organized as follows: We describe our technique for calculating higher-order lensing distortions and power spectra in section 2.2. Derivations of the different types of corrections to the shear and reduced shear power spectra are given in section 2.3.1 through section 2.3.4. We quantify the impact of these corrections on future surveys in section 2.4 and discuss our results in section 2.5.

2.2 Calculational Method

In this section we derive the higher-order lensing distortions following Hirata and Seljak (2003), and introduce our technique and notation for calculating power spectrum corrections.

Throughout this calculation we assume a flat universe and work in the flat sky approximation. We use a unit system based on setting the speed of light $c = 1$, which makes potentials dimensionless. We use the Einstein summation convention and sum over all Roman indices appearing twice in a term. Lower case, italic type Roman indices $a, b, c, \ldots = 1, 2$ are used to for Cartesian components of two dimensional vectors and tensors; capital case, italic type Roman indices $I, J, K, \ldots = 1, 2$ are used for the components of polars which are defined with reference to a Cartesian coordinate system but have different transformation properties. Greek indices are used for redshift slices.

2.2.1 Lensing Distortion Tensor

We work in the flat sky approximation and choose the sky to lie in the xy-plane. Photons travel roughly along the $-\hat{\mathbf{z}}$ direction and are deflected by the Newtonian potential Φ generated by the nonrelativistic matter inhomogeneities. As long as their deflection from the $-\hat{\mathbf{z}}$ direction is small, they observe a metric (e.g., Hirata and Seljak 2003)

$$ds^2 = a^2(\tau)\left[-(1+2\Phi)d\tau^2 + (1-2\Phi)\left(d\chi^2 + \chi^2(dn_x^2 + dn_y^2)\right)\right], \tag{2.2}$$

where a is the scale factor, χ is the comoving radial distance, and $\mathbf{n}$ is the angular coordinate of the photon path on the sky. We calculate the deflection angle of a light ray from its null geodesic equation

$$\frac{\mathrm{d}}{\mathrm{d}\chi}\left(\frac{\mathrm{d}\mathbf{n}}{\mathrm{d}\chi}\chi\right) = -2\frac{\partial\Phi(\mathbf{x}(\mathbf{n},\chi);z(\chi))}{\partial\mathbf{n}}\chi\,, \tag{2.3}$$

where $\Phi(\mathbf{x};z)$ is the Newtonian potential at position $\mathbf{x}$ and redshift z, with initial conditions $\mathbf{n}(\chi = 0) = \mathbf{n}_0$ and $\partial_\chi\mathbf{n}(\chi = 0) = 0$.

To first order in Φ, the integration is performed along the unperturbed photon trajectory, this is the so-called Born approximation. Taylor expanding equation (2.3) to third-order in Φ we obtain a perturbative solution for the deflection angle $\mathbf{d} \equiv \mathbf{n} - \mathbf{n}_0$,

$$\begin{aligned}
n_i(z_\mathrm{s}) &= n_{0i} + d_i^{(1)}(z_\mathrm{s}) + d_i^{(2)}(z_\mathrm{s}) + d_i^{(3)}(z_\mathrm{s}) \\
&= n_{0i} + d^{(1)}(z_\mathrm{s}) - 2\int_0^{\chi_\mathrm{s}}\mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^2\Phi_{,ia}(\chi)d_a^{(1)}(\chi) \\
&\quad - 2\int_0^{\chi_\mathrm{s}}\mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^2\left(\frac{1}{2}\chi\Phi_{,iab}(\chi)d_a^{(1)}(\chi)d_b^{(1)}(\chi) + \Phi_{,ia}(\chi)d_a^{(2)}(\chi)\right)
\end{aligned} \tag{2.4}$$

$$\begin{aligned}
&= n_{0i} - 2\int_0^{\chi_\mathrm{s}}\mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi\Phi_{,i}(\chi) + 4\int_0^{\chi_\mathrm{s}}\mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^2\int_0^{\chi}\mathrm{d}\chi' W(\chi',\chi)\chi'\Phi_{,ia}(\chi)\Phi_{,a}(\chi') \\
&\quad - \left(4\int_0^{\chi_\mathrm{s}}\mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^3\int_0^{\chi}\mathrm{d}\chi' W(\chi',\chi)\chi'\int_0^{\chi}\mathrm{d}\chi'' W(\chi'',\chi)\chi''\Phi_{,iab}(\chi)\Phi_{,a}(\chi')\Phi_{,b}(\chi'')\right. \\
&\quad \left. + 8\int_0^{\chi_\mathrm{s}}\mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^2\int_0^{\chi}\mathrm{d}\chi' W(\chi',\chi)\chi'^2\int_0^{\chi'}\mathrm{d}\chi'' W(\chi'',\chi')\chi''\Phi_{,ia}(\chi)\Phi_{,ab}(\chi')\Phi_{,b}(\chi'')\right),
\end{aligned} \tag{2.5}$$

where $W(\chi',\chi) = \left(\frac{1}{\chi'} - \frac{1}{\chi}\right)\Theta(\chi - \chi')$ with $\Theta(x)$ the Heaviside step function. Here $\chi_\mathrm{s} = \chi(z_\mathrm{s})$ is the comoving distance of a source at redshift z_s, commata represent comoving spatial transverse derivatives. These spatial derivatives are evaluated at the unperturbed position $\Phi(\chi) = \Phi(\mathbf{n}_0\chi,\chi;z(\chi))$ unless otherwise indicated. The first- and second-order deflection angles are identical to those found by Hirata and Seljak (2003) [1]. The third-order deflection angles are caused by the two types of second-order transverse displacement in the Taylor expansion of $\Phi(\mathbf{x};z)$ shown in equation (2.4). We discuss the difference between these terms after equation (2.8).

---

[1] Our notation differs from Hirata and Seljak (2003) in using spatial instead of angular derivatives to simplify comparison with Cooray and Hu (2002), Dodelson, Shapiro and White (2006), Shapiro (2009)

The distortion of a light ray is then described by the Jacobian matrix

$$\mathbf{A}(\mathbf{n}_0, z_\mathrm{s}) = \frac{\partial \mathbf{n}(z_\mathrm{s})}{\partial \mathbf{n}_0} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2-\omega \\ -\gamma_2+\omega & 1-\kappa+\gamma_1 \end{pmatrix} , \tag{2.6}$$

where γ_I are the cartesian components of the shear, and ω induces an (unobservable) rotation of the image. Using (2.5), the distortion tensor $\psi_{ij} = \delta_{ij} - A_{ij}$ is given by

$$\psi_{ij}(\mathbf{n}_0, z_\mathrm{s}) = \psi_{ij}^{(1)}(\mathbf{n}_0, z_\mathrm{s}) + \psi_{ij}^{(2)}(\mathbf{n}_0, z_\mathrm{s}) + \psi_{ij}^{(3A)}(\mathbf{n}_0, z_\mathrm{s}) + \psi_{ij}^{(3B)}(\mathbf{n}_0, z_\mathrm{s}) + \psi_{ij}^{(3C)}(\mathbf{n}_0, z_\mathrm{s}), \tag{2.7}$$

where

$$\begin{aligned}
\psi_{ij}^{(1)}(\mathbf{n}_0, z_\mathrm{s}) &= 2\int_0^{\chi_\mathrm{s}} \mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^2 \Phi_{,ij}(\chi) , \\
\psi_{ij}^{(2)}(\mathbf{n}_0, z_\mathrm{s}) &= -4\int_0^{\chi_\mathrm{s}} \mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^2 \int_0^{\chi} \mathrm{d}\chi' W(\chi',\chi)\chi' \left\{\Phi_{,ia}(\chi)\chi'\Phi_{,aj}(\chi') + \chi\Phi_{,ija}(\chi)\Phi_{,a}(\chi')\right\} , \\
\psi_{ij}^{(3A)}(\mathbf{n}_0, z_\mathrm{s}) &= +4\underbrace{\int_0^{\chi_\mathrm{s}} \mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^4 \int_0^{\chi} \mathrm{d}\chi' W(\chi',\chi)\chi' \int_0^{\chi} \mathrm{d}\chi'' W(\chi'',\chi)}_{\Longrightarrow \chi'<\chi,\, \chi''<\chi} \chi''\Phi_{,ijab}(\chi)\Phi_{,a}(\chi')\Phi_{,b}(\chi'') , \\
\psi_{ij}^{(3B)}(\mathbf{n}_0, z_\mathrm{s}) &= +8\underbrace{\int_0^{\chi_\mathrm{s}} \mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^3 \int_0^{\chi} \mathrm{d}\chi' W(\chi',\chi)\chi'^2 \int_0^{\chi} \mathrm{d}\chi'' W(\chi'',\chi)}_{\Longrightarrow \chi'<\chi,\, \chi''<\chi} \chi''\Phi_{,iab}(\chi)\Phi_{,aj}(\chi')\Phi_{,b}(\chi'') , \\
\psi_{ij}^{(3C)}(\mathbf{n}_0, z_\mathrm{s}) &= +8\underbrace{\int_0^{\chi_\mathrm{s}} \mathrm{d}\chi W(\chi,\chi_\mathrm{s})\chi^2 \int_0^{\chi} \mathrm{d}\chi' W(\chi',\chi)\chi'^2 \int_0^{\chi'} \mathrm{d}\chi'' W(\chi'',\chi')\chi''}_{\Longrightarrow \chi''<\chi'<\chi} \\
&\qquad \times \frac{\partial}{\partial n_{0,j}}[\Phi_{,ia}(\chi)\Phi_{,ab}(\chi')\Phi_{,b}(\chi'')] ,
\end{aligned} \tag{2.8}$$

where we have used the symmetry of the integrals over χ' and χ'' in the derivation of $\psi_{ij}^{(3B)}$. This calculation automatically includes the “Born correction” and “lens-lens coupling” corrections considered by Cooray and Hu (2002). Compared to their approach, we find additional terms $\psi_{ij}^{(3C)}$ which give the third-order corrections caused by three lenses placed at different locations along the line of sight ($\chi'' < \chi' < \chi$), namely the derivatives of the last term in equation (2.8). These include the two terms previously considered by Shapiro and Cooray (2006), however, we will show in section 2.3.1 that within the Limber approximation, the 3C term does not contribute to the shear power spectrum at $\mathcal{O}(\Phi^4)$.

The convergence, shear, and rotation are expressible in terms of ψ_{ij} by the usual rules $\kappa =$

$\frac{1}{2}(\psi_{11} + \psi_{22})$, $\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22})$, $\gamma_2 = \frac{1}{2}(\psi_{12} + \psi_{21})$, and $\omega = \frac{1}{2}(\psi_{12} - \psi_{21})$.

Note that while our derivation of the deflection angle is based on the small angle approximation $d \ll 1$, in the flat sky approximation the elements of the distortion matrix need not be as small.

2.2.2 Fourier Space: First Order

Since we work in terms of power spectra, we need to transform these equations to Fourier space. In the flat-sky approximation,

$$\psi_{ij}(\mathbf{n}_0, z_s) = \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2}\tilde{\psi}_{ij}(\mathbf{l}, z_s)\mathrm{e}^{\mathrm{i}\mathbf{l}\cdot\mathbf{n}_0}. \tag{2.9}$$

The angular cross-power spectra of two fields Γ and Γ' is then defined by

$$\langle\tilde{\Gamma}(\mathbf{l})\tilde{\Gamma}'(\mathbf{l}')\rangle = (2\pi)^2 C_{\Gamma\Gamma'}(l)\delta_\mathrm{D}(\mathbf{l} + \mathbf{l}') \tag{2.10}$$

with δ_D the Dirac delta function, which has units $[\delta_\mathrm{D}(\mathbf{x})] = [x]^{-n}$ where n is the dimension of $\mathbf{x}$. Potentials are functions of a three-dimensional position variable. Following Dodelson and Zhang (2005), we use $\tilde{\phi}$ to denote the Fourier transform of the potential in the angular (transverse) variables only

$$\tilde{\phi}\,(\mathbf{l};\chi) \equiv \frac{1}{\chi^2}\int \frac{\mathrm{d}k_3}{2\pi}\tilde{\Phi}(\mathbf{l}/\chi, k_3; z(\chi))\mathrm{e}^{\mathrm{i}k_3\chi}\ . \tag{2.11}$$

Then the spatial derivatives of the potential can be expressed in terms of the angular Fourier transform $\tilde{\phi}$ as

$$\Phi_{,i_1 i_2 \dots i_M}(\mathbf{n}_0\chi;\chi) = \frac{\mathrm{i}^M}{\chi^M}\int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} l_{i_1} l_{i_2} \dots l_{i_M}\tilde{\phi}(\mathbf{l};\chi)\mathrm{e}^{\mathrm{i}\mathbf{l}\cdot\mathbf{n}_0}. \tag{2.12}$$

Applying this to the first term from equation (2.8) and using the relation between convergence, shear and ψ_{ij}, we arrive at the well-known first order results for convergence and shear

$$\tilde{\kappa}^{(1)}(\mathbf{l}, z_s) = \frac{1}{2}\left(\tilde{\psi}^{(1)}_{11}(\mathbf{l}, z_s) + \tilde{\psi}^{(1)}_{22}(\mathbf{l}, z_s)\right) = -l^2\int_0^{\chi_s}\mathrm{d}\chi W(\chi, \chi_s)\tilde{\phi}(\mathbf{l};\chi) \ \text{ and } \ \tilde{\gamma}^{(1)}_I(\mathbf{l}, z_s) = T_I(\mathbf{l})\tilde{\kappa}^{(1)}(\mathbf{l}, z_s)\ . \tag{2.13}$$

Here $T_1(\mathbf{l}) = \cos(2\phi_l)$ and $T_2(\mathbf{l}) = \sin(2\phi_l)$, where ϕ_l is the azimuthal angle of $\mathbf{l}$.

We generally decompose the shear components into tangential (or E-mode) shear, γ_E and cross (or B-mode) shear, γ_B,

$$\tilde{\gamma}_\mathrm{E}(\mathbf{l}, z_s) = \delta_{IJ}T_I(\mathbf{l})\tilde{\gamma}_J(\mathbf{l}, z_s); \quad \tilde{\gamma}_\mathrm{B}(\mathbf{l}, z_s) = \epsilon_{IJ}T_I(\mathbf{l})\tilde{\gamma}_J(\mathbf{l}, z_s)\,, \tag{2.14}$$

with ϵ_{IJ} the two-dimensional Levi-Civita tensor. To first order, $\tilde{\gamma}_{\rm E}^{(1)}(\mathbf{l}, z_{\rm s}) = \tilde{\kappa}^{(1)}(\mathbf{l}, z_{\rm s})$ and $\tilde{\gamma}_{\rm B}^{(1)}(\mathbf{l}, z_{\rm s}) = \tilde{\omega}^{(1)}(\mathbf{l}, z_{\rm s}) = 0$. Their power spectra can be obtained under the Limber approximation (Kaiser 1992, Dodelson and Zhang 2005, equation (15)),

$$\langle\tilde{\phi}(\mathbf{l};\chi)\tilde{\phi}(\mathbf{l}';\chi')\rangle = (2\pi)^2\delta_{\rm D}(\mathbf{l}+\mathbf{l}')\frac{\delta_{\rm D}(\chi-\chi')}{\chi^2}P_\Phi\left(l/\chi; z(\chi)\right), \tag{2.15}$$

where $P_\Phi\left(l/\chi; z(\chi)\right)$ is the three-dimensional power spectrum of the potential at redshift $z(\chi)$. The lensing tomography cross-spectra between two source redshift slices at z_α and z_β (with $z_\alpha < z_\beta$) then read

$$C_{\tilde{\gamma}_{\rm E}}^{(11)}(l; z_\alpha, z_\beta) = C_{\tilde{\kappa}}^{(11)}(l; z_\alpha, z_\beta) = l^4\int_0^{\chi_\alpha} {\rm d}\chi\frac{W(\chi,\chi_\alpha)W(\chi,\chi_\beta)}{\chi^2}P_\Phi\left(l/\chi; z(\chi)\right)\,, \tag{2.16}$$

and

$$C_{\tilde{\gamma}_{\rm B}}^{(11)}(l; z_\alpha, z_\beta) = C_{\tilde{\omega}}^{(11)}(l; z_\alpha, z_\beta) = 0\,, \tag{2.17}$$

where the superscripts denote the order of expansion in the potential.

2.2.3 Fourier Space: Second Order

To work to second order, we need the usual convolution theorem for the product of two fields U and V is

$$[\widetilde{UV}](\mathbf{l}) \equiv [\tilde{U} * \tilde{V}](\mathbf{l}) = \int\frac{{\rm d}^2\mathbf{l}'}{(2\pi)^2}\tilde{U}(\mathbf{l}')\tilde{V}(\mathbf{l}-\mathbf{l}')\,. \tag{2.18}$$

Introducing

$$\mathfrak{M}\left(\mathbf{l}', \mathbf{l}-\mathbf{l}'; z_{\rm s}\right) = \int_0^{\chi_{\rm s}} {\rm d}\chi W(\chi,\chi_{\rm s})\int_0^{\chi} {\rm d}\chi' W(\chi',\chi)\tilde{\phi}\left(\mathbf{l}';\chi\right)\tilde{\phi}\left(\mathbf{l}-\mathbf{l}';\chi'\right), \tag{2.19}$$

and using the second term from equation (2.8) and the relation between convergence, rotation, shear and ψ_{ij}, the second-order corrections to convergence, rotation and shear can be written as

$$\begin{aligned}
\tilde{\kappa}^{(2)}(\mathbf{l}, z_{\rm s}) &= -2\int\frac{{\rm d}^2\mathbf{l}'}{(2\pi)^2}\left[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')\right]\mathbf{l}\cdot\mathbf{l}'\mathfrak{M}\left(\mathbf{l}', \mathbf{l}-\mathbf{l}'; z_{\rm s}\right)\,, & (2.20)\\
\tilde{\omega}^{(2)}(\mathbf{l}, z_{\rm s}) &= -2\int\frac{{\rm d}^2\mathbf{l}'}{(2\pi)^2}\left[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')\right]ll'\sin(\phi_{l'})\mathfrak{M}\left(\mathbf{l}', \mathbf{l}-\mathbf{l}'; z_{\rm s}\right) & (2.21)
\end{aligned}$$

and

$$\tilde{\gamma}_I^{(2)}(\mathbf{l}, z_\mathrm{s}) = -2\int \frac{\mathrm{d}^2\mathbf{l}'}{(2\pi)^2}\left[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')\right] l l' G_I(\mathbf{l},\mathbf{l}')\mathfrak{M}\left(\mathbf{l}', \mathbf{l}-\mathbf{l}'; z_\mathrm{s}\right). \tag{2.22}$$

Here the superscript refers to the order of expansion in Φ, and we define $G_1(\mathbf{l},\mathbf{l}') = \cos(\phi_l + \phi_{l'})$ and $G_2(\mathbf{l},\mathbf{l}') = \sin(\phi_l + \phi_{l'})$. When we work beyond first order in the lensing potential, the shear becomes a nonlinear function of the gravitational potential Φ. Hence the power spectrum of the shear depends on the higher-order correlation functions of Φ. Therefore we need the Limber approximation for these higher-order correlation functions. For the bispectrum, equation (2.15) generalizes to

$$\left\langle \tilde{\phi}\left(\mathbf{l}_1;\chi_1\right)\tilde{\phi}\left(\mathbf{l}_2;\chi_2\right)\tilde{\phi}\left(\mathbf{l}_3;\chi_3\right)\right\rangle = (2\pi)^2\delta_\mathrm{D}(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3)\frac{\delta_\mathrm{D}(\chi_1-\chi_2)\delta_\mathrm{D}(\chi_1-\chi_3)}{\chi_1^4}B_\Phi\left(\frac{\mathbf{l}_1}{\chi_1},\frac{\mathbf{l}_2}{\chi_1},\frac{\mathbf{l}_3}{\chi_1};z(\chi_1)\right); \tag{2.23}$$

and for the trispectrum,

$$\begin{aligned}\left\langle \tilde{\phi}\left(\mathbf{l}_1;\chi_1\right)\tilde{\phi}\left(\mathbf{l}_2;\chi_2\right)\tilde{\phi}\left(\mathbf{l}_3;\chi_3\right)\tilde{\phi}\left(\mathbf{l}_4;\chi_4\right)\right\rangle_\mathrm{c} &= (2\pi)^2\delta_\mathrm{D}(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3+\mathbf{l}_4)\frac{\delta_\mathrm{D}(\chi_1-\chi_2)\delta_\mathrm{D}(\chi_1-\chi_3)}{\chi_1^4}\\ &\quad\times\frac{\delta_\mathrm{D}(\chi_1-\chi_4)}{\chi_1^2}T_\Phi\left(\frac{\mathbf{l}_1}{\chi_1},\frac{\mathbf{l}_2}{\chi_1},\frac{\mathbf{l}_3}{\chi_1},\frac{\mathbf{l}_4}{\chi_1};z(\chi_1)\right),\end{aligned} \tag{2.24}$$

where the subscript "c" denotes a connected function.

As an example, we consider the correlation of two $\mathfrak{M}$ functions,

$$\begin{aligned}\left\langle\mathfrak{M}\left(\mathbf{l}',\mathbf{l}-\mathbf{l}';z_\alpha\right)\mathfrak{M}\left(\mathbf{l}''',\mathbf{l}''-\mathbf{l}''';z_\beta\right)\right\rangle &= \int_0^{\chi_\alpha}\mathrm{d}\chi\int_0^{\chi_\beta}\mathrm{d}\chi''\int_0^{\chi}\mathrm{d}\chi'\int_0^{\chi''}\mathrm{d}\chi'''\\ &\quad\times W(\chi,\chi_\alpha)W(\chi',\chi)W(\chi'',\chi_\beta)W(\chi''',\chi'')\\ &\quad\times\left\langle\tilde{\phi}\left(\mathbf{l}';\chi\right)\tilde{\phi}\left(\mathbf{l}-\mathbf{l}';\chi'\right)\tilde{\phi}\left(\mathbf{l}''';\chi''\right)\tilde{\phi}\left(\mathbf{l}''-\mathbf{l}''';\chi'''\right)\right\rangle.\end{aligned} \tag{2.25}$$

The expectation value here can be broken up into a Gaussian (Wick's theorem) piece and a connected (non-Gaussian) piece. The connected piece vanishes because the δ_D-functions in equation (2.24) force $\chi = \chi' = \chi'' = \chi'''$ where the window functions vanish. Of the three possible contractions for the Gaussian term, the only one that survives is $\chi'' = \chi > \chi''' = \chi'$. Thus,

$$\left\langle\mathfrak{M}\left(\mathbf{l}',\mathbf{l}-\mathbf{l}';z_\alpha\right)\mathfrak{M}\left(\mathbf{l}''',\mathbf{l}''-\mathbf{l}''';z_\beta\right)\right\rangle = (2\pi)^4\delta_\mathrm{D}(\mathbf{l}'+\mathbf{l}''')\delta_\mathrm{D}(\mathbf{l}-\mathbf{l}'+\mathbf{l}''-\mathbf{l}''')M\left(l',|\mathbf{l}-\mathbf{l}'|;z_\alpha,z_\beta\right), \tag{2.26}$$

where we have introduced the mode-coupling integral

$$M\left(l, l'; z_\alpha, z_\beta\right) = \int_0^{\chi_\alpha} \mathrm{d}\chi \frac{W(\chi, \chi_\alpha) W(\chi, \chi_\beta)}{\chi^2} \int_0^{\chi} \mathrm{d}\chi' \frac{W^2(\chi', \chi)}{\chi'^2} P_\Phi\left(l/\chi; z(\chi)\right) P_\Phi\left(l'/\chi'; z(\chi')\right). \quad (2.27)$$

Note that equation (2.26) is true even for a non-Gaussian density field.

The third-order terms each require specialized treatment, so we handle them on a case-by-case basis below.

2.3 The Corrections to the Power Spectrum

We can now calculate the higher-order contributions to the reduced shear power spectrum by Taylor expanding the reduced shear in terms of the shear and convergence to contain all terms up to $O(\Phi^4)$,

$$\left\langle \tilde{g}_{\mathrm{E/B}}(\mathbf{l}, z_\alpha)\ \tilde{g}_{\mathrm{E/B}}(\mathbf{l}', z_\beta) \right\rangle \approx \left\langle (\tilde{\gamma} * (1 + \tilde{\kappa} + \tilde{\kappa} * \tilde{\kappa}))_{\mathrm{E/B}}(\mathbf{l}, z_\alpha)\ (\tilde{\gamma} * (1 + \tilde{\kappa} + \tilde{\kappa} * \tilde{\kappa}))_{\mathrm{E/B}}(\mathbf{l}', z_\beta) \right\rangle\ , \quad (2.28)$$

where $*$ denotes a convolution, and where the shear and convergence need to be expanded in terms of the potential according to equation (2.8) and projected into E/B components using equation (2.14).

As the power spectra depend only on the magnitude of $\mathbf{l}$, we can choose $\mathbf{l} \| \hat{\mathbf{x}}$, which implies $T(\mathbf{l}) = (1, 0)$ and thus $\tilde{\gamma}_{\mathrm{E}}(\mathbf{l}) = \tilde{\gamma}_1(\mathbf{l})$, and simplifies the calculations without loss of generality. Consider for example the correction to the E-mode power spectrum arising from the correlation between second-order corrections,

$$\begin{aligned}
\left\langle \tilde{g}_{\mathrm{E}}^{(2)}(\mathbf{l}, z_\alpha)\ \tilde{g}_{\mathrm{E}}^{(2)}(\mathbf{l}', z_\beta) \right\rangle &= \left\langle \left(\tilde{\gamma}^{(2)} + \tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}} (\mathbf{l}, z_\alpha)\ \left(\tilde{\gamma}^{(2)} + \tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}} (\mathbf{l}', z_\beta) \right\rangle \\
&= \left\langle \tilde{\gamma}_1^{(2)}(\mathbf{l}, z_\alpha)\ \tilde{\gamma}_{\mathrm{E}}^{(2)}(\mathbf{l}', z_\beta) \right\rangle + \left\langle \left(\tilde{\gamma}_1^{(1)} * \tilde{\kappa}^{(1)}\right)(\mathbf{l}, z_\alpha)\ \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}} (\mathbf{l}', z_\beta) \right\rangle \\
&\quad + \left\langle \tilde{\gamma}_1^{(2)}(\mathbf{l}, z_\alpha)\ \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}} (\mathbf{l}', z_\beta) \right\rangle + \left\langle \left(\tilde{\gamma}_1^{(1)} * \tilde{\kappa}^{(1)}\right)(\mathbf{l}, z_\alpha)\ \gamma_{\mathrm{E}}^{(2)}(\mathbf{l}', z_\beta) \right\rangle \\
&= \left\langle \tilde{\gamma}_1^{(2)}(\mathbf{l})\ T_I(\mathbf{l}') \tilde{\gamma}_I^{(2)}(\mathbf{l}') \right\rangle_{\alpha\beta} + \left\langle \left(\tilde{\gamma}_1^{(1)} * \tilde{\kappa}^{(1)}\right)(\mathbf{l})\ T_I(\mathbf{l}') \left(\tilde{\gamma}_I^{(1)} * \tilde{\kappa}^{(1)}\right)(\mathbf{l}') \right\rangle_{\alpha\beta} \\
&\quad + 2 \left\langle \tilde{\gamma}_1^{(2)}(\mathbf{l})\ T_I(\mathbf{l}') \left(\tilde{\gamma}_I^{(1)} * \tilde{\kappa}^{(1)}\right)(\mathbf{l}') \right\rangle_{\alpha\beta}\ , \qquad (2.29)
\end{aligned}$$

where in the last step we have rewritten the E-mode component using equation (2.14) and where we define the symmetrized expectation value

$$\langle A(\mathbf{l}) B(\mathbf{l}') \rangle_{\alpha\beta} = \frac{1}{2} \left[\left\langle A(\mathbf{l}, z_\alpha) B(\mathbf{l}', z_\beta) \right\rangle + \left\langle A(\mathbf{l}, z_\beta) B(\mathbf{l}', z_\alpha) \right\rangle \right], \quad (2.30)$$

to shorten our notation.

Noting $\tilde{\gamma}_{\mathrm{B}}^{(1)}(\mathbf{l}) = 0$ and $\langle \tilde{\gamma}_{\mathrm{E}}(\mathbf{l})\tilde{\gamma}_{\mathrm{B}}(\mathbf{l}')\rangle = 0$, we can expand equation (2.28) to $\mathcal{O}(\Phi^4)$:

$$
\begin{aligned}
\Big\langle \tilde{g}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\ \tilde{g}_{\mathrm{E}}(\mathbf{l}', z_\beta)\Big\rangle &= C^{(11)}_{\tilde{\gamma}_{\mathrm{E}}}(l; z_\alpha, z_\beta) + \underbrace{\Delta C^{(12)}_{\tilde{g}_{\mathrm{E}}}(l; z_\alpha, z_\beta)}_{\mathcal{O}(\Phi^3)\text{ reduced shear}} + \underbrace{\Delta C^{(13)}_{\tilde{\gamma}_{\mathrm{E}}}(l; z_\alpha, z_\beta) + \Delta C^{(22)}_{\tilde{\gamma}_{\mathrm{E}}}(l; z_\alpha, z_\beta)}_{\mathcal{O}(\Phi^4)\text{ shear}} \\
&\quad + \underbrace{\Delta C^{(13)}_{\tilde{g}_{\mathrm{E}}}(l; z_\alpha, z_\beta) + \Delta C^{(22)}_{\tilde{g}_{\mathrm{E}}}(l; z_\alpha, z_\beta)}_{\mathcal{O}(\Phi^4)\text{ reduced shear}} \\
&= C^{(11)}_{\tilde{\gamma}_{\mathrm{E}}}(l; z_\alpha, z_\beta) + 2\left\langle \tilde{\gamma}^{(1)}_1(\mathbf{l})\ T_I(\mathbf{l}')\left(\tilde{\gamma}^{(1)}_I * \tilde{\kappa}^{(1)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta} + 2\left\langle \tilde{\gamma}^{(1)}_1(\mathbf{l})\ T_I(\mathbf{l}')\tilde{\gamma}^{(3)}_I(\mathbf{l}')\right\rangle_{\alpha\beta} \\
&\qquad + \left\langle \tilde{\gamma}^{(2)}_1(\mathbf{l})\ T_I(\mathbf{l}')\tilde{\gamma}^{(2)}_I(\mathbf{l}')\right\rangle_{\alpha\beta} \\
&\quad + 2\left\{\left\langle \tilde{\gamma}^{(1)}_1(\mathbf{l})\ T_I(\mathbf{l}')\left(\tilde{\gamma}^{(1)}_I * \tilde{\kappa}^{(2)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta} + \left\langle \tilde{\gamma}^{(1)}_1(\mathbf{l})\ T_I(\mathbf{l}')\left(\tilde{\gamma}^{(2)}_I * \tilde{\kappa}^{(1)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta}\right. \\
&\qquad \left. + \left\langle \tilde{\gamma}^{(1)}_1(\mathbf{l})\ T_I(\mathbf{l}')\left(\tilde{\gamma}^{(1)}_I * \tilde{\kappa}^{(1)} * \tilde{\kappa}^{(1)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta}\right\} \\
&\quad + \left\{2\left\langle \tilde{\gamma}^{(2)}_1(\mathbf{l})\ T_I(\mathbf{l}')\left(\tilde{\gamma}^{(1)}_I * \tilde{\kappa}^{(1)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta} + \left\langle \left(\tilde{\gamma}^{(1)}_1 * \tilde{\kappa}^{(1)}\right)(\mathbf{l})\ T_I(\mathbf{l}')\left(\tilde{\gamma}^{(1)}_I * \tilde{\kappa}^{(1)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta}\right\},
\end{aligned}
\tag{2.31}
$$

$$
\begin{aligned}
\Big\langle \tilde{g}_{\mathrm{B}}(\mathbf{l}, z_\alpha)\ \tilde{g}_{\mathrm{B}}(\mathbf{l}', z_\beta)\Big\rangle &= \Delta C^{(22)}_{\tilde{\gamma}_{\mathrm{B}}}(l; z_\alpha, z_\beta) + \Delta C^{(22)}_{\tilde{g}_{\mathrm{B}}}(l; z_\alpha, z_\beta) \\
&= \left\langle \tilde{\gamma}^{(2)}_2(\mathbf{l})\ \epsilon_{IJ} T_I(\mathbf{l}')\tilde{\gamma}^{(2)}_J(\mathbf{l}')\right\rangle_{\alpha\beta} + 2\left\langle \tilde{\gamma}^{(2)}_2(\mathbf{l})\ \epsilon_{IJ} T_I(\mathbf{l}')\left(\tilde{\gamma}^{(1)}_J * \tilde{\kappa}^{(1)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta} \\
&\quad + \left\langle \left(\tilde{\gamma}^{(1)}_2 * \tilde{\kappa}^{(1)}\right)(\mathbf{l})\ \epsilon_{IJ} T_I(\mathbf{l}')\left(\tilde{\gamma}^{(1)}_J * \tilde{\kappa}^{(1)}\right)(\mathbf{l}')\right\rangle_{\alpha\beta},
\end{aligned}
\tag{2.32}
$$

where we have omitted terms such as $\Delta C^{(12)}_{\tilde{\gamma}_{\mathrm{E}}}$ which vanish under the Limber approximation.

2.3.1 Multiple-Deflection Shear Corrections

The shear-only corrections come in two flavors: the "22" (second-order–second-order) terms and the "13" terms. The "12" terms are mathematically of order Φ^3, and hence one might expect them to be present if the matter bispectrum is non-zero. However, they vanish in the Limber approximation due to the $W(\chi', \chi)$ factor in equation (2.19), which is zero whenever $\chi' = \chi$.

The "22" B-mode shear correction can be written as

$$\begin{aligned}\left\langle\tilde{\gamma}^{(2)}_{\mathrm{B}}(\mathbf{l},z_\alpha)\tilde{\gamma}^{(2)}_{\mathrm{B}}(\mathbf{l}'',z_\beta)\right\rangle &= 2\epsilon_{IJ}T_I(\mathbf{l})\int\frac{\mathrm{d}^2l'}{(2\pi)^2}\,[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')]\,ll'G_J(\mathbf{l},\mathbf{l}')\;2\epsilon_{HK}T_H(\mathbf{l}'')\\ &\times\int\frac{\mathrm{d}^2l'''}{(2\pi)^2}\,[\mathbf{l}'''\cdot(\mathbf{l}''-\mathbf{l}''')]\,l''l'''G_K(\mathbf{l}'',\mathbf{l}''')\left\langle\mathfrak{M}\,(\mathbf{l}',\mathbf{l}-\mathbf{l}';z_\alpha)\,\mathfrak{M}\left(\mathbf{l}''',\mathbf{l}''-\mathbf{l}''';z_\beta\right)\right\rangle\\ &= (2\pi)^2\delta(\mathbf{l}+\mathbf{l}'')4l^2\int\frac{\mathrm{d}^2l'}{(2\pi)^2}\,(l'\sin\phi_{l'})^2\,[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')]^2\,M\left(l',|\mathbf{l}-\mathbf{l}'|;z_\alpha,z_\beta\right),\end{aligned} \tag{2.33}$$

where we have used equations (2.14, 2.22, 2.26) and $\phi_l = 0$ repeatedly. By comparison with equation (2.20) one can see that $\Delta C^{(22)}_{\tilde{\gamma}_{\mathrm{B}}} = \Delta C^{(22)}_{\tilde{\omega}}$. Similarly,

$$\Delta C^{(22)}_{\tilde{\gamma}_{\mathrm{E}}}(l;z_\alpha,z_\beta) = 4l^2\int\frac{\mathrm{d}^2\mathbf{l}'}{(2\pi)^2}\,(l'\cos\phi_{l'})^2\,[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')]^2\,M\left(l',|\mathbf{l}-\mathbf{l}'|;z_\alpha,z_\beta\right), \tag{2.34}$$

and

$$\Delta C^{(22)}_{\tilde{\kappa}}(l;z_\alpha,z_\beta) = 4\int\frac{\mathrm{d}^2\mathbf{l}'}{(2\pi)^2}\,(\mathbf{l}\cdot\mathbf{l}')^2\,[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')]^2\,M\left(l',|\mathbf{l}-\mathbf{l}'|;z_\alpha,z_\beta\right)\;. \tag{2.35}$$

The integrals in equations (2.34, 2.35) are dominated by angular scales corresponding to the peak of the matter power spectrum, which is at scales much larger than those typically probed by lensing: If we define $\mathbf{l}_{\mathrm{c}} = \mathbf{l}-\mathbf{l}'$, then for small l_{c} (compared to l of lensing experiments) the contribution to these integrals scales as $\int \mathrm{d}^2\mathbf{l}_{\mathrm{c}}l_{\mathrm{c}}^2\cos^2(\mathbf{l},\mathbf{l}_{\mathrm{c}})M(l,l_{\mathrm{c}};z_\alpha,z_\beta)$. Assuming an effective power-law index $n_{\mathrm{s}}^{\mathrm{eff}}$ for the nonlinear matter power spectrum $P_{\delta,\mathrm{n}l}(k)$, the l_{c}-dependence of $M(l,l_{\mathrm{c}};z_\alpha,z_\beta)$ scales as $l_c^{n_{\mathrm{s}}^{\mathrm{eff}}-4}$. So the contribution to the integral per logarithmic range in l_{c} scales as $l_{\mathrm{c}}^{n_{\mathrm{s}}^{\mathrm{eff}}}$, which is dominated by scales corresponding to the peak of the matter power spectrum.

The "13" correction in principle has three parts: those arising from the 3A, 3B, and 3C terms of equation (2.8). Let us consider the 3B term first. The expectation value of the product of two Fourier modes is

$$\begin{aligned}\langle\psi^{(1)}_{ab}(\mathbf{l},z_\alpha)\psi^{(3B)}_{ij}(\mathbf{L},z_\beta)\rangle \;=\; & 16\int_0^{\chi_\alpha}\mathrm{d}\chi\int_0^{\chi_\beta}\mathrm{d}\chi_1\int_0^{\chi_1}\mathrm{d}\chi_1'\int_0^{\chi_1}\mathrm{d}\chi_1''\int\frac{\mathrm{d}^2\mathbf{L}'}{(2\pi)^2}\int\frac{\mathrm{d}^2\mathbf{L}''}{(2\pi)^2}\\ &\times W(\chi,\chi_\alpha)W(\chi_1,\chi_\beta)W(\chi_1',\chi_1)W(\chi_1'',\chi_1)\\ &\times l_al_bL'_cL'_jL''_d(\mathbf{L}-\mathbf{L}'-\mathbf{L}'')_i(\mathbf{L}-\mathbf{L}'-\mathbf{L}'')_c(\mathbf{L}-\mathbf{L}'-\mathbf{L}'')_d\\ &\times\langle\tilde{\phi}(\mathbf{l};\chi)\tilde{\phi}(\mathbf{L}-\mathbf{L}'-\mathbf{L}'';\chi_1)\tilde{\phi}(\mathbf{L}';\chi_1')\tilde{\phi}(\mathbf{L}'';\chi_1'')\rangle\;.\end{aligned} \tag{2.36}$$

In the Limber approximation, the only nonvanishing contraction is at $\chi = \chi_1$ and $\chi_1' = \chi_1''$. The

$\delta_{\rm D}$-functions then enforce $\mathbf{L}'_1 = -\mathbf{L}''_1$ and $\mathbf{L} = -\mathbf{l}$. We thus find

$$\langle \psi^{(1)}_{ab}(\mathbf{l}, z_\alpha)\psi^{(3B)}_{ij}(\mathbf{L}, z_\beta)\rangle = (2\pi)^2\delta_{\rm D}(\mathbf{l}+\mathbf{L})16\int_0^{\min(\chi_\alpha,\chi_\beta)} {\rm d}\chi \int_0^{\chi} {\rm d}\chi'_1 \frac{W(\chi,\chi_\alpha)W(\chi,\chi_\beta)W^2(\chi'_1,\chi)}{\chi^2{\chi'_1}^2}$$
$$\times \int \frac{{\rm d}^2\mathbf{L}'}{(2\pi)^2} P_\Phi\left(l/\chi; z(\chi)\right) P_\Phi\left(L'/\chi'_1; z(\chi'_1)\right) l_a l_b L'_c L'_j L'_d l_i l_c l_d. \qquad (2.37)$$

The integrand is odd under $\mathbf{L}' \to -\mathbf{L}'$, and hence the “13B” correction to the shear power spectrum vanishes.

The “13C” correction is zero because the restriction $\chi'' < \chi' < \chi$ in equation (2.8) implies that there are no allowed contractions within the independent lens plane approximation. This leaves us with the “13A” correction, which is similar to “13B,” except with the replacement $L'_j \to l_j$. The choice $\mathbf{l}\|\hat{\mathbf{x}}$ implies that the only nonvanishing component of “13A” is $\langle\psi^{(1)}_{11}(\mathbf{l}, z_\alpha)\psi^{(3A)}_{11}(\mathbf{L}, z_\beta)\rangle$. Hence we find

$$\Delta C^{(13)}_{\tilde{\gamma}_{\rm E}}(l; z_\alpha, z_\beta) = \Delta C^{(13A)}_{\tilde{\gamma}_{\rm E}}(l; z_\alpha, z_\beta) = \Delta C^{(13A)}_{\tilde{\kappa}}(l; z_\alpha, z_\beta) = -4l^4 \int \frac{{\rm d}^2\mathbf{l}'}{(2\pi)^2} \left(\mathbf{l}\cdot\mathbf{l}'\right)^2 M(l, l'; z_\alpha, z_\beta)\,. \qquad (2.38)$$

There is no “13” B-mode shear or rotation power spectrum because $\tilde{\gamma}^{(1)}_{\rm B}(\mathbf{l}, z_\alpha)$ and $\tilde{\omega}^{(1)}(\mathbf{l}, z_\alpha)$ vanish.

The dimensionless shear power spectrum, $\Delta^{2(11)}_{\tilde{\gamma}_{\rm E}}(l) = l(l+1)C^{(11)}_{\tilde{\gamma}_{\rm E}}(l)/(2\pi)^2$ scales as $\Delta^{2(11)}_{\tilde{\gamma}_{\rm E}}(l) \propto l^{n_{\rm s}^{\rm eff}+2}$, while the corrections $\Delta^{2(13)}_{\tilde{\gamma}_{\rm E}}(l)$ and $\Delta^{2(22)}_{\tilde{\gamma}_{\rm E}}(l)$ scale as $l^{n_{\rm s}^{\rm eff}+4}$. The main contribution to these corrections at large l is the bulk deflection on small scales by large wavelength density perturbations which causes only small local distortions. Thus the “22” and “13” terms largely cancel, similar to the perturbative calculation of the one-loop correction to the density power spectrum (e.g., Vishniac 1983). As these corrections diverge for large l and have opposite sign, their numerical difference needs to be evaluated carefully.[2]

The dotted lines in figure 2.1 illustrate their magnitude for $z_\alpha = z_\beta = 1$ using the fitting formula of Smith et al. (2003b) for the nonlinear matter power spectrum with the transfer function from Efstathiou, Bond and White (1992) for the numerical integration. Here the combined E-mode correction is negative at small l and positive for $l \gtrsim 4200$. These corrections are at least 4 orders of magnitude smaller that the linear theory result $C^{(11)}_{\tilde{\gamma}_{\rm E}}$.

Note that unlike the results of Cooray and Hu (2002), our calculations agree with the expected

---

[2] Apply a variable transform $\mathbf{l}'' = \mathbf{l} - \mathbf{l}'$ to $\Delta C^{(22)}_{\tilde{\gamma}_{\rm E}}$ and cancel diverging contributions at $\mathbf{l}''$ by rewriting the integral as $\Delta C^{(22)}_{\tilde{\gamma}_{\rm E}} + \Delta C^{(13)}_{\tilde{\gamma}_{\rm E}} = 4\int\frac{{\rm d}^2\mathbf{l}''}{(2\pi)^2}\left(\mathbf{l}\cdot(\mathbf{l}''+\mathbf{l})\right)^2\left(\mathbf{l}''\cdot(\mathbf{l}''+\mathbf{l})\right)^2\left(M(|\mathbf{l}+\mathbf{l}''|, l''; z_\alpha, z_\beta) - M(l, l''; z_\alpha, z_\beta)\right) + \int\frac{{\rm d}^2\mathbf{l}''}{(2\pi)^2}\left(\left(\mathbf{l}\cdot(\mathbf{l}''+\mathbf{l})\right)^2\left(\mathbf{l}''\cdot(\mathbf{l}''+\mathbf{l})\right)^2 - l^4\left(\mathbf{l}\cdot\mathbf{l}''\right)^2\right)M(l, l''; z_\alpha, z_\beta)$, where the azimuthal integration of the second term can be done analytically.

equivalence between the tangential shear and convergence (cf. equations (2.34, 2.35, 2.38)), as well as between cross shear and rotation power spectra (cf. discussion after equations (2.33, 2.38)).

2.3.2 Reduced Shear Corrections

The same methodology used for the corrections to the shear power spectra can also be used to compute the reduced shear terms in equation (2.31). Corrections to the reduced shear power spectra which combine second-order and first-order distortions contribute through two Wick contractions, for example,

$$
\begin{aligned}
\left\langle \tilde{\gamma}^{(2)}_{\mathrm{B}}(\mathbf{l},\mathbf{z}_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{B}}(\mathbf{l}'',\mathbf{z}_\beta)\right\rangle = &-2\epsilon_{IJ}T_I(\mathbf{l})\int\frac{\mathrm{d}^2l'}{(2\pi)^2}\left[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')\right] ll' G_J(\mathbf{l},\mathbf{l}')\epsilon_{HK}T_H(\mathbf{l}'') \\
&\times\int\frac{\mathrm{d}^2l'''}{(2\pi)^2}l'''^2 T_K(\mathbf{l}''-\mathbf{l}''')|\mathbf{l}''-\mathbf{l}'''|^2\int_0^{\chi_\beta}\mathrm{d}\chi'' W(\chi'',\chi_\beta) \\
&\times\int_0^{\chi_\beta}\mathrm{d}\chi''' W(\chi''',\chi_\beta)\left\langle \mathfrak{M}\left(\mathbf{l}',\mathbf{l}-\mathbf{l}';z_\alpha\right)\tilde{\phi}\left(\mathbf{l}''';\chi''\right)\tilde{\phi}\left(\mathbf{l}''-\mathbf{l}''';\chi'''\right)\right\rangle \\
= &-(2\pi)^2\delta_{\mathrm{D}}(\mathbf{l}+\mathbf{l}'')\int\frac{\mathrm{d}^2l'}{(2\pi)^2}\left[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')\right]|\mathbf{l}-\mathbf{l}'|^2 ll'^3\sin(\phi_{l'}) \\
&\times\int_0^{\chi_\beta}\mathrm{d}\chi\frac{W(\chi,\chi_\alpha)W(\chi,\chi_\beta)}{\chi^2}\int_0^{\chi}\mathrm{d}\chi'\frac{W(\chi',\chi_\beta)W(\chi',\chi)}{\chi'^2} \\
&\times 2\Big\{\sin(2\phi_{l'})+\sin(2\phi_{\mathbf{l}-\mathbf{l}'})\Big\}\,P_\Phi(l'/\chi;z(\chi))P_\Phi(|\mathbf{l}-\mathbf{l}'|/\chi';z(\chi'))\,,
\end{aligned}
\tag{2.39}
$$

where we have used $\phi_l = 0$ and $\epsilon_{IJ}T_I(\mathbf{l}')T_J(\mathbf{l}'') = \sin(2\phi_{l''} - 2\phi_{l'})$.

Corrections to the reduced shear power spectra which combine only first-order distortions contribute through all Wick contractions plus a connected contribution, for example,

$$
\begin{aligned}
\left\langle\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{B}}(\mathbf{l},z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{B}}(\mathbf{l}'',z_\beta)\right\rangle = &\,\epsilon_{IJ}T_I(\mathbf{l})\int\frac{\mathrm{d}^2l'}{(2\pi)^2}T_J(\mathbf{l}')\epsilon_{HK}T_H(\mathbf{l}'')\int\frac{\mathrm{d}^2l'''}{(2\pi)^2}T_K(\mathbf{l}''') \\
&\times\left\langle\tilde{\kappa}^{(1)}(\mathbf{l}',z_\alpha)\tilde{\kappa}^{(1)}(\mathbf{l}-\mathbf{l}',z_\alpha)\tilde{\kappa}^{(1)}(\mathbf{l}''',z_\beta)\tilde{\kappa}^{(1)}(\mathbf{l}''-\mathbf{l}''',z_\beta)\right\rangle \\
= &\,(2\pi)^2\delta_{\mathrm{D}}(\mathbf{l}+\mathbf{l}'')\int\frac{\mathrm{d}^2l'}{(2\pi)^2}\sin(2\phi_{l'}) \\
&\times\Big\{(\sin(2\phi_{l'})+\sin(2\phi_{\mathbf{l}'-\mathbf{l}}))C^{(11)}_{\tilde{\gamma}_{\mathrm{E}}}(l';z_\alpha,z_\beta)C^{(11)}_{\tilde{\gamma}_{\mathrm{E}}}(\mathbf{l}'-\mathbf{l};z_\alpha,z_\beta) \\
&+\int\frac{\mathrm{d}^2l'''}{(2\pi)^2}\sin(2\phi_{l'''})T_\kappa(\mathbf{l}',\mathbf{l}-\mathbf{l}',\mathbf{l}''',-\mathbf{l}-\mathbf{l}''';z_\alpha,z_\alpha,z_\beta,z_\beta)\Big\},
\end{aligned}
\tag{2.40}
$$

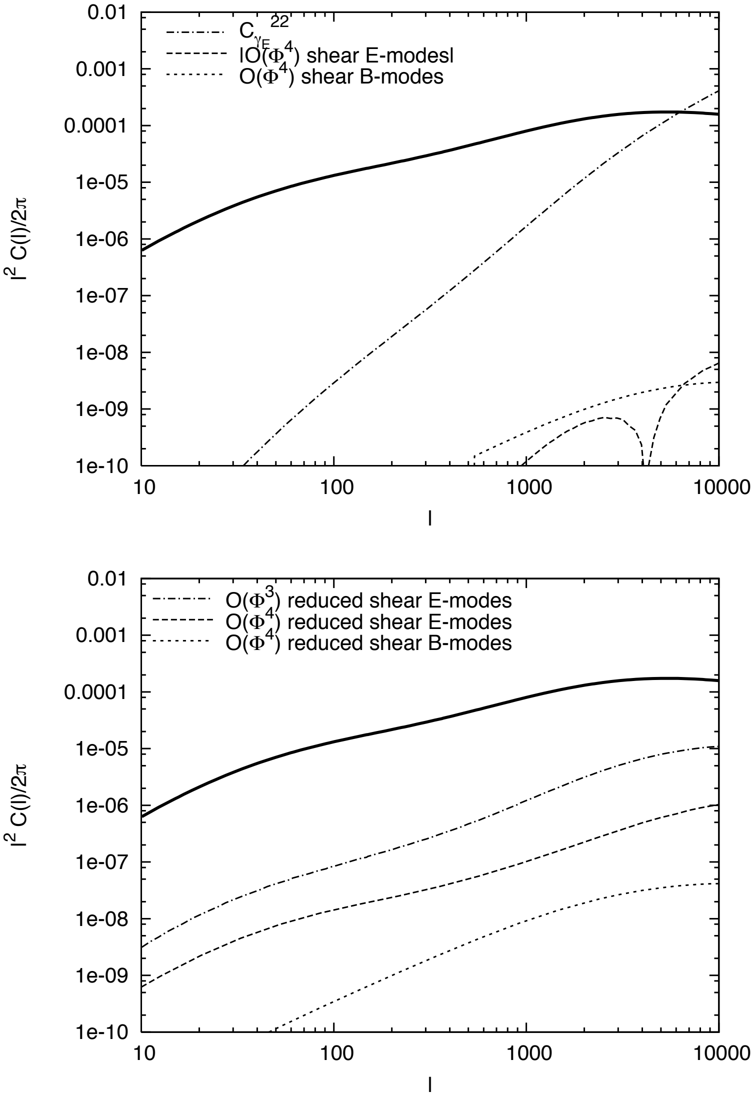


Figure 2.1 Linear-order shear power spectrum (thick solid line; equation (2.13)) and corrections up to $O(\Phi^4)$.

Top. The dashed/short dashed lines show the fourth-order corrections to the E/B-mode shear power spectra that arise from relaxing the Born approximation and including lens-lens coupling in the calculation of the shear (section 2.3.1; cf. Cooray and Hu (2002)). The E-mode correction is negative at small l and positive for $l \gtrsim 4200$. The dashed-dotted line illustrates term $C_{\gamma_E}^{(22)}$ (cf. equation (2.34)) which contributes to the E-mode shear correction, the divergency is cancelled by equation (2.38).

Bottom. The dashed/short dashed lines show the combined fourth-order corrections to the reduced shear E/B-mode power spectra (section 2.3.2, table 2.1). The dashed-dotted line shows the third-order correction to the reduced shear E-mode power spectrum.

We assume a source redshift $z_\alpha = z_\beta = 1$ and use the transfer function from Efstathiou, Bond and White (1992), the fitting formula of Smith et al. (2003b) for the nonlinear matter power spectrum, and the fitting formula of Scoccimarro and Couchman (2001) for the nonlinear matter bispectrum. This figure assumes a flat ΛCDM cosmology with $(\Omega_\mathrm{m}, \Omega_\mathrm{b}, \sigma_8, h, n) = (0.3, 0.05, 0.9, 0.7, 1)$ to enable comparison with previous calculations.

Table 2.1 Contributions of order $O(\Phi^3)$ and $O(\Phi^4)$ to the reduced shear E-mode power spectrum

Type	Term	Contribution to $C_{\tilde{g}_{\mathrm{E}}}(l)$
$O(\Phi^3)$ reduced shear	$\left\langle \tilde{\gamma}^{(1)}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$2\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\cos(2\phi_{l'})B_\kappa(\mathbf{l},\mathbf{l}',-\mathbf{l}-\mathbf{l}';z_\alpha,z_\beta,z_\beta)$ cf. (Shapiro 2009)
shear with multiple deflections	$\left\langle \tilde{\gamma}^{(2)}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\tilde{\gamma}^{(2)}_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$4l^2\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\left(l'\cos(\phi_{l'})\right)^2\left(\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')\right)^2 M\left(l', \vert\mathbf{l}-\mathbf{l}'\vert;\chi\right)$
	$\left\langle \tilde{\gamma}^{(1)}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\tilde{\gamma}^{(3)}_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$-4l^4\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}(\mathbf{l}\cdot\mathbf{l}')^2 M(l, l'; z_\alpha, z_\beta)$
reduced shear with multiple deflections	$\left\langle \tilde{\gamma}^{(1)}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(2)}\right)_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$-2\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}(\mathbf{l}\cdot\mathbf{l}')^2\, l^2 l'^2\cos(2\phi_{l'})$ $\times\Big\{\int_0^{\chi_\beta}\mathrm{d}\chi\frac{W(\chi,\chi_\alpha)W(\chi,\chi_\beta)}{\chi^2}\int_0^{\chi}\mathrm{d}\chi'\frac{W(\chi',\chi_\beta)W(\chi',\chi)}{\chi'^2}P_\Phi(l/\chi;z(\chi))P_\Phi(l'/\chi';z(\chi'))$ $+\int_0^{\chi_\beta}\mathrm{d}\chi\frac{W^2(\chi,\chi_\beta)}{\chi^2}\int_0^{\chi}\mathrm{d}\chi'\frac{W(\chi',\chi_\alpha)W(\chi',\chi)}{\chi'^2}P_\Phi(l'/\chi;z(\chi))P_\Phi(l/\chi';z(\chi'))\Big\}$
	$\left\langle \tilde{\gamma}^{(1)}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(2)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$-2\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\mathbf{l}\cdot\mathbf{l}'\, l^2 l'^2\vert\mathbf{l}+\mathbf{l}'\vert$ $\times\Big\{l\cos(\phi_{\mathbf{l}+\mathbf{l}'})\int_0^{\chi_\beta}\mathrm{d}\chi\frac{W(\chi,\chi_\alpha)W(\chi,\chi_\beta)}{\chi^2}\int_0^{\chi}\mathrm{d}\chi'\frac{W(\chi',\chi_\beta)W(\chi',\chi)}{\chi'^2}P_\Phi(l/\chi;z(\chi))P_\Phi(l'/\chi';z(\chi'))$ $+l'\cos(\phi_{\mathbf{l}+\mathbf{l}'}+\phi_{l'})\int_0^{\chi_\beta}\mathrm{d}\chi\frac{W^2(\chi,\chi_\beta)}{\chi^2}\int_0^{\chi}\mathrm{d}\chi'\frac{W(\chi',\chi_\alpha)W(\chi',\chi)}{\chi'^2}P_\Phi(l'/\chi;z(\chi))P_\Phi(l/\chi';z(\chi'))\Big\}$
	$\left\langle \tilde{\gamma}^{(2)}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$-2\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\left[\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}')\right]\vert\mathbf{l}-\mathbf{l}'\vert^2 l l'^3\cos(\phi_{l'})\left\{\cos(2\phi_{l'})+\cos(2\phi_{\mathbf{l}-\mathbf{l}'})\right\}$ $\times\int_0^{\chi_\beta}\mathrm{d}\chi\frac{W(\chi,\chi_\alpha)W(\chi,\chi_\beta)}{\chi^2}\int_0^{\chi}\mathrm{d}\chi'\frac{W(\chi',\chi_\beta)W(\chi',\chi)}{\chi'^2}P_\Phi(l'/\chi;z(\chi))P_\Phi(\vert\mathbf{l}-\mathbf{l}'\vert/\chi';z(\chi'))$
reduced shear with single deflection	$\left\langle \tilde{\gamma}^{(1)}_{\mathrm{E}}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$C^{(11)}_{\tilde{\gamma}_{\mathrm{E}}}(l;z_\alpha,z_\beta)\sigma^2_{\tilde{\gamma}_{\mathrm{E}}}(z_\beta)+\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\cos(2\phi_{l'})\int\frac{\mathrm{d}^2 l'''}{(2\pi)^2}T_\kappa(\mathbf{l},\mathbf{l}',\mathbf{l}''',-\mathbf{l}-\mathbf{l}-\mathbf{l}''';z_\alpha,z_\beta,z_\beta,z_\beta)$
	$\left\langle\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\mathrm{E}}(\mathbf{l}'', z_\beta)\right\rangle$	$\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\cos(2\phi_{l'})\Big\{\left(\cos(2\phi_{l'})+\cos(2\phi_{\mathbf{l}'-\mathbf{l}})\right)C^{(11)}_{\tilde{\gamma}_{\mathrm{E}}}(l';z_\alpha,z_\beta)C^{(11)}_{\tilde{\gamma}_{\mathrm{E}}}(\vert\mathbf{l}'-\mathbf{l}\vert;z_\alpha,z_\beta)$ $+\int\frac{\mathrm{d}^2 l'''}{(2\pi)^2}\cos(2\phi_{l'''})T_\kappa(\mathbf{l}',\mathbf{l}-\mathbf{l}',\mathbf{l}''',-\mathbf{l}-\mathbf{l}''';z_\alpha,z_\alpha,z_\beta,z_\beta)\Big\}$

Table 2.2 Correction of order $O(\Phi^4)$ to the reduced shear B-mode power spectrum

Type	Term	Contribution to $C_{\tilde{g}_\mathrm{B}}(l)$
shear with multiple deflections	$\left\langle \tilde{\gamma}^{(2)}_\mathrm{B}(\mathbf{l}, z_\alpha)\tilde{\gamma}^{(2)}_\mathrm{B}(\mathbf{l''}, z_\beta)\right\rangle$	$4l^2 \int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\,(l' \sin(\phi_{l'}))^2\,(\mathbf{l'}\cdot(\mathbf{l}-\mathbf{l'}))^2\,M\,(l', \lvert\mathbf{l}-\mathbf{l'}\rvert;\chi)$
reduced shear with multiple deflections	$\left\langle \tilde{\gamma}^{(2)}_\mathrm{B}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_\mathrm{B}(\mathbf{l''}, z_\beta)\right\rangle$	$-2\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\,[\mathbf{l'}\cdot(\mathbf{l}-\mathbf{l'})]\,\lvert\mathbf{l}-\mathbf{l'}\rvert^2 l l'^3 \sin(\phi_{l'})\,\{\sin(2\phi_{l'}) + \sin(2\phi_{\mathbf{l}-\mathbf{l'}})\}$ $\times\int_0^{\chi_\beta}\mathrm{d}\chi\frac{W(\chi,\chi_\alpha)W(\chi,\chi_\beta)}{\chi^2}\int_0^{\chi}\mathrm{d}\chi'\,\frac{W(\chi',\chi_\beta)W(\chi',\chi)}{\chi'^2}P_\Phi(l'/\chi;z(\chi))P_\Phi(\lvert\mathbf{l}-\mathbf{l'}\rvert/\chi';z(\chi'))$
reduced shear with single deflection	$\left\langle\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_\mathrm{B}(\mathbf{l}, z_\alpha)\left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_\mathrm{B}(\mathbf{l''}, z_\beta)\right\rangle$	$\int\frac{\mathrm{d}^2 l'}{(2\pi)^2}\sin(2\phi_{l'})\Big\{(\sin(2\phi_{l'}) + \sin(2\phi_{\mathbf{l'}-\mathbf{l}}))C^{(11)}_{\tilde{\gamma}_\mathrm{E}}(l';z_\alpha,z_\beta)C^{(11)}_{\tilde{\gamma}_\mathrm{E}}(\lvert\mathbf{l'}-\mathbf{l}\rvert;z_\alpha,z_\beta)$ $+\int\frac{\mathrm{d}^2 l'''}{(2\pi)^2}\sin(2\phi_{l'''})T_\kappa(\mathbf{l'},\mathbf{l}-\mathbf{l'},\mathbf{l'''},-\mathbf{l}-\mathbf{l'''};z_\alpha,z_\alpha,z_\beta,z_\beta)\Big\}$

where we have omitted a term which only contributes to the $l = 0$ mode, and where T_κ is the lensing tomography convergence trispectrum (Cooray and Hu 2001) which we model with the halo model of large-scale structure (e.g., Seljak 2000, Cooray and Sheth 2002) as summarized in appendix A. Here, the Gaussian contribution, which is the dominant term on relevant angular scales, is simply a convolution of the standard $O(\Phi^2)$ lensing tomography cross-spectra with some geometrical projection factors. Note that in the halo model framework the connected contribution to the B-mode spectrum is downweighted by the geometric projection factors, especially *one-halo* and *(13) two-halo* are strongly suppressed. The connected E-mode terms given in table 2.1 has opposite angular symmetry and the connected part starts to dominate the signal above $l \sim 8000$.

The analytic expressions for all contributions to the fourth-order tangential reduced shear cross-spectra are summarized in table 2.1. Figure 2.1 illustrates the numerical values of the different corrections. The fourth-order reduced shear corrections of the lensing E-mode power spectrum reach the percent level at small angular scales and hence may be relevant for future weak lensing experiments. Reduced shear generates a small amount of B-mode power, which is about 4 magnitudes smaller than the E-mode signal, and is less than the level of B-mode power generated by observational systematics.

2.3.3 Relation Between Ellipticities and Reduced Shear

The linear relation between some measure of image ellipticity and reduced shear (2.1) is only valid in the limit of very weak lensing ($\kappa \ll 1$, $|\gamma| \ll 1$). In general the relation between image ellipticity and reduced shear depends on the ellipticity measure under consideration. As an example we consider two definitions of the complex image ellipticity here:

$$\boldsymbol{\varepsilon} = \frac{1-r}{1+r} e^{2i\phi}, \tag{2.41}$$

and

$$\mathbf{e} = \frac{1-r^2}{1+r^2} e^{2i\phi}, \tag{2.42}$$

where $r \leq 1$ is the minor to major axis ratio of the image, and ϕ is the position angle of the major axis. The latter is frequently employed in observational studies (Bernstein and Jarvis 2002), the former is more of theoretical interest due to its simple transformation properties. The full relation between ellipticity and complex reduced shear $\mathbf{g} = g_1 + ig_2$ is given by

$$\boldsymbol{\varepsilon} = \frac{\boldsymbol{\varepsilon}^{(\mathrm{s})} + \mathbf{g}}{1 + \mathbf{g}^* \boldsymbol{\varepsilon}^{(\mathrm{s})}}, \quad \text{and} \quad \mathbf{e} = \frac{\mathbf{e}^{(\mathrm{s})} + 2\mathbf{g} + \mathbf{g}^2 \mathbf{e}^{(\mathrm{s})*}}{1 + |\mathbf{g}|^2 + 2\mathcal{R}\left(\mathbf{g}\mathbf{e}^{(\mathrm{s})*}\right)}, \tag{2.43}$$

where $\mathcal{R}(\mathbf{z})$ is the real part of a complex number $\mathbf{z}$, $\mathbf{e}^{(\mathrm{s})}$ and $\boldsymbol{\varepsilon}^{(\mathrm{s})}$are the intrinsic ellipticities of the source and where we only consider $|\gamma| < 1$, which is certainly true for cosmic shear. The linear relation $\langle \boldsymbol{\varepsilon} \rangle = \mathbf{g}$ is exact (Seitz and Schneider 1997), as can be shown using the residue theorem. In the second case, using a Taylor expansion (Schneider and Seitz 1995, Mandelbaum et al. 2006), the ellipticities can be written as

$$\langle \mathbf{e} \rangle = c_1 \mathbf{g} + c_3 |\mathbf{g}|^2 \mathbf{g} + O(g^5) \approx \left(2 - \mathrm{e}^{(\mathrm{s})^2}\right) \mathbf{g} + \left(-2 + 5\mathrm{e}^{(\mathrm{s})^2} - 3\mathrm{e}^{(\mathrm{s})^4}\right) |\mathbf{g}|^2 \mathbf{g}, \tag{2.44}$$

where $\mathrm{e}^{(\mathrm{s})}$ is the absolute value of the intrinsic ellipticity of the source galaxies. In the practical case of a distribution of intrinsic source ellipticities, one should replace the powers of $\mathrm{e}^{(\mathrm{s})}$ by their moments $\langle \mathrm{e}^{(\mathrm{s})\,n} \rangle$. Shear is typically estimated by taking the mean observed ellipticity $\langle \mathbf{e} \rangle$ and dividing by the response factor c_1. To $O(\Phi^4)$, this shear estimator reads

$$\hat{\mathbf{g}} = \frac{\langle \mathbf{e} \rangle}{c_1} = \mathbf{g} + \frac{c_3}{c_1} |\mathbf{g}|^2 \mathbf{g}. \tag{2.45}$$

The last term gives rise to one additional contribution to the power spectrum of $\hat{g}_{\mathrm{E}}$:

$$\begin{aligned}
2\frac{c_3}{c_1} \left\langle \tilde{\gamma}_{\mathrm{E}}^{(1)}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(1)} * \tilde{\gamma}^{(1)} * \tilde{\gamma}^{(1)}\right)_{\mathrm{E}}(\mathbf{l}', z_\beta) \right\rangle &= 2\frac{c_3}{c_1}\, \delta_{IJ} T_I(\mathbf{l}') \int \frac{\mathrm{d}^2 l''}{(2\pi)^2} \int \frac{\mathrm{d}^2 l'''}{(2\pi)^2} T_H(\mathbf{l}''') T_H(\mathbf{l}'') T_J(\mathbf{l}' - \mathbf{l}'' - \mathbf{l}''') \\
&\quad \times \left\langle \tilde{\kappa}^{(1)}(\mathbf{l}, z_\alpha) \tilde{\kappa}^{(1)}(\mathbf{l}''', z_\beta) \tilde{\kappa}^{(1)}(\mathbf{l}'', z_\beta) \tilde{\kappa}^{(1)}(\mathbf{l}' - \mathbf{l}'' - \mathbf{l}''', z_\beta) \right\rangle \\
&= (2\pi)^2 \delta_{\mathrm{D}}(\mathbf{l} + \mathbf{l}')\, 2\frac{c_3}{c_1} \int \frac{\mathrm{d}^2 l''}{(2\pi)^2} \Big\{ (2\cos^2(2\phi_{l''}) + 1) \\
&\quad \times C_{\tilde{\gamma}_{\mathrm{E}}}^{(11)}(l; z_\alpha, z_\beta) C_{\tilde{\gamma}_{\mathrm{E}}}^{(11)}(l''; z_\beta, z_\beta) \\
&\quad + \int \frac{\mathrm{d}^2 l'''}{(2\pi)^2} \cos(2\phi_{l''} - 2\phi_{l'''}) \cos(2\phi_{-\mathbf{l}-\mathbf{l}''-\mathbf{l}'''}) \\
&\quad \times T_\kappa(\mathbf{l}, \mathbf{l}'', \mathbf{l}''', -\mathbf{l} - \mathbf{l}'' - \mathbf{l}'''; z_\alpha, z_\beta, z_\beta, z_\beta) \Big\} \\
&= (2\pi)^2 \delta_{\mathrm{D}}(\mathbf{l} + \mathbf{l}')\, 2\frac{c_3}{c_1} \Big\{ 2C_{\tilde{\gamma}_{\mathrm{E}}}^{(11)}(l; z_\alpha, z_\beta) \sigma_{\tilde{\gamma}_{\mathrm{E}}}^2(z_\beta) \\
&\quad + \int \frac{\mathrm{d}^2 l'' \mathrm{d}^2 l'''}{(2\pi)^4} \cos(2\phi_{l''} - 2\phi_{l'''}) \cos(2\phi_{-\mathbf{l}-\mathbf{l}''-\mathbf{l}'''}) \\
&\quad \times T_\kappa(\mathbf{l}, \mathbf{l}'', \mathbf{l}''', -\mathbf{l} - \mathbf{l}'' - \mathbf{l}'''; z_\alpha, z_\beta, z_\beta, z_\beta) \Big\},
\end{aligned} \tag{2.46}$$

where we have performed the angular integration of the Gaussian contribution in the last step and introduced the shear dispersion

$$\sigma^2_{\tilde{\gamma}_{\rm E}}(z_\beta) = \int \frac{dl'}{2\pi} l' C^{(11)}_{\tilde{\gamma}_{\rm E}}(l'; z_\beta, z_\beta). \tag{2.47}$$

For the case of the $\boldsymbol{\varepsilon}$ ellipticity, linearity implies $c_1 = 1$ and $c_3 = 0$. In this case, the correction of equation (2.46) vanishes. For the case of the **e** ellipticity, we have

$$\frac{c_3}{c_1} = \frac{-2 + 5\mathrm{e}^{(\mathrm{s})^2} - 3\mathrm{e}^{(\mathrm{s})^4}}{2 - \mathrm{e}^{(\mathrm{s})^2}}. \tag{2.48}$$

The magnitude of this corrections for the **e** ellipticity with $\langle \mathrm{e}^{(\mathrm{s})2} \rangle^{1/2} = 0.6$ is illustrated in figure 2.2.

2.3.4 Lensing Bias Corrections

Galaxies are only selected for shear measurement if they are large enough and bright enough to measure their shape. As lensing changes the observed brightness and size of the lensed galaxies, the number of galaxies selected above some magnitude and size threshold is correlated with the lensing field (this is the well know magnification bias, and the size bias discussed in Schmidt et al. (2009b)). Neglecting source clustering, the normalized observed galaxy overdensity due to lensing magnification is given by Schmidt et al. (2008), Hui, Gaztañaga and Loverde (2007)

$$1+\delta_{\rm lens}(\mathbf{n}) = \frac{1 + q\kappa(\mathbf{n}) + C_1\kappa^2(\mathbf{n}) + C_2|\gamma|^2(\mathbf{n})}{1 + C_1\langle\kappa^2\rangle + C_2\langle|\gamma|^2\rangle} \approx 1+q\kappa(\mathbf{n})+C_1\left(\kappa^2(\mathbf{n}) - \left\langle\kappa^2\right\rangle\right)+C_2\left(|\gamma|^2(\mathbf{n}) - \left\langle|\gamma|^2\right\rangle\right), \tag{2.49}$$

where we expanded the magnification to second order,[3] and where $C_1 = q(q+1)/2$ and $C_2 = q/2$. The parameter q is determined by the slope of the luminosity and radius distribution of the sample galaxies and typically $q \sim 1 - 2$ (Schmidt et al. 2009b).

Hence the sampling of the shear field measured from galaxy pairs is modulated by the lensing magnification implying that the observed shear depends on the true shear and the galaxy overdensity

$$g^{\rm obs}_I(\mathbf{n}) = g_I(\mathbf{n})\left(1 + \delta_{\rm lens}(\mathbf{n})\right) \approx \gamma_i(\mathbf{n})\Big\{1 + \kappa(\mathbf{n}) + \kappa^2(\mathbf{n}) + q\kappa(\mathbf{n}) + q\kappa^2(\mathbf{n}) + C_1[\kappa^2(\mathbf{n}) - \langle\kappa^2\rangle] + C_2[|\gamma(\mathbf{n})|^2 - \langle|\gamma|^2\rangle]\Big\}. \tag{2.50}$$

---

[3] We note that $\langle\kappa\rangle = 0$. This is because by rotational symmetry the mean deflection angle $\langle d\rangle = 0$, and therefore its derivative $\left\langle\psi_{ij}\right\rangle = 0$.

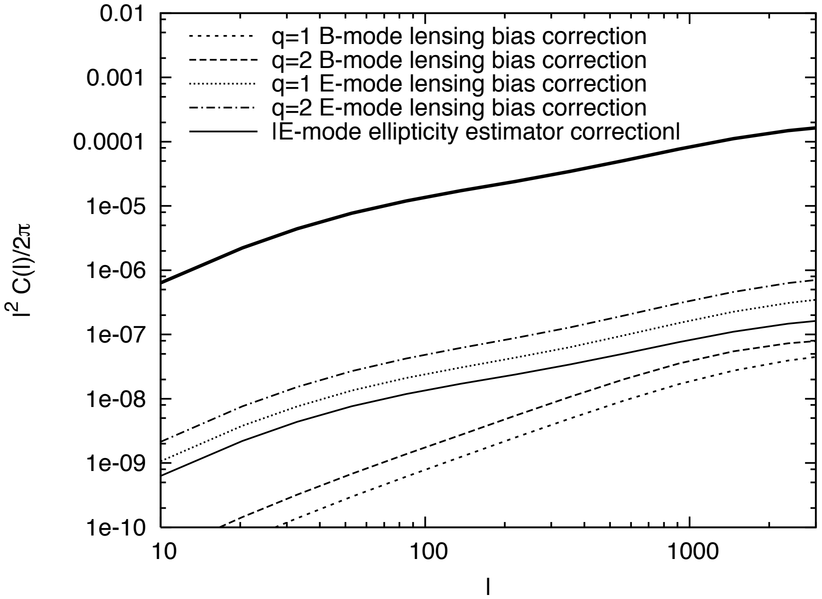


Figure 2.2 Linear-order shear power spectrum (thick solid line; equation (2.13)) and $O(\Phi^4)$ lensing bias and ellipticity estimator corrections.

The short dashed (dashed) lines show the lensing bias corrections to the B-mode shear power spectrum (equation 2.61) assuming $q = 1$ ($q = 2$). The dotted (dashed-dotted) lines show the lensing bias corrections to the E-mode shear power spectrum (equation 2.60) assuming $q = 1$ ($q = 2$).

The fine solid line illustrates the magnitude of the correction arising from the conversion between ellipticity and reduced shear equation (2.46) for the **e** ellipticity with $\langle e^{(s)2}\rangle^{1/2} = 0.6$. This correction is negative and its normalization depends on the distribution of source galaxies (see section 2.3.3 for details).

This figure uses the same cosmology and source redshifts as figure 2.1

The standard pair based estimator for the reduced shear correlation functions $\xi_{ab} = \langle g_a g_b \rangle$ then becomes (for details see Schmidt et al. 2009a)

$$\left\langle \hat{\xi}_{IJ}(\theta) \right\rangle = \left\langle \frac{1}{\mathcal{N}} g_I^{\mathrm{obs}}(\mathbf{n}) g_J^{\mathrm{obs}}(\mathbf{n}+\theta) \right\rangle, \tag{2.51}$$

where $\mathcal{N}$ is the observed number of galaxy pairs with separation θ relative to that expected for a random distribution; this is just the $\frac{DD}{RR}$ correlation function estimator (Peebles and Hauser 1974). For large-angle surveys, $\mathcal{N}$ converges to the correlation function,

$$\mathcal{N} \to 1 + \langle \delta_{\mathrm{lens}}(\mathbf{n}) \delta_{\mathrm{lens}}(\mathbf{n}+\theta) \rangle. \tag{2.52}$$

Therefore we may write

$$\left\langle \hat{\xi}_{IJ}(\theta) \right\rangle = \frac{\left\langle g_I^{\mathrm{obs}}(\mathbf{n}) g_J^{\mathrm{obs}}(\mathbf{n}+\theta) \right\rangle}{1 + \langle \delta_{\mathrm{lens}}(\mathbf{n}) \delta_{\mathrm{lens}}(\mathbf{n}+\theta) \rangle}. \tag{2.53}$$

This can be converted to products of correlation functions by conversion to a geometric series,

$$\left\langle \hat{\xi}_{IJ}(\theta) \right\rangle = \left\langle g_I^{\mathrm{obs}}(\mathbf{n}) g_J^{\mathrm{obs}}(\mathbf{n}+\theta) \right\rangle \sum_{\upsilon=0}^{\infty} (-1)^{\upsilon} \langle \delta_{\mathrm{lens}}(\mathbf{n}) \delta_{\mathrm{lens}}(\mathbf{n}+\theta) \rangle^{\upsilon}; \tag{2.54}$$

we then note that the υ term in this expansion is of order $O(\Phi^{2+2\upsilon})$. Since $\langle \hat{\xi}_{IJ}(\theta) \rangle$ is desired to $O(\Phi^4)$, it suffices to keep only the $\upsilon = 0$ and $\upsilon = 1$ terms. Moreover, in the $\upsilon = 1$ term, we only require the lowest-order expansion of the correlation function $\langle \delta_{\mathrm{lens}}(\mathbf{n}) \delta_{\mathrm{lens}}(\mathbf{n}+\theta) \rangle$, i.e.

$$\langle \delta_{\mathrm{lens}}(\mathbf{n}) \delta_{\mathrm{lens}}(\mathbf{n}+\theta) \rangle = q^2 \langle \kappa(\mathbf{n}) \kappa(\mathbf{n}+\theta) \rangle + O(\Phi^3). \tag{2.55}$$

We also need only the lowest-order expansion of $\left\langle g_I^{\mathrm{obs}}(\mathbf{n}) g_J^{\mathrm{obs}}(\mathbf{n}+\theta) \right\rangle$ in the $\upsilon = 1$ term, i.e., we can approximate it as $\langle \gamma_I(\mathbf{n}) \gamma_J(\mathbf{n}+\theta) \rangle$. Thus we reduce equation (2.53) to

$$\left\langle \hat{\xi}_{IJ}(\theta) \right\rangle \approx \left\langle g_I^{\mathrm{obs}}(\mathbf{n}) g_J^{\mathrm{obs}}(\mathbf{n}+\theta) \right\rangle - q^2 \langle \gamma_I(\mathbf{n}) \gamma_J(\mathbf{n}+\theta) \rangle \langle \kappa(\mathbf{n}) \kappa(\mathbf{n}+\theta) \rangle \ . \tag{2.56}$$

A straightforward generalization to cross-correlations between different redshift slices gives

$$\left\langle \hat{\xi}_{IJ}(\theta, z_\alpha, z_\beta) \right\rangle \approx \left\langle g_I^{\mathrm{obs}}(\mathbf{n}, z_\alpha) g_J^{\mathrm{obs}}(\mathbf{n}+\theta, z_\beta) \right\rangle - q^2 \left\langle \gamma_I(\mathbf{n}, z_\alpha) \gamma_J(\mathbf{n}+\theta, z_\beta) \right\rangle \left\langle \kappa(\mathbf{n}, z_\alpha) \kappa(\mathbf{n}+\theta, z_\beta) \right\rangle \ . \tag{2.57}$$

We now turn to practical computation. The terms involving $\left\langle g_I^{\mathrm{obs}}(\mathbf{n}, z_\alpha) g_J^{\mathrm{obs}}(\mathbf{n}+\theta, z_\beta) \right\rangle$ are all identical to terms that we have calculated previously, except with additional factors of q, q^2, C_1, and/or C_2, and hence present no new difficulties. The final subtraction term is the product of two expectation values and hence is different from terms that we have previously considered. This "product correction" can be evaluated by noting that its contribution to the observed correlation function is the product of the shear and convergence correlation functions. In Fourier space, this means that its contribution to the power spectrum is the convolution of the shear and convergence power spectra:

$$\Delta C^{\rm prod}_{\tilde{\gamma}_I\tilde{\gamma}_J}(\mathbf{l}) = -q^2 \int \frac{d^2 l'}{(2\pi)^2} C_{\gamma_I\gamma_J}(l') C_{\kappa\kappa}(|\mathbf{l}-\mathbf{l}'|), \tag{2.58}$$

where all power spectra carry the redshift indices z_α, z_β. Specializing to the case where $\mathbf{l}$ is along the x coordinate axis, and recalling that the E-mode shear and convergence power spectra are equal, we can then infer a contribution to the observed E-mode power spectrum

$$\Delta C^{\rm prod}_{\tilde{\gamma}_{\rm E}}(\mathbf{l}) = -q^2 \int \frac{d^2 l'}{(2\pi)^2} \cos^2(2\phi_{\mathbf{l}'})\, C^{(11)}_{\tilde{\gamma}_{\rm E}}(l') C^{(11)}_{\tilde{\gamma}_{\rm E}}(|\mathbf{l}-\mathbf{l}'|); \tag{2.59}$$

the B-mode contribution is similar except for the replacement $\cos^2 \to \sin^2$.

Similar to equations (2.28, 2.31), we now expand $\left\langle \tilde{g}^{\rm obs}_{\rm E/B}(\mathbf{l}) \tilde{g}^{\rm obs}_{\rm E/B}(\mathbf{l}'') \right\rangle$ to find the fourth-order power spectrum corrections $\Delta C^{\rm LB}_{\tilde{\gamma}_{\rm E/B}}$ which arise from lensing bias,

$$\begin{aligned}
\Delta C^{\rm LB}_{\tilde{\gamma}_{\rm E}}(l; z_\alpha, z_\beta) \;=\;& (2\pi)^2 \delta_{\rm D}(\mathbf{l}+\mathbf{l}'') \Big\langle\; 2q \left\langle \tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(2)}\right)_{\rm E}(\mathbf{l}'', z_\beta) \right\rangle \\
&+\; 2q \left[\left\langle \tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(2)} * \tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}'', z_\beta) \right\rangle + \left\langle \tilde{\gamma}^{(2)}_{\rm E}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}'', z_\beta) \right\rangle \right] \\
&+\; (2q + q^2) \left\langle \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}'', z_\beta) \right\rangle \\
&+\; 2C_2 \left[\left\langle \tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(1)} * \tilde{\gamma}^{(1)} * \tilde{\gamma}^{(1)}\right)_{\rm E}(\mathbf{l}'', z_\beta) \right\rangle_{\rm c} + C^{(11)}_{\tilde{\gamma}_{\rm E}}(l; z_\alpha, z_\beta) \sigma^2_{\tilde{\gamma}_{\rm E}}(z_\beta) \right] \\
&+\; 2q \left\langle \tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}'', z_\beta) \right\rangle + 2C_1 \left\langle \tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l}, z_\alpha) \left(\tilde{\gamma}^{(1)} * \tilde{\kappa}^{(1)} * \tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}'', z_\beta) \right\rangle_{\rm c} \Big\rangle_{\alpha\beta} \\
&-\; q^2 \int \frac{{\rm d}^2 l'}{(2\pi)^2} \cos^2(2\phi_{\mathbf{l}'}) C^{(11)}_{\tilde{\gamma}_{\rm E}}(l'; z_\alpha, z_\beta) C^{(11)}_{\tilde{\gamma}_{\rm E}}(|\mathbf{l}' - \mathbf{l}|; z_\alpha, z_\beta)\,,
\end{aligned} \tag{2.60}$$

and

$$\begin{aligned}\Delta C^{\rm LB}_{\tilde{\gamma}_{\rm B}}(l;z_\alpha,z_\beta) \;=\;& (2\pi)^2\delta_{\rm D}(\mathbf{l}+\mathbf{l}'')\Big\langle\, 2q\left\langle \tilde{\gamma}^{(2)}_{\rm B}(\mathbf{l},z_\alpha)\left(\tilde{\gamma}^{(1)}*\tilde{\kappa}^{(1)}\right)_{\rm B}(\mathbf{l}'',z_\beta)\right\rangle \\ &+(2q+q^2)\left\langle\left(\tilde{\gamma}^{(1)}*\tilde{\kappa}^{(1)}\right)_{\rm B}(\mathbf{l},z_\alpha)\left(\tilde{\gamma}^{(1)}*\tilde{\kappa}^{(1)}\right)_{\rm B}(\mathbf{l}'',z_\beta)\right\rangle\Big\rangle_{\alpha\beta} \\ &-q^2\int\frac{\mathrm{d}^2l'}{(2\pi)^2}\sin^2(2\phi_{\mathbf{l}'})C^{(11)}_{\tilde{\gamma}_{\rm E}}(l';z_\alpha,z_\beta)C^{(11)}_{\tilde{\gamma}_{\rm E}}(|\mathbf{l}'-\mathbf{l}|;z_\alpha,z_\beta)\,. \end{aligned} \tag{2.61}$$

In equation (2.60) we have simplified the terms which involve the variance of shear or convergence, e.g., the term in equation (2.57) which is proportional to C_1 becomes

$$\begin{aligned} &C_1\left\langle\tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l},z_\alpha)\left(\tilde{\gamma}^{(1)}*\tilde{\kappa}^{(1)}*\tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}',z_\beta)\right\rangle - C_1\left\langle\tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l},z_\alpha)\tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l}',z_\beta)\right\rangle\left\langle\kappa^2(z_\beta)\right\rangle \\ &\qquad = C_1\left\langle\tilde{\gamma}^{(1)}_{\rm E}(\mathbf{l},z_\alpha)\left(\tilde{\gamma}^{(1)}*\tilde{\kappa}^{(1)}*\tilde{\kappa}^{(1)}\right)_{\rm E}(\mathbf{l}',z_\beta)\right\rangle_{\rm c}\,. \end{aligned} \tag{2.62}$$

Here the second term is canceled by the disconnected part of the first term arising from the Wick contraction $C_1\langle\tilde{\gamma}\tilde{\gamma}\rangle\langle\tilde{\kappa}\tilde{\kappa}\rangle$, the two other Wick contractions of this term vanish after azimuthal integration. An explicit expression for the connected term is given in table 2.1.

For the redshift range and cosmology considered in this work, the second term and third in equation (2.61) are the dominant contributions. These terms partly cancel and on scales $l \gtrsim 50$ lensing bias effectively increases the B-mode power spectrum by approximately a factor $(1+2q)$, which is smaller than the findings of Schmidt et al. (2009a) who only considered the Gaussian contribution to the second term in equation (2.61). The B-mode signal is largest for small angular scaled and high source redshifts. Assuming $q \leq 2$ and a WMAP5 cosmology (Komatsu et al. 2009), for sources at $z \leq 3$ and in the range $l \leq 10000$ the B-mode power spectrum is suppressed by at least a factor 500 (a factor 3000 for $z \leq 1$) compared to the shear E-mode power spectrum.

Lensing bias gives rise to a third-order correction discussed by Schmidt et al. (2009a), which is q times the reduced shear correction analyzed by Shapiro (2009). The fourth-order E-mode correction generated by lensing bias equation (2.60) is more complicated and we will discuss its impact on the E-mode power spectrum in section 2.4.

The lensing bias E-mode and B-mode corrections are illustrated in figure 2.2 assuming a source redshift $z_\alpha = z_\beta = 1$. Due to uncertainties in modeling the nonlinear clustering of matter on small scales we restrict our analysis to $l \leq 3000$, on these scales the lensing bias corrections are below 1%.

Table 2.3 Z values for the $O(\Phi^4)$ corrections for different ellipticity estimators with lensing bias.

estimator	$q = 0$	$q = 1$	$q = 2$
$\boldsymbol{\varepsilon}$	1.14	3.19	5.31
$\mathbf{e}$, $\langle e^{(s)2}\rangle^{1/2} = 0.6$	0.12	2.13	4.25

2.4 Impact on Future Surveys

The corrections derived in section 2.3 generate a small amount of B-mode power, and have a $\lesssim$1% effect on the ellipticity E-mode power spectrum. These are well below the error bars of current surveys and therefore have no significant effect on published results. However, future "Stage IV" surveys such as LSST, JDEM, and Euclid will be sensitive to subpercent effects. We can quantify the importance of the higher-order lensing corrections by comparing the corrections to the power spectrum $\Delta C(l; z_\alpha, z_\beta)$ to their covariance matrix. Quantitatively,

$$Z = \sqrt{\sum_{l\alpha\beta l'\alpha'\beta'} \{\mathrm{Cov}^{-1}[C(l; z_\alpha, z_\beta), C(l'; z_{\alpha'}, z_{\beta'})]\}\Delta C(l; z_\alpha, z_\beta)\Delta C(l'; z_{\alpha'}, z_{\beta'})} \tag{2.63}$$

represents the number of sigmas at which the corrected and uncorrected power spectra could be distinguished by that survey. Corrections with $Z \ll 1$ are negligible in comparison with statistical errors, whereas corrections with $Z \gg 1$ must be known to high accuracy to make full use of the data set. We have computed equation (2.63) assuming a WMAP5 cosmology (Komatsu et al. 2009) for a model survey with a surface density of 30 galaxies/arcmin^2, median redshift $z_{\mathrm{med}} = 1.1$, and sky coverage of 10^4 deg^2, as appropriate for some of the proposed versions of JDEM. The power spectra were computed in 14 redshift slices and 12 l-bins with a maximum multipole of $l_{\mathrm{max}} = 3000$. The algorithm for computing the covariance matrix is as described in Appendix A.2.d of the JDEM Figure of Merit Science Working Group report (Albrecht et al. 2009). Without lensing bias ($q = 0$), we find $Z = 1.14$ for the linear ellipticity estimator $\boldsymbol{\varepsilon}$; for the standard estimator $\mathbf{e}$ and for an rms ellipticity[4] $\langle e^{(s)2}\rangle^{1/2} = 0.6$, we find $Z = 0.12$. Including the lensing bias corrections from section 2.3.4 increases the significance of the corrections as detailed in table 2.3. Note that the table includes only the $O(\Phi^4)$ corrections, and does *not* include the $O(\Phi^3)$ corrections that have previously been considered (Shapiro 2009, Schmidt et al. 2008). Thus, the perturbative corrections to the weak lensing approximation are expected to be at the level of $\sim 1\sigma$-4σ. These corrections will have to be

---

[4] The rms ellipticity here includes both the + and × components, so it is $\sqrt{2}$ times the rms per axis.

taken into account for future surveys, but given that they are only $\sim 1\sigma$-5σ and should be accurately calculable (either directly via ray-tracing simulations, or by analytic expression in terms of the moments of the density field, which can be determined from N-body simulations), they should not represent a fundamental difficulty.

2.5 Discussion

We have calculated the reduced shear power spectra perturbatively to fourth-order in the gravitational potential, accounting for the differences between shear and reduced shear, relaxing the Born approximation, and including lens–lens coupling in the calculation of shear and convergence. The full set of corrections to the reduced shear power spectra are given in table 2.1 (E-mode) and table 2.2 (B-mode). The ellipticity power spectrum contains additional contributions, equation (2.46), which arises from the nonlinearity of the shear estimator and depends on the specific definition of ellipticity used, and equation (2.60) which is caused by lensing bias. Through order Φ^4, this is the full set of corrections to the power spectrum arising from the lensing process itself. All corrections have been derived within the Limber approximation, and the analysis of "12" type multiple-deflection corrections is left for future work. Other corrections associated with the source galaxy population, such as source clustering and intrinsic alignments, are not treated in this analysis. We find that, depending on the properties of the source galaxy population and on the type of shear estimator used, these corrections will be at the $\sim 1\sigma$-4σ level, and thus should be included in the analysis of future precision cosmology weak lensing experiments.

That said, we caution that there are other areas in which the theory of weak lensing needs work if it is to meet ambitious future goals. Current fitting formula of the nonlinear dark matter power spectrum have an accuracy of about 10% at arcminute scales (Smith et al. 2003b) and the uncertainty exceeds 30% for $l > 10000$ (Hilbert et al. 2009), due to this difficulty in modeling the nonlinear gravitational clustering angular scales of $l > 3000$ are likely to be excluded from parameter fits to cosmic shear measurements. Utilizing near-future N-body simulations it will become possible to determine the nonlinear dark matter power spectrum with percent level accuracy (e.g., Heitmann et al. 2008, 2009). However, this does not account for the effect of baryons, which will likely be important at halo scales and depend critically on the details of baryonic processes (cooling, feedback) involved. Baryons in dark matter halos which are able to cool modify the structure of the dark matter halo through adiabatic contraction (Blumenthal et al. 1986, Gnedin et al. 2004), causing deviations of the inner halo profile from the simple NFW form and changing the halo mass–halo concentration relation (e.g., Rudd, Zentner and Kravtsov 2008, Pedrosa, Tissera and Scannapieco 2009). The latter can be constrained though galaxy-galaxy lensing (Mandelbaum et al. 2006), or could be internally self-calibrated in a weak lensing survey via its preferential effect on the small-scale power spectrum (Zentner, Rudd and Hu 2008). Baryons in the intergalactic medium may make up about 10% of the mass in the universe, and if their distribution on Megaparcec scales has been strongly affected by non-gravitational processes then they could pose a problem for precise calculation of the matter power spectrum (see Levine and Gnedin 2006, for an extreme and probably unrealistic example).

Given these uncertainties in modeling the nonlinear matter distribution and that all the corrections derived in this work are integrals over the nonlinear matter power spectrum, bispectrum and trispectrum, we refrain from calculating $O(\Phi^5)$ and higher corrections. We expect that the corrections derived here are sufficient to model the perturbative relation between the nonlinear matter distribution and the lensing distortion in weak lensing surveys for the foreseeable future.

Acknowledgements

We thank Wayne Hu, Fabian Schmidt, Peter Schneider, and Chaz Shapiro for useful discussions.

Chapter 3

A new Third-Order Cosmic Shear Statistics: Separating E/B-Mode Correlations on a Finite Interval*

3.1 Introduction

Cosmic shear, the distortion of light from distant galaxies by the tidal gravitational field of the intervening large-scale structure, is an excellent tool to probe the matter distribution in the universe. The statistics of the image distortions are related to the statistical properties of the large-scale matter distribution and the geometry of the universe, and can thereby be used to constrain cosmology. Current results already demonstrate the power of cosmic shear observations at constraining the clustering amplitude σ_8 and the matter density $\Omega_{\rm m}$ (e.g., Fu et al. 2008, Schrabback et al. 2007, Huff et al. 2011). Furthermore, cosmic shear provides an ideal tool to study dark energy through measuring the growth of structure with large future surveys like KIDS[1], DES[2], LSST[3] (Ivezic et al. 2008), or Euclid[4] (Laureijs et al. 2011). The large volume probed by these surveys will enable us to measure not only the power spectrum, but also higher-order statistics with unprecedented precision. As the evolved density field is non-Gaussian, the three-point correlation function and its Fourier space equivalent, the bispectrum, contain significant cosmological information complementary to the more commonly used two-point statistics and are a powerful tool for breaking parameter degeneracies (Takada and Jain 2004).

---

*This chapter was adapted from A new third-order cosmic shear statistics: Separating E/B-mode correlations on a finite interval, Elisabeth Krause, Peter Schneider and Tim Eifler; *MNRAS*, accepted (2012). Reproduced here with permission, copyright (2012) by Wiley-Blackwell.

[1] http://www.astro-wise.org/projects/KIDS

[2] http://www.darkenergysurvey.org/

[3] http://www.lsst.org/lsst/

[4] http://sci.esa.int/euclid

The upcoming weak lensing experiments will limit the statistical uncertainties to the percent level. In order to extract cosmological information from these cosmic shear experiments, the increased data quality needs to be accompanied by a thorough treatment of a wide range of systematic errors, from photometric redshifts and galaxy shape measurements to the removal of astrophysical contaminants.

If the shear estimated from observed galaxy shapes is solely caused by gravitational lensing, then it should consist only of a “gradient component,” the so-called E-mode shear. B-modes (or curl components) cannot be generated by gravitational light deflection to leading order, and higher-order corrections are expected to be very small. Hence observing any B-mode pattern indicates remaining systematics in the shear analysis.

Decomposing the observed shear field directly into E/B-modes (e.g Bunn et al. 2003) is complicated by the complex mask geometry of weak lensing observations. At the two-point statistics level, an E/B-mode decomposition is commonly performed using the aperture mass dispersion (Schneider et al. 1998) and related measures (e.g., Crittenden et al. 2002), which can be calculated from the measured shear two-point correlation function (2PCF) and is thus not affected by the masking geometry. However, these methods assume that the 2PCF is known either from $\theta = 0$ to some finite angular value (aperture mass dispersion) or to arbitrarily large separations. However, in reality the 2PCF can only be measured on a finite interval $[\theta_{\rm min}, \theta_{\rm max}]$, where the lower boundary is caused by inability to measure the shape of image pairs with very small angular separation. As Kilbinger, Schneider and Eifler (2006) pointed out, lack of shear-correlation measurements on small scales leads to an underestimation of the aperture mass dispersion on small scales and causes an apparent mixing of E- and B-modes with this type of estimator. Schneider and Kilbinger (2007), Eifler, Schneider and Krause (2010); and Schneider, Eifler and Krause (2010) develop statistical measures for an exact E/B-mode decomposition based on 2PCFs known only on a finite interval $[\theta_{\rm min}, \theta_{\rm max}]$.

At the three-point statistics level, Jarvis, Bernstein and Jain (2004) and Schneider, Kilbinger and Lombardi (2005) introduced E/B-mode separating shear measures which assume knowledge of the 3PCF down to arbitrarily small scales. Shi, Schneider and Joachimi (2011) derived a general condition for the E/B-mode decomposition of lensing three-point statistics, but the construction of filter functions with finite support based on this condition is far from straight forward. In this chapter we derive an extension of the 2PCF ring statistics (Schneider and Kilbinger 2007, Eifler, Schneider and Krause 2010) to an exact E/B-mode decomposition of shear three-point correlation functions on a finite interval.

In order to constrain cosmology with third-order shear statistics, it is important to obtain the corresponding predictions from a large suite of cosmological numerical simulations in a reasonable time and with limited computational effort. We facilitate this by giving an expression of the third-order ring statistics in terms of the convergence field, thereby avoiding the time-consuming calculation of the shear 3PCF for each simulation.

3.2 Shear Three-Point Correlation Functions

We first introduce the shear three-point correlation function (3PCF): Consider a triangle in the complex plane with vertices $\mathbf{X}_i$ and let $\gamma_\mu(\mathbf{X}_i)$, $\mu = 1, 2$ be the Cartesian components of the shear at point $\mathbf{X}_i$. Unless otherwise noted, we will assume that the triangle is oriented such that $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ are ordered counterclockwise around the triangle. We define $\mathbf{x}_1 = \mathbf{X}_1 - \mathbf{X}_3$ and $\mathbf{x}_2 = \mathbf{X}_2 - \mathbf{X}_3$ to be the sides of this triangle (c.f. figure 3.1). We will use $\mathbf{x}_i$ to refer to complex numbers or vectors interchangeably, and denote their magnitude as x_i.

The Cartesian components of the shear 3PCF are defined as

$$\gamma_{\mu\nu\lambda}(\mathbf{x}_1, \mathbf{x}_2) \equiv \left\langle \gamma_\mu(\mathbf{X}_1)\gamma_\nu(\mathbf{X}_2)\gamma_\lambda(\mathbf{X}_3) \right\rangle , \tag{3.1}$$

where we have assumed that the shear field is statistically homogeneous so that $\gamma_{\mu\nu\lambda}$ depends only on the side vectors $\mathbf{x}_i$. Since one cannot form a trilinear scalar from the product of three shears, the behavior of the Cartesian components of the shear 3PCF under rotations is complicated. In order to write the 3PCF in terms of tangential (γ_t) and cross components ($\gamma_\times$) of the shear which are parity eigenstates and have relatively simple transformation properties, one can project the complex Cartesian shear $\gamma^\mathrm{c} = \gamma_1 + \mathrm{i}\gamma_2$ into tangential and cross component with respect to a chosen direction $\mathbf{a}_i$ with polar angle α_i,

$$\gamma(\mathbf{X}_i; \alpha_i) \equiv \gamma_\mathrm{t}(\mathbf{X}_i; \alpha_i) + \mathrm{i}\gamma_\times(\mathbf{X}_i; \alpha_i) = -\left[\gamma_1(\mathbf{X}_i) + \mathrm{i}\gamma_2(\mathbf{X}_i)\right] \mathrm{e}^{-2\mathrm{i}\alpha_i} = -\gamma^\mathrm{c}(\mathbf{X}_i)\mathrm{e}^{-2\mathrm{i}\alpha_i} = -\gamma^\mathrm{c}(\mathbf{X}_i)\, \mathbf{a}_i^{*2}/a_i^2 \,. \tag{3.2}$$

If the directions of projection α_i are defined in terms of the vertices $\mathbf{X}_i$ and thus do not depend on an external coordinate system, then the tangential and cross shear are invariant under rotations of the triangle (Schneider and Lombardi 2003, Takada and Jain 2003, Zaldarriaga and Scoccimarro 2003), and the 3PCF of these shear projections will only depend on the side *lengths* x_i and the orientation of the triangle (clockwise or counterclockwise). In the following we will use the centroid projection, where the shear at vertex $\mathbf{X}_i$ is projected along the direction $\mathbf{q}_i$ connecting $\mathbf{X}_i$ with the centroid $\bar{\mathbf{X}} = (\mathbf{X}_1 + \mathbf{X}_2 + \mathbf{X}_3)/3$, and α_i is the polar angle of this projection direction (see figure 3.1 for an illustration).

Following Schneider and Lombardi (2003) we define the (complex) natural components of the 3PCF, which have relatively simple transformation properties

$$\begin{aligned}
\Gamma^{(0)}(x_1, x_2, x_3) &\equiv \langle \gamma\,(\mathbf{X}_1; \alpha_1)\, \gamma\,(\mathbf{X}_2; \alpha_2)\, \gamma\,(\mathbf{X}_3; \alpha_3)\rangle = -\,\langle \gamma^{\rm c}\,(\mathbf{X}_1)\, \gamma^{\rm c}\,(\mathbf{X}_2)\, \gamma^{\rm c}\,(\mathbf{X}_3)\rangle\, {\rm e}^{-2{\rm i}(\alpha_1+\alpha_2+\alpha_3)}\,, \\
\Gamma^{(1)}(x_1, x_2, x_3) &\equiv \langle \gamma^*\,(\mathbf{X}_1; \alpha_1)\, \gamma\,(\mathbf{X}_2; \alpha_2)\, \gamma\,(\mathbf{X}_3; \alpha_3)\rangle = -\,\langle \gamma^{\rm c*}\,(\mathbf{X}_1)\, \gamma^{\rm c}\,(\mathbf{X}_2)\, \gamma^{\rm c}\,(\mathbf{X}_3)\rangle\, {\rm e}^{-2{\rm i}(-\alpha_1+\alpha_2+\alpha_3)}\,, \\
\Gamma^{(2)}(x_1, x_2, x_3) &\equiv \langle \gamma\,(\mathbf{X}_1; \alpha_1)\, \gamma^*\,(\mathbf{X}_2; \alpha_2)\, \gamma\,(\mathbf{X}_3; \alpha_3)\rangle\,, \\
\Gamma^{(3)}(x_1, x_2, x_3) &\equiv \langle \gamma\,(\mathbf{X}_1; \alpha_1)\, \gamma\,(\mathbf{X}_2; \alpha_2)\, \gamma^*\,(\mathbf{X}_3; \alpha_3)\rangle\ . \qquad (3.3)
\end{aligned}$$

$\Gamma^{(0)}$ is invariant under cyclic permutations of arguments; the other three components transform into each other: $\Gamma^{(1)}(x_1, x_2, x_3) = \Gamma^{(2)}(x_3, x_1, x_2) = \Gamma^{(3)}(x_2, x_3, x_1)$, etc. A different parameterization of oriented triangles is in terms of two sides and their inner angle, e.g., x_1, x_2, and ϕ (c.f. figure 3.1). We choose the convention $\phi \in [-\pi, \pi]$, such that $\phi > 0$ corresponds to $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ being ordered counter-clock-wise (“positive orientation”) and $\phi < 0$ corresponds to clock-wise ordering (”negative orientation”).

3.3 E/B-Mode Separation

To construct integrals which separate third-order E- and B-mode correlations we start from the circle statistics $C(\theta)$ (Crittenden et al. 2002, Schneider, van Waerbeke and Mellier 2002), which geometrically separates E- and B-modes by measuring the mean tangential and cross component of the shear on a circle of radius θ around the origin,

$$C(\theta) = C_{\rm t}(\theta) + {\rm i}C_\times(\theta) = \frac{1}{2\pi}\int_0^{2\pi} {\rm d}\psi\ \left(\gamma_{\rm t} + {\rm i}\gamma_\times\right)(\theta, \psi; \psi) = -\frac{1}{2\pi}\int_0^{2\pi} {\rm d}\psi\ {\rm e}^{-2{\rm i}\psi}\gamma^{\rm c}\,(\theta, \psi)\ , \qquad (3.4)$$

where ψ is the polar angle on the circle, and in the last step we have rotated the tangential and radial shear into the cartesian components. Following Schneider and Kilbinger (2007) we now consider the shear inside an annulus $\vartheta_1 \leq \theta \leq \vartheta_2$ and define the ring statistics $\mathcal{R}$,

$$\mathcal{R} = \mathcal{R}_{\rm t} + {\rm i}\mathcal{R}_\times = \int_{\vartheta_1}^{\vartheta_2} {\rm d}\theta\, W(\theta; \vartheta_1, \vartheta_2)\, C(\theta)\,, \tag{3.5}$$

which is a function of two radii ϑ_1 and ϑ_2, and where $W(\theta; \vartheta_1, \vartheta_2)$ is a normalized weight function

$$\int_{\vartheta_1}^{\vartheta_2} {\rm d}\theta\; W(\theta; \vartheta_1, \vartheta_2) = 1\,, \tag{3.6}$$

and $W = 0$ outside the annulus, i.e., if $\theta < \vartheta_1$ or $\theta > \vartheta_2$. From this definition we construct the third-order ring statistics as the correlation of the weighted mean shear in three concentric annuli with radii $\vartheta_1 \leq \theta_1, \leq \vartheta_2 < \vartheta_3 \leq \theta_2 \leq \vartheta_4 < \vartheta_5 \leq \theta_3 \leq \vartheta_6$ (cf. figure 3.1),

$$\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle (\boldsymbol{\vartheta}) = \int_{\vartheta_1}^{\vartheta_2} {\rm d}\theta_1 W(\theta_1; \vartheta_1, \vartheta_2) \int_{\vartheta_3}^{\vartheta_4} {\rm d}\theta_2 W(\theta_2; \vartheta_3, \vartheta_4) \int_{\vartheta_5}^{\vartheta_6} {\rm d}\theta_3 W(\theta_3; \vartheta_5, \vartheta_6)\, \langle C(\theta_1)\, C(\theta_2)\, C(\theta_3)\rangle \tag{3.7}$$

$$\langle \mathcal{R}^*\mathcal{R}\mathcal{R}\rangle (\boldsymbol{\vartheta}) = \int_{\vartheta_1}^{\vartheta_2} {\rm d}\theta_1 W(\theta_1; \vartheta_1, \vartheta_2) \int_{\vartheta_3}^{\vartheta_4} {\rm d}\theta_2 W(\theta_2; \vartheta_3, \vartheta_4) \int_{\vartheta_5}^{\vartheta_6} {\rm d}\theta_3 W(\theta_3; \vartheta_5, \vartheta_6)\, \langle C^*(\theta_1)\, C(\theta_2)\, C(\theta_3)\rangle\,, \tag{3.8}$$

where we have used $\boldsymbol{\vartheta} = (\vartheta_1, ..., \vartheta_6)$ to denote a six-tuple of radii. Expanding these correlators in terms of the mean tangential and cross shear yields

$$\begin{aligned} \langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle (\boldsymbol{\vartheta}) &= [\langle \mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\rangle - \langle \mathcal{R}_\times\mathcal{R}_\times\mathcal{R}_{\rm t}\rangle - \langle \mathcal{R}_\times\mathcal{R}_{\rm t}\mathcal{R}_\times\rangle - \langle \mathcal{R}_{\rm t}\mathcal{R}_\times\mathcal{R}_\times\rangle]\,(\boldsymbol{\vartheta}) \\ &\quad + {\rm i}\,[-\langle \mathcal{R}_\times\mathcal{R}_\times\mathcal{R}_\times\rangle + \langle \mathcal{R}_\times\mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\rangle + \langle \mathcal{R}_{\rm t}\mathcal{R}_\times\mathcal{R}_{\rm t}\rangle + \langle \mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\mathcal{R}_\times\rangle]\,(\boldsymbol{\vartheta}) \end{aligned} \tag{3.9}$$

$$\begin{aligned} \langle \mathcal{R}^*\mathcal{R}\mathcal{R}\rangle (\boldsymbol{\vartheta}) &= [\langle \mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\rangle + \langle \mathcal{R}_\times\mathcal{R}_\times\mathcal{R}_{\rm t}\rangle + \langle \mathcal{R}_\times\mathcal{R}_{\rm t}\mathcal{R}_\times\rangle - \langle \mathcal{R}_{\rm t}\mathcal{R}_\times\mathcal{R}_\times\rangle]\,(\boldsymbol{\vartheta}) \\ &\quad + {\rm i}\,[\langle \mathcal{R}_\times\mathcal{R}_\times\mathcal{R}_\times\rangle - \langle \mathcal{R}_\times\mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\rangle + \langle \mathcal{R}_{\rm t}\mathcal{R}_\times\mathcal{R}_{\rm t}\rangle + \langle \mathcal{R}_{\rm t}\mathcal{R}_{\rm t}\mathcal{R}_\times\rangle]\,(\boldsymbol{\vartheta})\,. \end{aligned} \tag{3.10}$$

Note that the imaginary parts of (3.9, 3.10) vanish in the absence of parity-violating modes.

We analogously define the correlators $\langle \mathcal{R}\mathcal{R}^*\mathcal{R}\rangle$ and $\langle \mathcal{R}\mathcal{R}\mathcal{R}^*\rangle$ and separate E- and B- modes via

$$\left\langle \mathcal{R}_{\rm E}^3 \right\rangle (\boldsymbol{\vartheta}) = \frac{1}{4}{\rm Re}\,[\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}^*\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}^*\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}\mathcal{R}^*\rangle]\,(\boldsymbol{\vartheta}) \tag{3.11}$$

$$\left\langle \mathcal{R}_{\rm E}\mathcal{R}_{\rm B}^2 \right\rangle (\boldsymbol{\vartheta}) = \frac{1}{4}{\rm Re}\,[-3\,\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}^*\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}^*\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}\mathcal{R}^*\rangle]\,(\boldsymbol{\vartheta}) \tag{3.12}$$

$$\left\langle \mathcal{R}_{\rm E}^2\mathcal{R}_{\rm B} \right\rangle (\boldsymbol{\vartheta}) = \frac{1}{4}{\rm Im}\,[3\,\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}^*\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}^*\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}\mathcal{R}^*\rangle]\,(\boldsymbol{\vartheta}) \tag{3.13}$$

$$\left\langle \mathcal{R}_{\rm B}^3 \right\rangle (\boldsymbol{\vartheta}) = \frac{1}{4}{\rm Im}\,[-\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}^*\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}^*\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}\mathcal{R}^*\rangle]\,(\boldsymbol{\vartheta})\,, \tag{3.14}$$

where equation (3.11) corresponds to pure E-mode correlations, and equation (3.14) to parity violating third-order B-mode correlations. Equation (3.12) is a parity invariant correlation between E- and B-modes, and equation (3.13) is a parity violating correlation between E- and B-modes.

For brevity, the mixed terms (3.12, 3.13) are generalized expressions which are sensitive to B-modes in any of the annuli, i.e $\left\langle \mathcal{R}_\mathrm{E}\mathcal{R}_\mathrm{B}^2 \right\rangle(\boldsymbol{\vartheta}) = (\langle R_\mathrm{t}R_\times R_\times\rangle + \langle R_\times R_\mathrm{t} R_\times\rangle + \langle R_\times R_\times R_\mathrm{t}\rangle)/3$, etc.. Instead one can also consider more localized B-mode measures like

$$\langle R_\times R_\times R_\mathrm{t}\rangle(\boldsymbol{\vartheta}) \quad = \quad \frac{1}{4}\mathrm{Re}\left[\langle -\mathcal{R}\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}^*\mathcal{R}\mathcal{R}\rangle + \langle \mathcal{R}\mathcal{R}^*\mathcal{R}\rangle - \langle \mathcal{R}\mathcal{R}\mathcal{R}^*\rangle\right](\boldsymbol{\vartheta})\,, \tag{3.15}$$

which picks up correlations with B modes in the innermost and middle annulus, but is insensitive to B-modes in the outer annulus. Such a localized test for B-mode correlations can help identifying the source of a B-mode contamination if equation (3.12) indicates the overall presence of B-modes.

3.4 Third-Order Ring Statistics

In this section we derive computationally advantageous expressions for the third-order ring statistics in terms of the shear 3PCF, and show their relation to the convergence bispectrum.

3.4.1 Relation to the Shear Three-Point Functions

We rewrite the third-order ring statistics in terms of the shear 3PCF by starting from the definition equation (3.7)

$$\begin{aligned}\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle(\boldsymbol{\vartheta}) \quad = \quad & -\int_{\vartheta_1}^{\vartheta_2} \mathrm{d}\theta_1\, W(\theta_1;\vartheta_1,\vartheta_2) \int_{\vartheta_3}^{\vartheta_4} \mathrm{d}\theta_2\, W(\theta_2;\vartheta_3,\vartheta_4) \int_{\vartheta_5}^{\vartheta_6} \mathrm{d}\theta_3\, W(\theta_3;\vartheta_5,\vartheta_6) \\ & \times \int_0^{2\pi} \frac{\mathrm{d}\psi_3}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\psi_2}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\psi_1}{2\pi}\, \mathrm{e}^{-2\mathrm{i}(\psi_1+\psi_2+\psi_3)} \left\langle \gamma^\mathrm{c}(\theta_1,\psi_1)\,\gamma^\mathrm{c}(\theta_2,\psi_2)\,\gamma^\mathrm{c}(\theta_3,\psi_3)\right\rangle\,.\end{aligned} \tag{3.16}$$

Noting that $\mathbf{X}_j = \theta_j \exp(\mathrm{i}\psi_j)$ and using equation (3.3), this can be rewritten as

$$\begin{aligned}\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle(\boldsymbol{\vartheta}) \quad = \quad & \int \frac{\mathrm{d}^2 X_1}{2\pi\,|\mathbf{X}_1|}\, W(|\mathbf{X}_1|;\vartheta_1,\vartheta_2) \int \frac{\mathrm{d}^2 X_2}{2\pi\,|\mathbf{X}_2|}\, W(|\mathbf{X}_2|;\vartheta_3,\vartheta_4) \int \frac{\mathrm{d}^2 X_3}{2\pi\,|\mathbf{X}_3|}\, W(|\mathbf{X}_3|;\vartheta_5,\vartheta_6) \\ & \times \exp(2\mathrm{i}(\alpha_1+\alpha_2+\alpha_3-\psi_1-\psi_2-\psi_3))\,\Gamma^{(0)}(\mathbf{X}_1-\mathbf{X}_3,\mathbf{X}_2-\mathbf{X}_3)\,,\end{aligned} \tag{3.17}$$

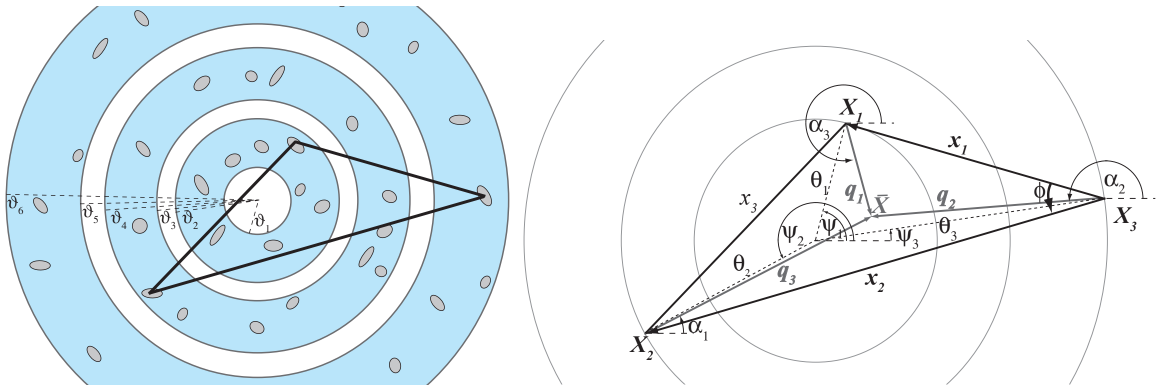


Figure 3.1 *Left:* Concept of the third-order ring statistics $\langle \mathcal{RRR} \rangle (\vartheta_1, \vartheta_2, ..., \vartheta_6)$. E/B-modes are separated on a finite interval by correlating the shear of galaxy triplets located within three concentric annuli, as illustrated by the thick black triangle. The minimum separation between galaxies in the above geometry is $\min(\vartheta_3 - \vartheta_2, \vartheta_5 - \vartheta_4)$ and the maximum separation is $\vartheta_6 + \vartheta_4$.
Right: Geometry of a triangle in the third-order ring statistics. The triangle vertices $\mathbf{X}_j$ are located on three concentric circles of radius θ_j and have polar angles ψ_j. The triangle centroid is $\bar{\mathbf{X}}$. In the centroid projection the shear at each $\mathbf{X}_j$ is projected onto the centroid along direction $\mathbf{q}_j$, the line connecting $\mathbf{X}_j$ with the centroid. φ_j is the orientation angle of vector $\mathbf{q}_j$. Finally, ϕ is the inner angle of the triangle at $\mathbf{X}_3$ which we will use when parameterizing a triangle in terms of two side lengths x_1, x_2 and angle ϕ.

where $\Gamma^{(0)}$ is the shear 3PCF measured relative to the centroid, so that the α_i are the directions of the point $\mathbf{X}_i$ to the centroid $\bar{\mathbf{X}} = (\mathbf{X}_1 + \mathbf{X}_2 + \mathbf{X}_3)/3$. Owing to circular symmetry, we can set $\psi_3 = 0$; equivalently, one can use relative polar angles $\Delta\psi_j = \psi_j - \psi_3$ and show that the integrand depends only on these relative angles.

As Γ is measured within discrete angular bins, while the weight functions and geometric factors in equation (3.17) can be evaluated continuously, it is numerically more stable to rewrite the third-order ring statistics such that only the three outermost integrals contain the shear 3PCF and the inner integrals can be evaluated numerically to arbitrary precision. With $\mathbf{x}_j = \mathbf{X}_j - \mathbf{X}_3 = \theta_j \mathrm{e}^{\mathrm{i}\psi_j} - \theta_3$ for $j = 1, 2$,

$$\begin{aligned} \langle \mathcal{RRR} \rangle (\boldsymbol{\vartheta}) &= \frac{1}{(2\pi)^2} \int \mathrm{d}^2 x_1 \int \mathrm{d}^2 x_2 \, \Gamma^{(0)}(\mathbf{x}_1, \mathbf{x}_2) \int \mathrm{d}\theta_3 \; W(\theta_3; \vartheta_5, \vartheta_6) \, W(|\mathbf{x}_1 + \boldsymbol{\theta}_3|; \vartheta_1, \vartheta_2) \\ & \quad \times W(|\mathbf{x}_2 + \boldsymbol{\theta}_3|; \vartheta_3, \vartheta_4) \frac{1}{|\mathbf{x}_1 + \boldsymbol{\theta}_3||\mathbf{x}_2 + \boldsymbol{\theta}_3|} \; \exp(2\mathrm{i}(\alpha_1 + \alpha_2 + \alpha_3 - \psi_1 - \psi_2) \,, \end{aligned} \tag{3.18}$$

where we have used $\boldsymbol{\theta}_3$ to denote a complex number with zero imaginary part for consistency. We have $\mathrm{e}^{2\mathrm{i}\alpha_j} = \mathbf{q}_j/\mathbf{q}_j^*$, with $\mathbf{q}_1 = (2\mathbf{x}_1 - \mathbf{x}_2)/3$, $\mathbf{q}_2 = (2\mathbf{x}_2 - \mathbf{x}_1)/3$, $\mathbf{q}_3 = -(\mathbf{x}_1 + \mathbf{x}_2)/3$, and

$\mathrm{e}^{\mathrm{i}\psi_j} = (\mathbf{x}_j + \boldsymbol{\theta}_3)/\theta_j$, so that

$$\mathrm{e}^{-2\mathrm{i}\psi_j} = \frac{\mathbf{x}_j^* + \boldsymbol{\theta}_3^*}{\mathbf{x}_j + \boldsymbol{\theta}_3} \, . \tag{3.19}$$

Thus,

$$\begin{aligned} \langle \mathcal{RRR} \rangle (\boldsymbol{\vartheta}) &= \frac{1}{(2\pi)^2} \int \mathrm{d}^2 x_1 \int \mathrm{d}^2 x_2 \, \Gamma^{(0)}(\mathbf{x}_1, \mathbf{x}_2) \int \mathrm{d}\theta_3 \; W(\theta_3; \vartheta_5, \vartheta_6) \, W(|x_1 + \boldsymbol{\theta}_3|; \vartheta_1, \vartheta_2) \\ &\quad \times W(|x_2 + \boldsymbol{\theta}_3|; \vartheta_3, \vartheta_4) \, \frac{1}{|\mathbf{x}_1 + \boldsymbol{\theta}_3| |\mathbf{x}_2 + \boldsymbol{\theta}_3|} \, \frac{\mathbf{q}_1 \mathbf{q}_2 \mathbf{q}_3}{\mathbf{q}_1^* \mathbf{q}_2^* \mathbf{q}_3^*} \, \frac{\mathbf{x}_1^* + \boldsymbol{\theta}_3^*}{\mathbf{x}_1 + \boldsymbol{\theta}_3} \, \frac{\mathbf{x}_2^* + \boldsymbol{\theta}_3^*}{\mathbf{x}_2 + \boldsymbol{\theta}_3} \, . \end{aligned} \tag{3.20}$$

Finally, if φ_i is the polar angle of $\mathbf{x}_i$, and $\phi = \varphi_2 - \varphi_1$ is the angle between $\mathbf{x}_2$ and $\mathbf{x}_1$, we obtain

$$\begin{aligned} \langle \mathcal{RRR} \rangle (\boldsymbol{\vartheta}) &= \frac{1}{(2\pi)^2} \int \mathrm{d}x_1 \; x_1 \int \mathrm{d}x_2 \; x_2 \int \mathrm{d}\phi \; \Gamma^{(0)}(x_1, x_2, \phi) \\ &\quad \times \int \mathrm{d}\theta_3 \, W(\theta_3; \vartheta_5, \vartheta_6) \int \mathrm{d}\varphi_1 \; W(|x_1 \mathrm{e}^{\mathrm{i}\varphi_1} + \theta_3|; \vartheta_1, \vartheta_2) \, W(|x_2 \mathrm{e}^{\mathrm{i}(\varphi_1 + \phi)} + \theta_3|; \vartheta_3, \vartheta_4) \\ &\quad \times \frac{1}{|\mathbf{x}_1 + \boldsymbol{\theta}_3| |\mathbf{x}_2 + \boldsymbol{\theta}_3|} \, \frac{\mathbf{q}_1 \mathbf{q}_2 \mathbf{q}_3}{\mathbf{q}_1^* \mathbf{q}_2^* \mathbf{q}_3^*} \, \frac{\mathbf{x}_1^* + \theta_3}{\mathbf{x}_1 + \theta_3} \, \frac{\mathbf{x}_2^* + \theta_3}{\mathbf{x}_2 + \theta_3} \\ &\equiv \frac{1}{(2\pi)^2} \int \mathrm{d}x_1 \; x_1 \int \mathrm{d}x_2 \; x_2 \int \mathrm{d}\phi \; \Gamma^{(0)}(x_1, x_2, \phi) \; \mathbf{Z}_0(x_1, x_2, \phi, \boldsymbol{\vartheta}) \, , \end{aligned} \tag{3.21}$$

where we have defined the complex filter function $\mathbf{Z}_0$ of the ring statistics in the last step. Note that the ratio of the $\mathbf{q}$'s does not depend on θ_3 and thus the evaluation of the filter function can be further simplified by reversing the order of integration and moving this phase factor to the outer (φ_1-) integral.

Expressions for the other correlations required for E/B-mode separation, which contain a complex conjugate ring statistic $\mathcal{R}^*$, are derived analogously. For the correlation involving the complex conjugate shear at vertex $\mathbf{X}_j$, the resulting expression analogous to equation (3.21) contains $\Gamma^{(j)}$ instead of $\Gamma^{(0)}$, $\mathbf{q}_j/\mathbf{q}_j^*$ is replaced by its complex conjugate (corresponding to $\alpha_j \to -\alpha_j$ in equation (3.3)), and for $j = 1, 2$ the factor $\mathbf{x}_j/\mathbf{x}_j^*$ is also replaced by its complex conjugate (corresponding to $\psi_j \to -\psi_j$ in the equivalent of equation (3.16)), e.g.,

$$\langle \mathcal{R}^* \mathcal{RR} \rangle (\boldsymbol{\vartheta}) = \frac{1}{(2\pi)^2} \int \mathrm{d}x_1 \; x_1 \int \mathrm{d}x_2 \; x_2 \int \mathrm{d}\phi \; \Gamma^{(1)}(x_1, x_2, \phi) \; \mathbf{Z}_1(x_1, x_2, \phi, \boldsymbol{\vartheta}) \, , \tag{3.22}$$

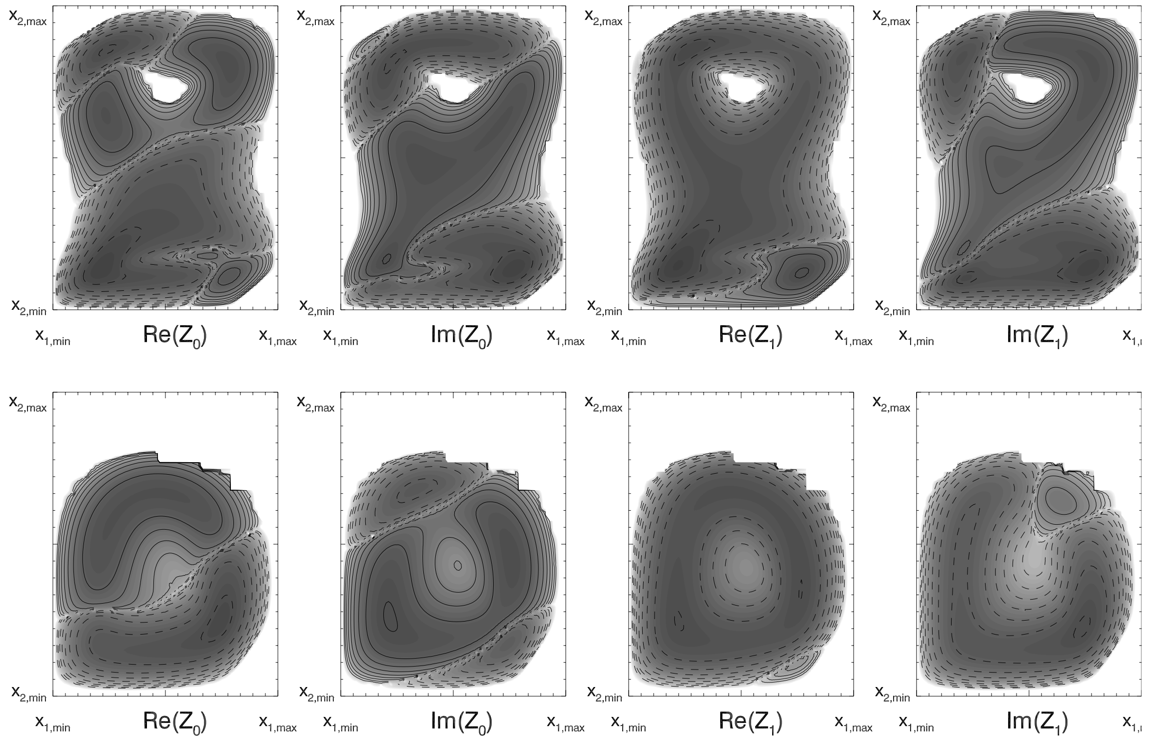


Figure 3.2 Contours of the filter functions $\mathbf{Z}_{0,1}$ as a function of angular scales $x_1 \in [\vartheta_5 - \vartheta_2, \vartheta_6 + \vartheta_2]$ and $x_2 \in [\vartheta_5 - \vartheta_4, \vartheta_6 + \vartheta_4]$ for $\boldsymbol{\vartheta} = (1', 2', 3', 4', 5', 6')$ with $\phi = \pi/8$ (top row) or $\phi = \pi/4$ (bottom row). Contour lines are evenly spaced with separation 0.5 dex ranging from $|Z| = 10^{-5}$ to $|Z| = 10^{-1.5}$, dashed lines indicate regions where Z is negative. The filter functions vanish if the triangle configuration (x_1, x_2, ϕ) is not allowed in the ring statistics geometry (c.f. Fig 3.1).

with

$$\begin{aligned}\mathbf{Z}_1(x_1, x_2, \phi, \boldsymbol{\vartheta}) \;\; = \;\; & \int \mathrm{d}\theta_3 \, W(\theta_3; \vartheta_5, \vartheta_6) \int \mathrm{d}\varphi_1 \, W(|x_1 \mathrm{e}^{\mathrm{i}\varphi_1} + \theta_3|; \vartheta_1, \vartheta_2) \, W(|x_2 \mathrm{e}^{\mathrm{i}(\varphi_1+\phi)} + \theta_3|; \vartheta_3, \vartheta_4) \\ & \times \frac{1}{|\mathbf{x}_1 + \boldsymbol{\theta}_3||\mathbf{x}_2 + \boldsymbol{\theta}_3|} \frac{\mathbf{q}_1^* \mathbf{q}_2 \mathbf{q}_3}{\mathbf{q}_1 \mathbf{q}_2^* \mathbf{q}_3^*} \frac{\mathbf{x}_1 + \theta_3}{\mathbf{x}_1^* + \theta_3} \frac{\mathbf{x}_2^* + \theta_3}{\mathbf{x}_2 + \theta_3} \, . \end{aligned} \tag{3.23}$$

In the computation of the ring statistics one can choose any (normalized) radial weight function W that fulfills $W(0; \vartheta_1, \vartheta_2) = 0$ even if $\vartheta_1 = 0$ (as the separation in tangential/cross shear is ill-defined on circle of radius $\theta = 0$). To be specific, we choose

$$W(\theta; \vartheta_i, \vartheta_j) = 30 \frac{(\theta - \vartheta_i)^2 (\vartheta_j - \theta)^2}{(\vartheta_j - \vartheta_i)^5} \, , \tag{3.24}$$

as in the computation of the second-order ring statistics (Schneider and Kilbinger 2007). The shape of the third-order ring statistics filter functions $Z_{0,1}$ based on this choice for W is illustrated in figure 3.2.

3.4.2 Relation to the Bispectrum

In order to rewrite the third-order ring statistics in terms of the bispectrum we first relate it to the lensing convergence field κ, which is easier to express in terms of the convergence bispectrum than the shear 3PCF (see Schneider, Kilbinger and Lombardi 2005, for details) as it contains fewer oscillatory phase factors. Expressing the ring statistics in terms of the convergence field also speeds up the measurement of $\left\langle \mathcal{R}_{\mathrm{E}}^3 \right\rangle$ in simulations considerably, as described below.

Consider the convergence field smoothed with a radially symmetric filter $U_\vartheta(\theta)$ with characteristic scale ϑ. If $U_\vartheta(\theta)$ is a compensated filter $\int \mathrm{d}\theta \; \theta \, U_\vartheta(\theta) = 0$, this convolution can be expressed in terms of the shear field as

$$\int \mathrm{d}^2\theta' \; U_\vartheta(|\boldsymbol{\theta}'|) \, \kappa(\boldsymbol{\theta}') = \int \mathrm{d}^2\theta \; Q_\vartheta(|\boldsymbol{\theta}'|) \, \gamma_{\mathrm{t}}(\boldsymbol{\theta}') \, , \tag{3.25}$$

where U and Q are related by (Kaiser, Squires and Broadhurst 1995, Schneider 1996, Schneider and Kilbinger 2007)

$$Q_\vartheta(\theta) = \frac{2}{\theta^2} \int_0^\theta \mathrm{d}\theta' U_\vartheta(\theta') - U_\vartheta(\theta) \quad \text{and} \quad U_\vartheta(\theta) = \int_\theta^\infty \frac{2\mathrm{d}\theta'}{\theta'} Q_\vartheta(\theta') - Q_\vartheta(\theta) \, . \tag{3.26}$$

As shown in Schneider and Kilbinger (2007) the definition of the ring statistics $\mathcal{R}$ (equation (3.5)) is equivalent to an aperture mass $M_{\rm ap}(\vartheta_i, \vartheta_j)$ with two characteristic scales if

$$Q_{\vartheta_i,\vartheta_j}(\theta) = \frac{W(\theta; \vartheta_i, \vartheta_j)}{2\pi\,\theta} \,. \tag{3.27}$$

As the relation between the filter Q and U does not depend on the shape of Q, we can calculate the corresponding compensated filter $U_{\vartheta_i,\vartheta_j}(\theta)$ as in equation (3.26). The left and middle panel of figure 3.3 show the ring statistics filter $W(\theta; \vartheta_i, \vartheta_j)$ and the corresponding aperture mass filter $U_{\vartheta_i,\vartheta_j}(\theta)$ for different choices of ring radii $(\vartheta_i, \vartheta_j)$. As expected from equation (3.26), U is constant for $\theta < \vartheta_i$, then becomes negative, and is zero for $\theta > \vartheta_j$.

Based on equations (3.25, 3.26, 3.27), the third-order ring statistics of a pure E-mode field can be computed directly from simulated convergence maps by convolving the convergence field with different filters $U_{\vartheta_i,\vartheta_j}$ and correlating three filtered maps. With this approach one does not need to calculate the shear 3PCF, which are computationally expensive (e.g., Jarvis, Bernstein and Jain 2004). We stress that calculating the ring statistics from the convergence field is only possible for obtaining predictions from simulations, which are B-mode free by construction and allow for the direct measurement of the convergence.

Expressing $\mathcal{R}_{\rm E}$ as the convolution of κ and $U_{\vartheta_i,\vartheta_j}$ also enables us to write down he third-order ring statistics of a pure E-mode field in terms of the convergence bispectrum $B_\kappa(l_1, l_2, l_3)$ (c.f. Schneider, Kilbinger and Lombardi 2005)

$$\begin{aligned} \left\langle \mathcal{R}_{\rm E}^3 \right\rangle (\boldsymbol{\vartheta}) \;&=\; \frac{1}{(2\pi)^3} \int {\rm d}l_1\, l_1 \int {\rm d}l_2\, l_2 \int {\rm d}\phi \; B_\kappa\left(l_1, l_2, \sqrt{l_1^2 + l_2^2 - 2 l_1 l_2 \cos\phi}\right) \\ &\quad \times \tilde{U}_{\vartheta_1,\vartheta_2}(l_1) \tilde{U}_{\vartheta_3,\vartheta_4}(l_2) \tilde{U}_{\vartheta_5,\vartheta_6}\left(\sqrt{l_1^2 + l_2^2 - 2 l_1 l_2 \cos\phi}\right) , \end{aligned} \tag{3.28}$$

with the Fourier transformed filter function $\tilde{U}(l) = \int {\rm d}\theta\, \theta\, {\rm J}_0(l\theta) U(\theta)$. The bispectrum filter functions for the third-order ring statistics are illustrated in the right panel of figure 3.3.

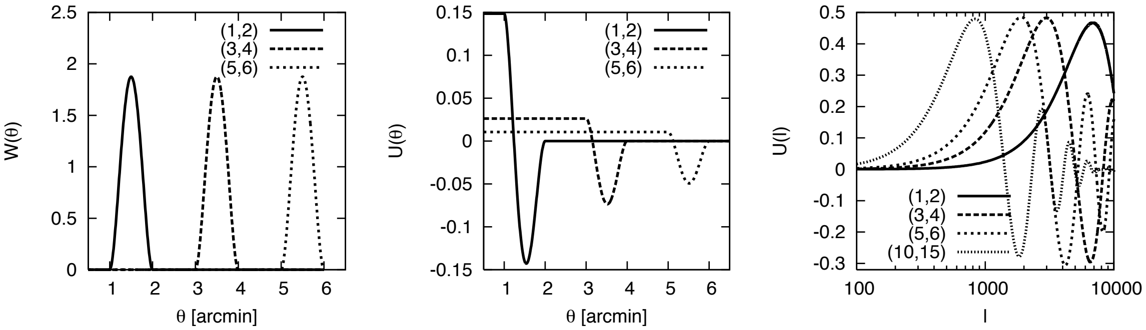


Figure 3.3 Filter functions associated with $\mathcal{R}$ for different ring radii $(\vartheta_i, \vartheta_j)$. The left panel shows the radial filter function of the ring statistics, $W(\theta; \vartheta_i, \vartheta_j)$; the corresponding compensated aperture mass filter function $U_{\vartheta_i,\vartheta_j}(\theta)$ is shown in the middle panel, and the right panel illustrates the corresponding Fourier transform $\tilde{U}_{\vartheta_i,\vartheta_j}(l)$.

3.5 Conclusion

Upcoming lensing surveys will provide data of unprecedented quality and enable us to conduct robust measurements of cosmic shear beyond the two-point level. These higher-order statistics contribute substantial information to cosmological constraints by breaking parameter degeneracies when combined with second-order shear statistics. Furthermore, three-point statistics have the potential to improve our understanding of systematics effects in the data, e.g., a detection of third-order B-modes can be an additional indicator for unsolved problems in the data analysis.

When extracting third-order information from a high-quality data set it is therefore essential to use robust and unbiased theoretical methods that meet the quality of the data. We have introduced the third-order ring statistics, which separates the shear 3PCF into third-order E/B- mode correlations on a finite interval $[\theta_{\rm min}, \theta_{\rm max}]$. Hence this statistic does not require knowledge of the 3PCF down to zero lag, where it is impossible to measure. Thus, unlike the third-order aperture mass statistics, it is not affected by apparent E/B-mode mixing (Kilbinger, Schneider and Eifler 2006).

Our main results are equations (3.21, 3.22, 3.23), which give compact expressions for the third-order ring statistics in terms of the shear 3PCF. Furthermore, in section 3.4.2 we give convenient expressions for computing the E-mode ring statistics from numerical simulations, and from the convergence bispectrum which facilitate the comparison with theoretical models for weak lensing three-point statistics (e.g., Valageas, Sato and Nishimichi 2011).

In addition to the cosmological information contained in the E-mode signal, our expression for third-order B-mode correlations opens a new window to detect remaining systematics in the data. For example, the various permutations of $\langle \mathcal{R}_{\rm E}\mathcal{R}_{\rm B}\mathcal{R}_{\rm E}\rangle$ allow for an association of B-modes with a specific angular scales.

For the analysis of future shear 3PCF measurements, we recommend using equation (3.11) to obtain a clean third-order E-mode signal, and equations (3.12, 3.15) to test for remaining B-mode correlations.

Acknowledgements

We thank Bhuvnesh Jain, Chris Hirata, and Mike Jarvis for useful discussions.

Chapter 4

Tidal Alignments as a Contaminant of the Galaxy Bispectrum [*]

4.1 Introduction

While the evolution of dark matter perturbations in the current ΛCDM model is well understood theoretically, the relation between the galaxy distribution and the large-scale (dark) matter distribution is complicated by the detailed physics of galaxy formation and different models may lead to different clustering properties of galaxies. In particular, while local theories of galaxy formation predict that the galaxy density fluctuations trace the matter fluctuations on large scales, they also predict that the two are related by the *bias parameter b*, which is in general not known *a priori* (Kaiser 1984). The unknown bias parameter represents a key problem for attempts to measure the growth of cosmological perturbations using galaxies.

In combination with the galaxy power spectrum, third-order galaxy clustering measures such as the bispectrum or (equivalently) the 3-point correlation function can be used to measure nonlinear galaxy bias and break the degeneracy between the normalization of the matter power spectrum, σ_8, and the linear galaxy bias. This enables one remove the effects of galaxy biasing and measure the cosmological growth of structure from the galaxy distribution (Fry 1994, Verde et al. 1998, Scoccimarro, Couchman and Frieman 1999, Verde, Heavens and Matarrese 2000), and thus constrain dark energy (e.g., Dolney, Jain and Takada 2006). Recently third order galaxy clustering has been analyzed by several authors using the bispectrum (Scoccimarro et al. 2001, Feldman et al. 2001, Verde et al. 2002, Kulkarni et al. 2007) and the three point correlation function (Jing and Börner

---

[*] This chapter was adapted from *Tidal alignments as a contaminant of the galaxy bispectrum*, Elisabeth Krause and Christopher M. Hirata; *MNRAS*, **410**, 2730 (2011). Reproduced here with permission, copyright (2010) by Wiley-Blackwell.

2004, Kayo et al. 2004, Nichol et al. 2006). Using mock catalogs from numerical simulations, Sefusatti et al. (2006) show that a combined analysis of the galaxy power spectrum and bispectrum including their cross-correlation contains significant information on galaxy bias and fundamental cosmological parameters and helps break parameter degeneracies of other cosmological probes.

The most important systematic errors in interpreting the observed galaxy clustering arise in the nonlinear regime,where the behavior of galaxy biasing and models of the (redshift space) galaxy power spectrum and bispectrum are difficult to model (see Smith, Sheth and Scoccimarro 2008, for the complications of a current model of the redshift space bispectrum). Recently Hirata (2009) showed that the alignment of galaxies by large-scale tidal fields can cause a systematic error in the determination of the linear redshift space distortion parameter β (Kaiser 1987): the alignment of galaxies with the tidal field (along the stretching axis of the field for large elliptical galaxies) in combination with a viewing direction dependent galaxy selection effect, e.g., preferential selection of galaxies which are observed along their long axis, will lead to a selection probability for galaxies which is modulated by the tidal field along the line of sight. This results in an anisotropy in redshift-space clustering with the same scale and angular dependence as the linear redshift-space effect. In this chapter we will explore the implications of such a tidal alignment contamination for the observed galaxy bispectrum and how it affects the measurement of galaxy bias parameters.

Throughout this work we assume a standard ΛCDM cosmology with the best-fit WMAP 7 (Komatsu et al. 2010) parameters, and assume Gaussian initial density perturbations.

We begin in section 4.2 with a derivation of the standard redshift space galaxy bispectrum and discuss toy models for physical processes that cause alignments of galaxy orientations with large-scale structure. In section 4.3 we explain how tidal alignments of galaxies in combination with an orientation dependent galaxy selection modify the observed galaxy distribution and calculate the corresponding corrections to the galaxy bispectrum. Using a Fisher matrix technique we then estimate the systematic error induced by tidal alignments to measurements of galaxy bias parameters from angular clustering in section 4.4. We conclude and discuss mitigation strategies in section 4.5.

4.2 Theoretical Background

In this section we derive the redshift space galaxy bispectrum to second-order in perturbation theory (for a review, see, e.g., Bernardeau et al. 2002), and discuss toy models for the alignment of galaxies with the large-scale tidal field.

4.2.1 Galaxy Bispectrum

The matter bispectrum B is defined as

$$\left\langle \tilde{\delta}(\mathbf{k}_1)\tilde{\delta}(\mathbf{k}_2)\tilde{\delta}(\mathbf{k}_3)\right\rangle \equiv (2\pi)^3 \delta_{\mathrm{D}}\left(\mathbf{k}_{123}\right) B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\ , \tag{4.1}$$

where $\tilde{\delta}(\mathbf{k})$ is the matter density contrast in Fourier space, δ_{D} the Dirac delta function, and $\mathbf{k}_{123} \equiv \mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$. The bispectrum vanishes for a Gaussian random field.

To second-order perturbation theory the density contrast is given by

$$\tilde{\delta}(\mathbf{k}) = \tilde{\delta}^{(1)}(\mathbf{k}) + \int \frac{\mathrm{d}^3\mathbf{k}_1}{(2\pi)^3} F_2\left(\mathbf{k}_1, \mathbf{k} - \mathbf{k}_1\right) \tilde{\delta}^{(1)}(\mathbf{k}_1)\tilde{\delta}^{(1)}(\mathbf{k} - \mathbf{k}_1), \tag{4.2}$$

with $\tilde{\delta}^{(1)}(\mathbf{k})$ the linear density contrast, and the second-order density kernel

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2k_1k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}\left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1k_2}\right)^2 . \tag{4.3}$$

Hence the matter bispectrum induced by nonlinear gravitational evolution at tree-level is given by

$$B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P(k_1)P(k_2) + 2 \text{ perm.}, \tag{4.4}$$

where $P(k)$ is the linear matter power spectrum, $\mathbf{k}_3 = -\mathbf{k}_1 - \mathbf{k}_2$ and “2 perm.” indicates that the 2 permutations $(\mathbf{k}_2, \mathbf{k}_3)$ and $(\mathbf{k}_1, \mathbf{k}_3)$ are also included in the summation.

Using the local bias approximation (e.g., Fry and Gaztanaga 1993), the galaxy density contrast δ_{g} can be expressed as a nonlinear function of the matter density contrast,

$$\delta_{\mathrm{g}}(\mathbf{x}) = b_1\delta(\mathbf{x}) + \frac{1}{2}b_2\delta(\mathbf{x})^2 + \cdots\ , \tag{4.5}$$

where the expansion coefficients are the linear (b_1) and nonlinear galaxy bias factors. In reality, galaxy biasing may be more complicated, especially on small scales, due to 1-halo terms (Seljak 2000) and nonlocal dependences such as the strength of the local tidal field (McDonald 2006, McDonald and Roy 2009). However, in simulations the local bias model is found to be a fair description of nonlinear halo clustering on large scales with an accuracy of a few percent (e.g., Marín et al. 2008, Guo and Jing 2009b, Manera and Gaztanaga 2009), which is sufficient at the level of this analysis.

Then the galaxy bispectrum B_g is related to the matter bispectrum via

$$B_\mathrm{g}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \simeq b_1^3 B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + b_1^2 b_2 \left(P(k_1)P(k_2) + 2\ \mathrm{perm.}\right), \tag{4.6}$$

and similarly for the galaxy power spectrum,

$$P_\mathrm{g}(k) = b_1^2 P(k). \tag{4.7}$$

To arrive at an expression for the redshift space galaxy bispectrum we have to transform radial coordinates to redshift space. In the plane-parallel approximation, the mapping from real space position $\mathbf{x}$ to coordinate $\mathbf{x}^\mathrm{s}$ in redshift space is given by

$$\mathbf{x}^\mathrm{s} = \mathbf{x} + \frac{\hat{\mathbf{n}} \cdot \mathbf{u}(\mathbf{x})}{Ha}\hat{\mathbf{n}}, \tag{4.8}$$

where $\mathbf{u}(\mathbf{x})$ is the peculiar velocity field, and $\hat{\mathbf{n}}$ is the direction of the line of sight. The velocity field is curl-free, $\nabla \times \mathbf{u}(\mathbf{x}) = 0$, at all orders in perturbation theory. Its divergence is given to linear-order in perturbation theory by

$$\mathrm{i}\mathbf{k} \cdot \tilde{\mathbf{u}}^{(1)}(\mathbf{k}) = aHf\tilde{\delta}^{(1)}(\mathbf{k}), \tag{4.9}$$

where $f = \mathrm{d}\ln(G)/\mathrm{d}\ln(a)$ is the logarithmic growth rate of linear perturbations (equal to roughly $\Omega_\mathrm{m}^{0.6}$ in general relativity). Higher-order contributions to $\nabla \cdot \mathbf{u}$ (Bernardeau et al. 2002) are analogous to equation (4.2), e.g.,

$$\mathrm{i}\mathbf{k} \cdot \tilde{\mathbf{u}}^{(2)}(\mathbf{k}) = aHf \int \frac{\mathrm{d}^3\mathbf{k}_1}{(2\pi)^3} G_2(\mathbf{k}_1, \mathbf{k} - \mathbf{k}_1)\tilde{\delta}^{(1)}(\mathbf{k}_1)\tilde{\delta}^{(1)}(\mathbf{k} - \mathbf{k}_1), \tag{4.10}$$

with the kernel

$$G_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{3}{7} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2k_1k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{4}{7}\left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1k_2}\right)^2. \tag{4.11}$$

Taking into account the Jacobian of this mapping of $\mathbf{x} \to \mathbf{x}^\mathrm{s}$ (equation 4.8), and approximating the peculiar velocity field by the second-order bulk velocity field, the galaxy density is redshift space is (Heavens, Matarrese and Verde 1998, Scoccimarro, Couchman and Frieman 1999)

$$\tilde{\delta}_\mathrm{g}^\mathrm{s}(\mathbf{k}^\mathrm{s}) = (b_1 + f\mu^2)\tilde{\delta}^{(1)}(\mathbf{k}^\mathrm{s}) + \int \frac{\mathrm{d}^3\mathbf{k}_1^\mathrm{s}}{(2\pi)^3} Z_2\left(\mathbf{k}_1^\mathrm{s}, \mathbf{k}^\mathrm{s} - \mathbf{k}_1^\mathrm{s}\right)\tilde{\delta}^{(1)}(\mathbf{k}_1^\mathrm{s})\tilde{\delta}^{(1)}(\mathbf{k}^\mathrm{s} - \mathbf{k}_1^\mathrm{s}), \tag{4.12}$$

where $\mathbf{k}^\mathrm{s}$ denotes a Fourier mode in redshift space, and $\mu \equiv \hat{\mathbf{k}} \cdot \hat{\mathbf{n}}$ is the cosine of the angle between the wave vector and the line of sight (we may analogously define μ_1, μ_{12}, etc.). The mode-coupling function Z_2 is

$$Z_2(\mathbf{k}_1, \mathbf{k}_2) = b_1 F_2(\mathbf{k}_1, \mathbf{k}_2) + f\mu_{12}^2 G_2(\mathbf{k}_1, \mathbf{k}_2) + \frac{f\mu_{12}k_{12}}{2}\left[\frac{\mu_1}{k_1}(b_1 + f\mu_2^2) + \frac{\mu_2}{k_2}(b_1 + f\mu_1^2)\right] + \frac{b_2}{2}, \quad (4.13)$$

Hence we can write the redshift space galaxy bispectrum as

$$B_{\mathrm{g}}^{\mathrm{s}}(\mathbf{k}_1^{\mathrm{s}}, \mathbf{k}_2^{\mathrm{s}}, \mathbf{k}_3^{\mathrm{s}}) \;=\; 2(b_1 + f\mu_1^2)(b_1 + f\mu_2^2)P(k_1^{\mathrm{s}})P(k_2^{\mathrm{s}})Z_2(\mathbf{k}_1^{\mathrm{s}}, \mathbf{k}_2^{\mathrm{s}}) + 2\,\mathrm{perm.}. \quad (4.14)$$

Note that this expression does not include the *Finger of God* effect due to the virialized motion of galaxies within a cluster (Jackson 1972), which is important when one of the k_i has a large line-of-sight component. While this effect is important even on weakly nonlinear scales, it is usually handled by phenomenological models (e.g., Hatton and Cole 1998, Verde et al. 1998, Scoccimarro, Couchman and Frieman 1999, Peacock et al. 2001), a compression of radial coordinates for galaxies living in the same cluster (e.g., Tegmark et al. 2004), or by reconstructing the redshift-space halo density field (Reid, Spergel and Bode 2009).

4.2.2 Toy Models of Tidal Alignments

4.2.2.1 Halo Shape Distortions: Linear Alignment

In the linear alignment model (Catelan, Kamionkowski and Blandford 2001) the shape and orientation of a galaxy are assumed to be determined by the shape of the halo it resides in. It is thought that the gravitational collapse of an initially spherical overdensity in a constant gravitational field leads to triaxial haloes, such that the halo will be prolate if the overdensity is stretched by the large-scale tidal field and oblate if it is compressed. This mechanism is believed to lead to a net correlation of halo orientations even though overdensities typically are not spherical, and such an alignment has been confirmed by simulations (e.g., Faltenbacher et al. 2009).

The relation between halo shape and galaxy shape is complicated by galaxy formation and differs between galaxy types (e.g., Faltenbacher et al. 2007), but at least for luminous red galaxies (LRGs) there is observational evidence for an alignment of the LRG with the major axis of its host (Binggeli 1982, Faltenbacher et al. 2007, Niederste-Ostholt et al. 2010). There are also correlations with large-scale structure (Binggeli 1982); with the Sloan Digital Sky Survey (SDSS) it has even been possible to measure the scale dependence of these correlations and show the consistency of their spectral index with the predictions of the linear tidal alignment model and the ΛCDM power spectrum (Heinis et al. 2007).

4.2.2.2 Tidal Torques: Quadratic Alignment

The orientation of a disc galaxy is determined by the direction of its angular momentum, which builds up due to tidal torquing during early stages of galaxy formation if the protogalaxy's inertia tensor is anisotropic and misaligned with the local shear field (Hoyle 1949, Sciama 1955, Peebles 1969, Doroshkevich 1970, White 1984, Crittenden et al. 2001). See Schäfer (2009) for a review of tidal torquing and the build up of angular momentum correlations.

Following Lee and Pen (2000), we parameterize the correlation between moment of inertia and the shear field by

$$\left\langle L_i L_j \right\rangle = \left\langle L^2 \right\rangle \left(\frac{1+\alpha}{3} \delta_{ij} - \alpha \hat{T}_{ih} \hat{T}_{hj} \right), \tag{4.15}$$

which is also the most general quadratic form possible. Here $\hat{T}_{ij}$ is the unit normalised traceless tidal field tensor ($\hat{T}_{ij}\hat{T}_{ij} = 1$) and α is a dimensionless coupling parameter, e.g. $\alpha = \frac{3}{5}$ at leading order in perturbation theory if shear and inertia tensor are mutually uncorrelated. It is also possible for α to be much smaller, e.g., if the angular momentum vector of the disk is only partially aligned with that of the host halo (e.g., van den Bosch et al. 2002).

Note that in nonlinear theory spin-induced alignments also have a linear contribution at large scales because the large-scale tidal field induces correlations of the small-scale tidal field and inertia tensor that lead to a nonzero contribution to $\langle L_i L_j \rangle$ (Hui and Zhang 2008), although this linear effect has not been observed for late-type galaxies despite several searches (Lee and Pen 2007, Heinis et al. 2007, Mandelbaum et al. 2011).

4.3 Tidal Alignment Contamination

As discussed in the previous section, the orientation of galaxies likely is not random but correlated with large-scale structure, and in combination with observational galaxy selection criteria which depend on the galaxy orientation relative to the line of sight, this may modify the observable galaxy distribution. Following Hirata (2009), we will now introduce the basic notation needed to discuss galaxy orientation and viewing direction dependent selection effects.

Let the galaxy orientation be described by the Euler angles (θ, ϕ, ψ) through a rotation matrix

$\mathbf{Q}(\theta, \phi, \psi)$. This matrix transforms "lab" frame coordinates to a coordinate system aligned with the galaxy. Due to tidal alignments the probability distribution $p(\mathbf{Q}|\mathbf{x})$ for the orientation of a galaxy at position $\mathbf{x}$ may be anisotropic and a function of the local environment of $\mathbf{x}$. The observational galaxy selection probability depends on the direction of the line of sight, $\hat{\mathbf{n}}$, and the galaxy orientation, specifically on the direction of the line of sight in the galaxy frame $\mathbf{Q}\hat{\mathbf{n}}$. We define

$$P \propto 1 + \Upsilon\left(\mathbf{Q}\hat{\mathbf{n}}, \mathbf{x}\right), \tag{4.16}$$

where the anisotropic part Υ is zero when averaged over all possible galaxy orientations or viewing directions.

The observable galaxy distribution $N(\text{selected})$ hence is modified compared to the true galaxy distribution $N(\text{true})$ by

$$\begin{aligned} \frac{N(\text{selected})}{N(\text{true})}(\hat{\mathbf{n}}|\mathbf{x}) &\propto \int_{\mathrm{SO}(3)} p(\mathbf{Q}|\mathbf{x})\left[1 + \Upsilon\left(\mathbf{Q}\hat{\mathbf{n}}, \mathbf{x}\right)\right] \, \mathrm{d}^3\mathbf{Q} \\ &= 1 + \int_{\mathrm{SO}(3)} p(\mathbf{Q}|\mathbf{x})\Upsilon\left(\mathbf{Q}\hat{\mathbf{n}}, \mathbf{x}\right) \, \mathrm{d}^3\mathbf{Q} \\ &\equiv 1 + \epsilon(\hat{\mathbf{n}}|\mathbf{x}), \end{aligned} \tag{4.17}$$

which is the average of equation (4.16) over the distribution of galaxy orientations, and where we have defined the orientation dependent selection function $\epsilon(\hat{\mathbf{n}}|\mathbf{x})$ in the last step. As the average of Υ over all galaxy orientations vanishes, equation (4.17) implies that ϵ vanishes if either the galaxy orientations are isotropically distributed or if the probability for selecting a galaxy is independent of $\mathbf{Q}\hat{\mathbf{n}}$, i.e., if $\Upsilon = 0$.

The observed galaxy density is modified by the orientation dependent selection function such that

$$1 + \delta_{\mathrm{g}}^{\mathrm{obs}}(\mathbf{x}^{\mathrm{s}}) = \left(\left[1 + \delta_{\mathrm{g}}(\mathbf{x})\right]\left[1 + \epsilon(\hat{\mathbf{n}}|\mathbf{x})\right]\right)^{\mathrm{s}}, \tag{4.18}$$

where the term in round brackets is the orientation modulated real space density of selected galaxies, and where the superscript s denotes the transform to redshift space. Expanding to second-order in the matter density field, this implies

$$\tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}}(\mathbf{k}^{\mathrm{s}}) = \tilde{\delta}_{\mathrm{g}}^{\mathrm{s}(1)}(\mathbf{k}^{\mathrm{s}}) + \tilde{\epsilon}^{\mathrm{s}(1)}(\hat{\mathbf{n}}|\mathbf{k}^{\mathrm{s}}) + \tilde{\delta}_{\mathrm{g}}^{\mathrm{s}(2)}(\mathbf{k}^{\mathrm{s}}) + \tilde{\epsilon}^{\mathrm{s}(2)}(\hat{\mathbf{n}}|\mathbf{k}^{\mathrm{s}}) + \int \frac{\mathrm{d}^3\mathbf{k}_1^{\mathrm{s}}}{(2\pi)^3}\tilde{\delta}_{\mathrm{g}}^{\mathrm{s}(1)}(\mathbf{k}_1^{\mathrm{s}})\tilde{\epsilon}^{\mathrm{s}(1)}(\hat{\mathbf{n}}|\mathbf{k}^{\mathrm{s}} - \mathbf{k}_1^{\mathrm{s}}). \tag{4.19}$$

In the following we calculate the impact of an orientation dependent selection function on the galaxy bispectrum by introducing models for the anisotropic galaxy selection function which are based on symmetry considerations and motivated by the toy models of tidal alignment discussed in section 4.3. First we extend the linear alignment model from Hirata (2009) to second-order in the density field, and then construct a new model the anisotropic galaxy selection function due to quadratic alignment.

4.3.1 Linear Alignment

In this subsection we construct a model for the anisotropic galaxy selection function ϵ based on the assumptions that the large-scale tidal fields induce a preferred direction in galaxy formation, and that the alignment is of linear-order in the tidal field. Additionally we require the average of $\epsilon(\hat{\mathbf{n}}|\mathbf{x})$ over the sky to vanish. Then the only possible contraction of the tidal field with the viewing direction $\hat{\mathbf{n}}$ is

$$\begin{aligned}\epsilon(\hat{\mathbf{n}}|\mathbf{x}) &= \frac{A_1}{4\pi G a^2 \bar{\rho}_{\mathrm{m}}(a)}\left(\hat{n}_i\hat{n}_j\nabla_i\nabla_j - \frac{1}{3}\nabla^2\right)\Psi(\mathbf{x}) \\ &= A_1\hat{n}_i\hat{n}_j\left(\nabla_i\nabla_j\nabla^{-2} - \frac{1}{3}\delta_{ij}\right)\delta(\mathbf{x}),\end{aligned} \tag{4.20}$$

where Ψ is the Newtonian potential, a is the scale factor, and where we have used the Poisson equation to write ϵ in terms of the dimensionless tidal field. A_1 is an expansion coefficient which encodes the degree to which galaxy orientations are non-random due to tidal fields and the strength of galaxy orientation-dependent selection effects. Note that *both* effects need to be present in order to have $A_1 \neq 0$.

To second-order in the linear matter density field the anisotropic selection function in Fourier space can be written as

$$\tilde{\epsilon}(\hat{\mathbf{n}}|\mathbf{k}) \approx A_1\left[\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{k}}\right)^2 - \frac{1}{3}\right]\left[\tilde{\delta}^{(1)}(\mathbf{k}) + \tilde{\delta}^{(2)}(\mathbf{k})\right]. \tag{4.21}$$

This expression is transformed to redshift space by Taylor expanding the real space expression and using equations (4.8, 4.9)

$$\begin{aligned}\epsilon^{\mathrm{s}}(\hat{\mathbf{n}}|\mathbf{x}^{\mathrm{s}}) &= \epsilon(\hat{\mathbf{n}}|\mathbf{x}) \approx \epsilon(\hat{\mathbf{n}}|\mathbf{x}^{\mathrm{s}}) + (\mathbf{x}-\mathbf{x}^{\mathrm{s}})\cdot\nabla\epsilon(\hat{\mathbf{n}}|\mathbf{x}^{\mathrm{s}}) + O(\delta^3) \\ &= \epsilon(\hat{\mathbf{n}}|\mathbf{x}^{\mathrm{s}}) + f\hat{\mathbf{n}}\cdot\nabla\,\nabla^{-2}\delta^{(1)}(\mathbf{x}^{\mathrm{s}})\,\hat{\mathbf{n}}\cdot\nabla\epsilon(\hat{\mathbf{n}}|\mathbf{x}^{\mathrm{s}})\,,\end{aligned} \tag{4.22}$$

and hence in Fourier space

$$\begin{aligned}\tilde{\epsilon}^{\mathrm{s}(1)}(\hat{\mathbf{n}}|\mathbf{k}^{\mathrm{s}}) &= \tilde{\epsilon}^{(1)}(\hat{\mathbf{n}}|\mathbf{k}^{\mathrm{s}}) && (4.23)\\ \tilde{\epsilon}^{\mathrm{s}(2)}(\hat{\mathbf{n}}|\mathbf{k}^{\mathrm{s}}) &= \tilde{\epsilon}^{(2)}(\hat{\mathbf{n}}|\mathbf{k}^{\mathrm{s}}) + \int \frac{\mathrm{d}^3\mathbf{k}_1^{\mathrm{s}}}{(2\pi)^3} f\mu_1\mu_{\mathbf{k}^{\mathrm{s}}-\mathbf{k}_1^{\mathrm{s}}} \frac{k_1^{\mathrm{s}}}{|\mathbf{k}^{\mathrm{s}}-\mathbf{k}_1^{\mathrm{s}}|}\tilde{\delta}^{(1)}(\mathbf{k}^{\mathrm{s}}-\mathbf{k}_1^{\mathrm{s}})\tilde{\epsilon}^{(1)}(\hat{\mathbf{n}}|\mathbf{k}_1^{\mathrm{s}}). && (4.24)\end{aligned}$$

Using this form for the selection function in combination with equation (4.19), we now calculate the galaxy bispectrum modulated by linear tidal alignments. Then the first-order observed density contrast is given by

$$\tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}(1)}(\mathbf{k}^{\mathrm{s}}) = \tilde{\delta}^{(1)}(\mathbf{k})\left(b_1 - \frac{1}{3}A_1 + (A_1+f)\mu_1^2\right) . \tag{4.25}$$

The different terms contributing to the observed galaxy bispectrum can be calculated as

$$\begin{aligned}\left\langle \tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}(1)}(\mathbf{k}_1^{\mathrm{s}})\tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}(1)}(\mathbf{k}_2^{\mathrm{s}})\ \tilde{\delta}_{\mathrm{g}}^{\mathrm{s}(2)}(\mathbf{k}_3^{\mathrm{s}})\right\rangle &= (2\pi)^3\delta_{\mathrm{D}}\left(\mathbf{k}_{123}^{\mathrm{s}}\right)P(k_1^{\mathrm{s}})P(k_2^{\mathrm{s}})\left\{2\ Z_2(\mathbf{k}_1^{\mathrm{s}},\mathbf{k}_2^{\mathrm{s}})\right\}\\ &\quad\times\left(b_1-\frac{1}{3}A_1+(A_1+f)\mu_1^2\right)\left(b_1-\frac{1}{3}A_1+(A_1+f)\mu_2^2\right) ,\\ \left\langle \tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}(1)}(\mathbf{k}_1^{\mathrm{s}})\tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}(1)}(\mathbf{k}_2^{\mathrm{s}})\ \tilde{\epsilon}^{\mathrm{s}(2)}(\hat{\mathbf{n}}|\mathbf{k}_3^{\mathrm{s}})\right\rangle &= (2\pi)^3\delta_{\mathrm{D}}\left(\mathbf{k}_{123}^{\mathrm{s}}\right)P(k_1^{\mathrm{s}})P(k_2^{\mathrm{s}})\left\{2A_1\left(\mu_{12}^2-\frac{1}{3}\right)F_2(\mathbf{k}_1,\mathbf{k}_2)\right.\\ &\quad\left.+A_1f\mu_1\mu_2\frac{k_1^{\mathrm{s}}}{k_2^{\mathrm{s}}}\left(\mu_1^2-\frac{1}{3}\right)+A_1f\mu_1\mu_2\frac{k_2^{\mathrm{s}}}{k_1^{\mathrm{s}}}\left(\mu_2^2-\frac{1}{3}\right)\right\}\\ &\quad\times\left(b_1-\frac{1}{3}A_1+(A_1+f)\mu_1^2\right)\left(b_1-\frac{1}{3}A_1+(A_1+f)\mu_2^2\right) ,\end{aligned} \tag{4.26}$$

and the contribution from the last term in equation (4.19) containing a convolution of first-order density contrast and anisotropic selection function,

$$\begin{aligned}\left\langle \tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}(1)}(\mathbf{k}_1^{\mathrm{s}})\tilde{\delta}_{\mathrm{g}}^{\mathrm{obs}(1)}(\mathbf{k}_2^{\mathrm{s}})\ \left(\tilde{\delta}_{\mathrm{g}}^{\mathrm{s}(1)}\otimes\tilde{\epsilon}^{\mathrm{s}(1)}\right)(\hat{\mathbf{n}},\mathbf{k}_3^{\mathrm{s}})\right\rangle &= (2\pi)^3\delta_{\mathrm{D}}\left(\mathbf{k}_{123}^{\mathrm{s}}\right)P(k_1^{\mathrm{s}})P(k_2^{\mathrm{s}})\\ &\quad\times\left(b_1-\frac{1}{3}A_1+(A_1+f)\mu_1^2\right)\left(b_1-\frac{1}{3}A_1+(A_1+f)\mu_2^2\right)\\ &\quad\times A_1\left\{\left(b_1+f\mu_1^2\right)\left(\mu_2-\frac{1}{3}\right)+\left(b_1+f\mu_2^2\right)\left(\mu_1-\frac{1}{3}\right)\right\}.\end{aligned} \tag{4.27}$$

Hence the galaxy bispectrum modulated by linear tidal alignments is given by

$$\begin{aligned}
B_{\rm g}^{\rm s,LA}(\mathbf{k}_1^{\rm s},\mathbf{k}_2^{\rm s},\mathbf{k}_3^{\rm s}) \;=\;& P(k_1^{\rm s})P(k_2^{\rm s})\left[b_1-\frac{A_1}{3}+(A_1+f)\mu_1^2\right]\left[b_1-\frac{A_1}{3}+(A_1+f)\mu_2^2\right]\\
&\times\Bigg\{2Z_2(\mathbf{k}_1^{\rm s},\mathbf{k}_2^{\rm s})+2A_1\left(\mu_{12}^2-\frac{1}{3}\right)F_2(\mathbf{k}_1^{\rm s},\mathbf{k}_2^{\rm s})+A_1b_1\left(\mu_1^2+\mu_2^2-\frac{2}{3}\right)\\
&\quad+A_1f\left[\mu_1\mu_2\left(\frac{k_2^{\rm s}}{k_1^{\rm s}}\left(\mu_2^2-\frac{1}{3}\right)+\frac{k_1^{\rm s}}{k_2^{\rm s}}\left(\mu_1^2-\frac{1}{3}\right)\right)+\left(6\mu_1^2\mu_2^2-\mu_1^2-\mu_2^2\right)/3\right]\Bigg\}\\
&+2\text{ perm.}
\end{aligned}\tag{4.28}$$

4.3.1.1 Transverse Galaxy Bispectrum

As the full redshift space bispectrum is a complicated function of configurations described by 5 parameters (3 parameters specifying triangle shape, and 2 angles describing the orientation with respect to the line of sight), we will now simplify equation (4.28) by considering only triangles in the plane of the sky ($\mu_i = 0$), which are the easiest to model and are the triangles observed in photometric redshift surveys. In this case, we find a galaxy bispectrum

$$B_{\rm g}^{\rm LA,\perp}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)=\left(b_1-\frac{A_1}{3}\right)^2\left[2\left(b_1-\frac{A_1}{3}\right)F_2\left(\mathbf{k}_1,\mathbf{k}_2\right)+b_2-\frac{2}{3}A_1b_1\right]P(k_1)P(k_2)+2\text{ perm.}$$

Comparing this expression to equation (4.6), one finds that the effect of linear tidal alignments on the transverse galaxy bispectrum can be described as a rescaling of the galaxy bias parameters

$$b_1\to b_1-\frac{A_1}{3},\quad b_2\to b_2-\frac{2}{3}A_1b_1.\tag{4.29}$$

Hirata (2009) found that the same rescaling of b_1 applies to the real-space ($\mu_i = 0$) galaxy power spectrum. Therefore, *the use of the real-space power spectrum and bispectrum to eliminate galaxy bias parameters and extract σ_8 is robust against linear tidal alignments*. However, this robustness does not extend to the $\mu_i \neq 0$ modes.

For later use, we also write out the systematic error in the transverse galaxy bispectrum induced by linear alignment

$$\Delta B_{\rm g}^{\rm LA,\perp}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)=\left[2\left(b_1^2A_1-b_1\frac{A_1^2}{3}+\frac{A_1^3}{27}\right)F_2\left(\mathbf{k}_1,\mathbf{k}_2\right)-b_1\frac{A_1}{3}\right]P(k_1)P(k_2)\;+2\text{ perm.}$$

4.3.1.2 Normalization

Following Hirata (2009), we use $A_1 \approx -0.024$ for LRG-type elliptical galaxies. This is a rough estimate which is based on the assumption that elliptical galaxies are optically thin triaxial systems, that the deviation from spherical symmetry can on average be related to the tidal field (with correlation strength B), on different models for the orientation dependence of a galaxy's apparent magnitude (parametrized by χ), and the slope of the galaxy luminosity function η:

$$A_1 = 2\eta\chi B \, . \tag{4.30}$$

While the total flux of an optically thin galaxy is not affected by tidal alignments, the average isophotal ellipticity and projected effective radius of a galaxy become a function of the tidal field.

The selection of galaxies in a survey will be modified by tidal alignment if part of the selection criteria is a magnitude cut, and if the apparent magnitude of a galaxy depends on its orientation. The apparent magnitude of a galaxy is nearly orientation independent if measured using Petrosian magnitudes or model magnitudes which are based on an accurate model for the radial profile, then at the level of the toy model considered by Hirata (2009) $\epsilon \approx 0$.

If galaxies are selected using isophotal magnitudes or aperture magnitudes, more light will be counted if a galaxy is viewed along its long axis than its short axis. The selection factor χ in equation (4.30) depends on the method used to measure galaxy fluxes (c.f. figure 2 in Hirata 2009), and it translates the fractional change in effective radius induced by intrinsic alignment to a fractional change in measured flux. This change in measured flux moves galaxies across the selection threshold, and it is translated into change in number density by assuming a luminosity function with slope $-\eta$.

The strength of the tidal alignment effect B is determined from measurements of the density-ellipticity cross-correlation function (Heinis et al. 2007). Our chosen normalization further assumes a LRG luminosity function with $\eta = 4.0$ and galaxy selection based on isophotal magnitudes measured within ~ 3 effective radii. Also note that this normalization is based on observations around $z = 0.3$, and should only be used near this redshift as the LRG luminosity function and the correlation between tidal field and galaxy orientation may show strong evolution with redshift.

4.3.2 Quadratic Alignment

The leading-order alignment of galactic angular momentum in tidal torque theories is quadratic in the tidal tensor because of the need for both a tidal field and an anisotropic inertia tensor on which it can act.

The anisotropic selection function for a disc galaxy is generally a function of its inclination i (defined by $\cos i = \hat{\mathbf{L}} \cdot \hat{\mathbf{n}}$). While i is in the range $0 \leq i \leq \pi$, we expect most selection criteria to be symmetric with respect to an observer being above or below the plane of the target, so it follows that the anisotropic part of the selection function contains only even-order spherical harmonics:

$$\Upsilon(\mathbf{Q}\hat{\mathbf{n}}, \mathbf{x}) = \sum_{J\geq 2,\,\mathrm{even}} c_J P_J(\cos i), \tag{4.31}$$

where P_J is a Legendre polynomial. Using equation (4.17), and noting that for a disk galaxy, we may replace the general integration over orientations $\mathbf{Q}$ ∈SO(3) with an integration over directions of the angular momentum vector $\hat{\mathbf{L}} \in S^2$, we may write

$$\epsilon(\hat{\mathbf{n}}|\mathbf{x}) = \sum_{J\geq 2,\,\mathrm{even}} c_J \int_{S^2} p(\hat{\mathbf{L}}|\mathbf{x}) P_J(\cos i)\, \mathrm{d}^2\hat{\mathbf{L}}. \tag{4.32}$$

Because the quadratic alignment model contains two factors of the tidal field, which are spin 2, $p(\hat{\mathbf{L}}|\mathbf{x})$ can contain spherical harmonics only through order $J \leq 4$. For simplicity, we will focus only on the quadrupolar $J = 2$ term in the sum (while noting that the hexadecapolar alignment $J = 4$ is in principle possible). Then equation (4.32) implies that

$$\epsilon(\hat{\mathbf{n}}|\mathbf{x}) \propto \langle P_2(\hat{\mathbf{L}} \cdot \hat{\mathbf{n}}) \rangle, \tag{4.33}$$

where the average is taken over the local probability distribution of $\hat{\mathbf{L}}$. Equivalently, using equation (4.15), we find that

$$\epsilon(\hat{\mathbf{n}}|\mathbf{x}) = \tilde{A}_2 \left(\hat{n}_i \hat{n}_j - \frac{1}{3}\delta_{ij} \right) \hat{T}_{ih} \hat{T}_{hj}. \tag{4.34}$$

We relate $\hat{T}_{ij}$ to the dimensionless shear field tensor T_{ij},

$$\begin{aligned} \tilde{T}_{ij}(\mathbf{k}) &= \frac{1}{4\pi G a^2 \bar{\rho}_{\mathrm{m}}(a)} \left(k_i k_j - \frac{1}{3}\delta_{ij} k^2 \right) \tilde{\Psi}(\mathbf{k}) \\ &= \left(\hat{k}_i \hat{k}_j - \frac{1}{3}\delta_{ij} \right) \tilde{\delta}(\mathbf{k}), \end{aligned} \tag{4.35}$$

by approximating the scalar $T^2 \equiv T_{ij}T_{ji}$ with its expected value C^2:

$$C^2 \equiv \langle T^2 \rangle = \frac{2}{3}\sigma^2(R), \tag{4.36}$$

i.e., we approximate $\hat{T}_{ij} \approx C^{-1}T_{ij}$. As this expression for the anisotropic selection function is already second-order in the density field, effects associated with mapping ϵ to redshift space only enter at higher-orders than considered in this analysis and in the following we will drop the superscript s to denote Fourier modes in redshift space.

Note that C^2 is proportional to the variance of the smoothed density field smoothed on the halo collapse scale R, since the density and tidal fields are both derived by taking second derivatives of the potential.

Then the contribution of quadratic alignment to the orientation dependent selection function can be written as

$$\begin{aligned}\tilde{\epsilon}^{(2)}(\hat{\mathbf{n}}|\mathbf{k}) &= \tilde{A}_2\left(\hat{n}_i\hat{n}_j - \frac{1}{3}\delta_{ij}\right)\int\frac{\mathrm{d}^3\mathbf{k}'}{(2\pi)^3}\hat{\tilde{T}}_{ih}(\mathbf{k})\hat{\tilde{T}}_{hj}(\mathbf{k}'') \\ &= A_2\hat{n}_i\hat{n}_j\int\frac{\mathrm{d}^3\mathbf{k}'}{(2\pi)^3}\left\{\left(\hat{k}'_i\hat{k}'_h - \frac{1}{3}\delta_{ih}\right)\left(\hat{k}''_h\hat{k}''_j - \frac{1}{3}\delta_{hj}\right) - \frac{1}{3}\delta_{ij}\left[\left(\hat{\mathbf{k}}'\cdot\hat{\mathbf{k}}''\right)^2 - \frac{1}{3}\right]\right\}\tilde{\delta}^{(1)}(\mathbf{k}')\tilde{\delta}^{(1)}(\mathbf{k}''),\end{aligned} \tag{4.37}$$

where $\mathbf{k}'' = \mathbf{k} - \mathbf{k}'$. This term contributes to the observed galaxy bispectrum via

$$\begin{aligned}\Delta B_{\mathrm{g}}^{\mathrm{QA}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) &= 2A_2\left(b_1 - \frac{A_1}{3} + (A_1+f)\mu_1^2\right)\left(b_1 - \frac{A_1}{3} + (A_1+f)\mu_2^2\right)P(k_1)P(k_2) \\ &\quad\times\left\{\mu_1\mu_2\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2 - \frac{1}{3}\left(\mu_1^2 + \mu_2^2 + (\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2)^2\right) + \frac{2}{9}\right\} \\ &\quad + 2\text{ perm.}\end{aligned} \tag{4.38}$$

Here $A_1 \neq 0$ if the galaxy population under consideration is also subject to linear alignment, and we have defined $A_2 \equiv \tilde{A}_2/C^2$.

4.3.2.1 Transverse Galaxy Bispectrum

The quadratic alignment model modifies the observed transverse galaxy bispectrum by

$$\Delta B_{\mathrm{g}}^{\mathrm{QA},\perp}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = \frac{2}{3}A_2 b_1^2\left[\frac{2}{3} - \left(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2\right)^2\right]P(k_1)P(k_2) + 2\text{ perm.} \tag{4.39}$$

Note that this systematic offset is independent of b_2, and its amplitude scales linearly with A_2 and quadratically with b_1. The systematic offset cannot be expressed as a simple rescaling of the galaxy bias parameters due to its shape dependence. Figure 4.1 illustrates its effect on the reduced transverse galaxy bispectrum

$$Q_{\mathrm{g}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{B_{\mathrm{g}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{P_{\mathrm{g}}(k_1)P_{\mathrm{g}}(k_2) + P_{\mathrm{g}}(k_1)P_{\mathrm{g}}(k_3) + P_{\mathrm{g}}(k_2)P_{\mathrm{g}}(k_3)}\ , \tag{4.40}$$

which is only mildly dependent on cosmology as the amplitude of fluctuations has been divided out. The shape and scale dependence of ΔQ_{g} is further illustrated in figure 4.2, which shows the systematic offset for all possible closed triangle configurations with $k_1 \geq k_2 \geq k3$, with the left plot showing triangles with $k_1 = 0.05h/\mathrm{Mpc}$ and the right plot showing triangles with $k_1 = 0.2h/\mathrm{Mpc}$. The systematic offset is negative for triangles which are close to collinear, and for the scales considered in this analysis it shows little scale dependence.

4.3.2.2 Normalization

Similar to the normalization of the linear alignment contamination outlined in Section 4.3.1.2, the magnitude of the observed contamination due to quadratic alignment again depends on (i) the orientation dependence of the recovered flux (continuum or line), (ii) the slope of the galaxy luminosity function, and (iii) the strength of the tidal alignment effect. We may use models for (i) and direct measurements for (ii), but (iii) is harder. For the linear alignment model we were able to use the observational constraints from the density-ellipticity cross-correlation function, but this is not an option here as the quadratic alignment contribution to two-point statistics vanish to leading order. Another option would be to set limits using the observed ellipticity variance, which must set an upper limit on α^2 (this was the approach followed in Crittenden et al. 2001 for estimating the intrinsic ellipticity correlation contamination of weak lensing surveys). We will take an even simpler approach here, and use some simple theoretical arguments on the value of α.

In the tidal torque model, the distribution of disk normal vectors $\hat{\mathbf{L}}$ given some tidal tensor $\hat{\mathbf{T}}$ can be approximated by (Crittenden et al. 2001)

$$p(\hat{\mathbf{L}}|\hat{\mathbf{T}}) \approx \frac{1}{4\pi}\left(1 + \frac{3\alpha}{2} - \frac{9\alpha}{2}\hat{L}_i\hat{L}_j\hat{T}_{ik}\hat{T}_{jk}\right). \tag{4.41}$$

For a geometrically thin disk with normal vector $\hat{\mathbf{L}}$ observed along the $\hat{z}$ axis, the inclination is

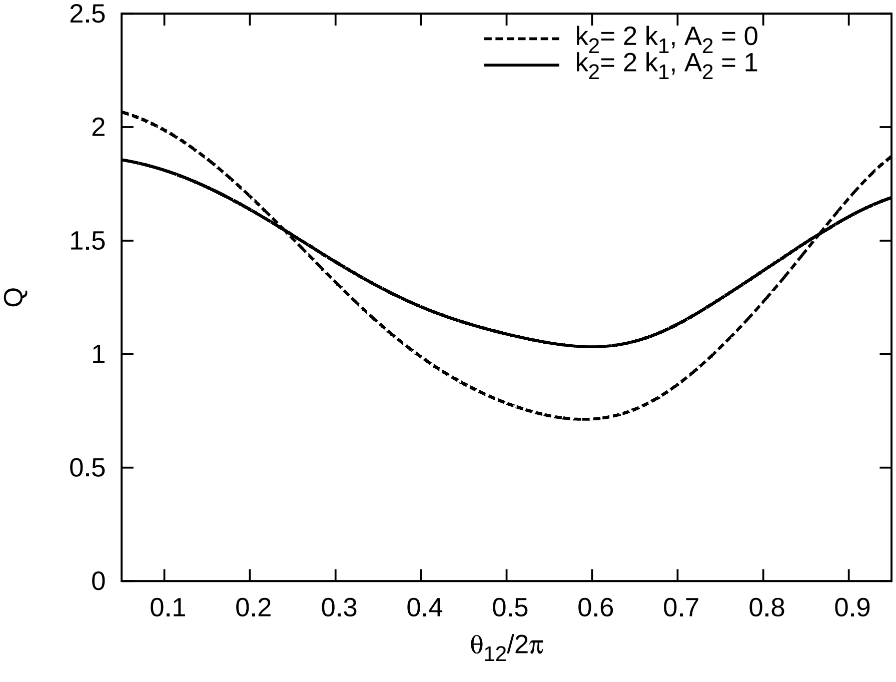


Figure 4.1 Effect of quadratic alignment on the reduced transverse galaxy bispectrum with $b_1 = 1$, $k_1 = 0.05h\,\mathrm{Mpc}^{-1}$, and where θ_{12} denotes the angle between $\mathbf{k}_1$ and $\mathbf{k}_2$, and for $A_2 = 1$.

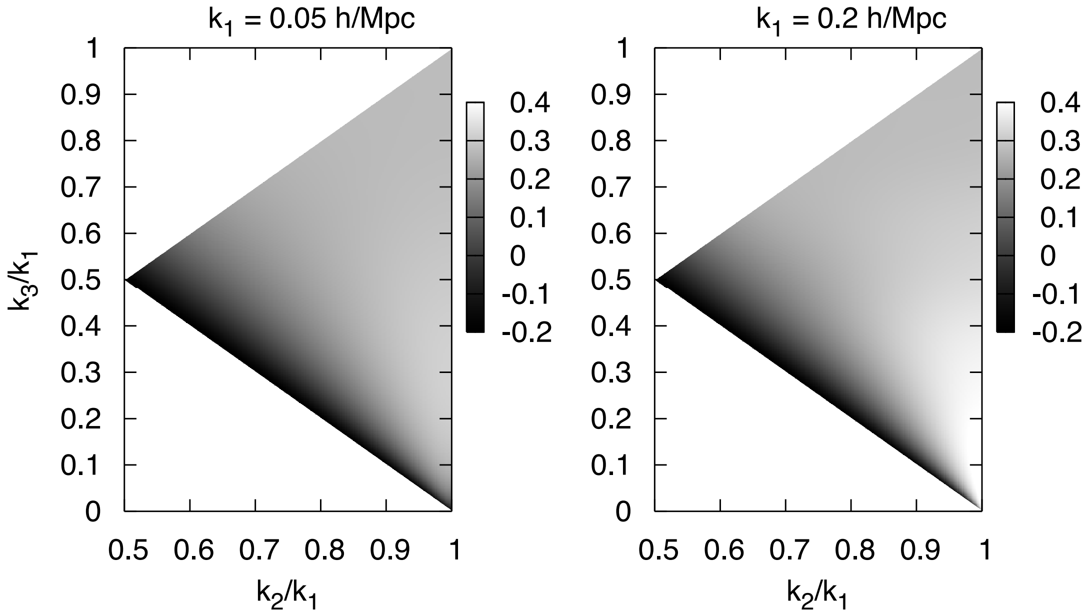


Figure 4.2 Systematic offset of the reduced transverse galaxy bispectrum due to quadratic alignment with $b_1 = 1$ and $A_2 = 1$ as a function of triangle shape and scale. Shown are all possible closed triangle configurations with $k_1 \geq k_2 \geq k3$ for a given k_1, areas in configuration space which do not correspond to a closed triangle are shown in white (located around the top and bottom left corner of each plot). Equilateral triangles are located in the upper right corner of the configuration space, isosceles triangles lie on the upper diagonal, and collinear ($\theta_{12} \rightarrow 0$) triangles near the lower diagonal.

$\cos i = \hat{L}_3$. The following constraints can be placed on α:

- Since $\hat{L}_i\hat{L}_j\hat{T}_{ik}\hat{T}_{jk}$ can take on any value between 0 and $\frac{2}{3}$, the requirement that $p(\hat{\mathbf{L}}|\hat{\mathbf{T}}) \geq 0$ sets the constraint $|\alpha| \leq \frac{2}{3}$.
- If one neglects correlations between the external tidal field and the moment of inertia tensor of the collapsing protogalaxy, one finds $\alpha = \frac{3}{5}$ (Lee and Pen 2000).
- The angular momentum of the disc of a galaxy may be disaligned from that of its host halo, due to e.g., torques between the disc and halo, or due to the disc containing only a specially selected subset of the halo's baryons. For a Gaussian distribution of disalingment angles with rms per axis Θ, the JM spherical harmonic component of $p(\hat{\mathbf{L}}|\hat{\mathbf{T}})$ is suppressed by a factor of $\exp[-J(J+1)\Theta^2/2]$; since we have a quadrupolar anisotropy ($J = 2$), α is suppressed by a factor of $\exp(-3\Theta^2)$.

The above arguments suggest that $|\alpha|$ of several tenths is plausible, but in no case should it exceed $\frac{2}{3}$. Also, while the simplest version of the tidal torque hypothesis implies $\alpha > 0$, there is no physical reason why negative values should not be allowed.

Next we determine the relation between an inclination dependent observed flux and the selection function ϵ: Assume a galaxy flux distribution with slope $\mathrm{d}\ln\bar{n}/\mathrm{d}\ln F_{\mathrm{m}in} = -\eta$. Then the number density of galaxies per logarithmic range in the intrinsic flux F_i per unit solid angle of disk orientation is

$$\mathcal{N}(F_\mathrm{i}, \hat{\mathbf{L}}) \propto F_\mathrm{i}^{-\eta} p(\hat{\mathbf{L}}|T) \,. \tag{4.42}$$

Let the observed, inclination dependent flux be $F(i) = F_\mathrm{i}\Phi(i)$. The number density of galaxies above some threshold flux F_0 then evaluates to

$$\begin{aligned} \mathcal{N}(> F_0) &\propto \int \mathrm{d}^2\hat{\mathbf{L}} \int_{F_0/\Phi(i)}^{\infty} \mathrm{d}\ln F_\mathrm{i} F_\mathrm{i}^{-\eta} p(\hat{\mathbf{L}}|\hat{\mathbf{T}}) \\ &\propto \int_0^{\pi} [\Phi(i)]^{\eta} \left[1 - \frac{9\alpha}{2}\left(\hat{T}_{3j}^2 - \frac{1}{3}\right) P_2(\cos(i))\right] \sin i \,\mathrm{d}i, \end{aligned} \tag{4.43}$$

where we have performed both the integral over ϕ and over F_i (since the latter is simply a power law), and defined $\hat{T}_{3j}^2 \equiv \hat{T}_{3j}\hat{T}_{3j}$. Defining

$$\psi = \frac{\int_0^{\pi} [\Phi(i)]^{\eta} P_2(\cos i) \sin i \,\mathrm{d}i}{\int_0^{\pi} [\Phi(i)]^{\eta} \sin i \,\mathrm{d}i}, \tag{4.44}$$

the anisotropic part of the observed galaxy count can be written as

$$\epsilon(\hat{z}|\mathbf{x}) = -\frac{9\alpha}{2}\psi\left(\hat{T}^2_{3j} - \frac{1}{3}\right). \tag{4.45}$$

Combining this with equation (4.34), we conclude that $\tilde{A}_2 = -\frac{9}{2}\alpha\psi$, and hence

$$A_2 = -\frac{9}{2}\frac{\alpha}{C^2}\psi = -\frac{27}{4}\frac{\alpha\psi}{\sigma^2_\delta(R)}. \tag{4.46}$$

The top-hat variance is related to the bias of the galaxies if the mass function is nearly universal (Sheth and Tormen 1999); for example, at $b = 1$ we have $\sigma^2_\delta(R) = 2.96$, whereas at $b = 2$ we have $\sigma^2_\delta(R) = 0.83$.

The last step in obtaining a numerical estimate for A_2 is evaluating the orientation dependent selection factor ψ. This requires a model for the angular distribution of emitted radiance $\Phi(i)$, which also determines the selection probability $p(i) \propto [\Phi(i)]^\eta$. Several geometric toy models for the vertical distributions of emitters and dust are discussed by Hirata (2009), and for galaxy distributions with $\eta \approx 2$ (appropriate for [O II] and Hα surveys), ψ is found to be of order a few tenths: for example, it is $\psi = 0.4$ in the optically thick slab model; $\psi = 0.23$ (0.30) in the uniform slab model with normal optical depth $\tau = 0.5$ (1.0); and $\psi = 0.26$ (0.37) in the sheet-in-slab model with $\tau = 0.5$ (1.0).

These toy models suggest that A_2 will be of order unity and we assume $A_2 = 1$ for illustrative purpose in the following analysis.[1] For application to any survey the normalization must be calculated based on the detailed selection criteria and galaxy distribution.

4.4 Fisher Matrix Analysis

We now estimate the parameter bias induced by a tidal alignment contamination by performing a Fisher matrix analysis for a survey with characteristics similar to the Dark Energy Survey (DES)[2], assuming that one would use the angular bispectrum of a slice of galaxies in photometric redshift space. A spectroscopic survey covering a similar volume and oversampling the density field ($nP > 1$) would of course yield tighter constraints, but a full Fisher analysis of such a survey including redshift space distortions and finger-of-God parameters is beyond the scope of this analysis.

---

[1] In principle, either *sign* of A_2 is allowed by our above calculations; for negative A_2 the direction of the parameter biases should be reversed.

[2] URL: `http://www.darkenergysurvey.org/`

4.4.1 Survey Characteristics and Analysis Details

Our fictitious survey has the same area as the DES, $\Omega = 5000$ square degrees. We assume a constant comoving galaxy density over the redshift range of interest and use a radial galaxy selection function of the form expected for the DES (Nock, Percival and Ross 2010),

$$\frac{\mathrm{d\,Prob}}{dz} \propto \left(\frac{z}{0.5}\right)^2 \exp\left(-\frac{z}{0.5}\right)^{1.5} , \tag{4.47}$$

In order to project out redshift space distortions we consider the angular clustering of galaxies projected over a finite radial distance. For our theoretical modeling the projection over a finite range in radial distance is equivalent to a projection over a finite redshift range, and we choose $0.4 \leq z \leq 0.6$. Observationally, this mapping is complicated by the distribution of photometric redshifts and the effect of redshift space distortions on the boundary of a region selected in redshift space (e.g., Padmanabhan et al. 2007, Nock, Percival and Ross 2010).

4.4.1.1 Binned Angular Multispectra and Covariances

We calculate the angular power and multispectra $\mathcal{P}_N$ using the Limber equation in Fourier space (Kaiser 1992, Buchalter, Kamionkowski and Jaffe 2000):

$$\mathcal{P}_N\left(\mathbf{l}_1 \ldots \mathbf{l}_N\right) = \int_{z=0.4}^{z=0.6} \mathrm{d}\chi\, \frac{\phi^N(\chi)}{\chi^{2N-2}} P_N\left(\frac{\mathbf{l}_1}{\chi}, \ldots, \frac{\mathbf{l}_N}{\chi}; \chi\right), \tag{4.48}$$

where P_N is the three dimensional N-point correlation function in Fourier space. In the following we use $\mathcal{P}$, $\mathcal{B}$, $\mathcal{T}$ to denote the angular *galaxy* power spectrum, bispectrum and trispectrum.

For a linear alignment contamination, the change in the observed angular galaxy bispectrum is described by the same bias parameter rescaling (equation (4.29)) as for the transverse galaxy bispectrum discussed above. The magnitude of the systematic offset in the angular galaxy bispectrum induced by a quadratic alignment contamination is proportional to $A_2 b_1^2$ and independent of b_2. As the angular projection mixes different physical scales, the exact configuration dependence and normalization of the angular bispectrum contamination depends strongly on the radial selection function (for details see Fry and Thomas 1999). As can be seen from figure 4.2 the systematic offset on the reduced transverse galaxy bispectrum is only weakly scale dependent, thus with our choice for the radial selection funtcion the angular reduced bispectrum has very similar shape dependence.

The Limber approximation requires the transverse scales under consideration to be significantly smaller than the radial projection depth, hence we limit our analysis to angular scales corresponding to comoving Fourier modes $k \geq 0.04h\,\mathrm{Mpc}^{-1}$. As our intrinsic alignment toy models and biasing approximation are not designed to describe in the nonlinear regime of structure formation, we will only consider angular frequencies corresponding to

$$0.04\,h\,\mathrm{Mpc}^{-1} \leq k \leq 0.2\,h\,\mathrm{Mpc}^{-1}. \tag{4.49}$$

We approximate the galaxy power spectrum by the linear matter power spectrum rescaled by the linear bias (equation 4.7); bispectra and trispectra on these scales are approximated by the tree-level perturbation theory in combination with local biasing (equation 4.5), i.e., using equations (4.7), (4.6), and (B.4). These are evaluated using transfer functions generated by CMBFAST (Seljak and Zaldarriaga 1996) for the best-fit WMAP 7 cosmology (Komatsu et al. 2010). Compared to an approach combining the halo model with halo occupation distribution modeling (e.g., Berlind and Weinberg 2002, Cooray and Sheth 2002) this is computationally much faster, the only model input is our biasing prescription and does not require halo models for intrinsic alignment. In the large-scale limits the halo models asymptote to the perturbation theory result, and at the scales of our analysis the galaxy power spectrum is fairly well described by perturbation theory (Cooray 2004, Smith, Sheth and Scoccimarro 2008). At redshift $z = 0$, Smith, Sheth and Scoccimarro (2008) find the reduced halo model bispectrum with $k_2 = 2k_1$ to be in close agreement with perturbation theory results at scales $k_1 \leq 0.1\ h/\mathrm{Mpc}$, except for collinear configuration ($\theta_{12} \to 0$). As we only consider triangle configurations with all angular frequencies $k_{1,2,3} \leq 0.2\ h\,\mathrm{Mpc}^{-1}$, the perturbation theory results should be sufficient at the level of this analysis. However, at scales smaller than $k \sim 0.1h/\mathrm{Mpc}$ Smith, Sheth and Scoccimarro (2008) and Guo and Jing (2009a) find the bispectrum measured from simulations to differ at the 10-20% level from the perturbation theory. Note that these systematic effects on the determination of bias parameters on small scales are larger than the tidal alignment contaminations discussed here.

We model the observed power spectrum by averaging the angular power spectrum over bins of width Δl,

$$\mathcal{P}(\bar{l}) \equiv \int_{\bar{l}-1/2\Delta l}^{\bar{l}+1/2\Delta l} \frac{\mathrm{d}l\ l}{\bar{l}\Delta l}\mathcal{P}(l), \tag{4.50}$$

and the corresponding covariance is given by

$$\mathrm{Cov}\left(\mathcal{P}(\bar{l}_1)\mathcal{P}(\bar{l}_2)\right) = \frac{1}{\Omega}\left\{\delta_{\bar{l}_1,\bar{l}_2}\frac{4\pi}{\bar{l}_1\Delta l}\left[\mathcal{P}(\bar{l}_1)+\frac{1}{\bar{n}}\right]^2 + \int_1\int_2 \mathcal{T}(\mathbf{l}_1,-\mathbf{l}_1,\mathbf{l}_2,-\mathbf{l}_2)\right\}\ , \tag{4.51}$$

where $\bar{n}$ is the average projected density of the galaxy population under consideration. Here the first term is a combination of Gaussian cosmic variance and shot noise. The second term involving the trispectrum of parallelogram configurations is the non-Gaussian power spectrum covariance.

The bispectrum is sampled with uniform binning Δl in all angular frequencies. Defining

$$\int_i \equiv \int_{\bar{l}_i-1/2\Delta l}^{\bar{l}_i+1/2\Delta l}\frac{\mathrm{d}l_i\ l_i}{\bar{l}_i\Delta l}, \tag{4.52}$$

the bin-averaged bispectrum is given by

$$\mathcal{B}(\bar{l}_1,\bar{l}_2,\bar{l}_3) \equiv \int_1\int_2\int_3 \mathcal{B}(l_1,l_2,l_3)\delta_{\mathrm{D}}(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3). \tag{4.53}$$

We approximate the expression from Joachimi, Shi and Schneider (2009) for the full non-Gaussian covariance of the bin-averaged bispectrum by

$$\begin{aligned}\mathrm{Cov}\left(\mathcal{B}(\bar{l}_1,\bar{l}_2,\bar{l}_3)\mathcal{B}(\bar{l}_4,\bar{l}_5,\bar{l}_6)\right) \quad = \quad & \frac{(2\pi)^3}{\Omega\bar{l}_1\bar{l}_2\bar{l}_3\Delta l^3}\Lambda^{-1}(\bar{l}_1,\bar{l}_2,\bar{l}_3)D_{\bar{l}_1,\bar{l}_2,\bar{l}_3,\bar{l}_4,\bar{l}_5,\bar{l}_6}\\ & \times\left[\mathcal{P}(\bar{l}_1)+\frac{1}{\bar{n}}\right]\left[\mathcal{P}(\bar{l}_2)+\frac{1}{\bar{n}}\right]\left[\mathcal{P}(\bar{l}_3)+\frac{1}{\bar{n}}\right]\\ & +\frac{2\pi\Lambda^{-1}(\bar{l}_1,\bar{l}_2,\bar{l}_3)\Lambda^{-1}(\bar{l}_4,\bar{l}_5,\bar{l}_6)}{\Omega}\delta_{\bar{l}_3,\bar{l}_4}\int_1\int_2\int_3\int_5\int_6\delta_{\mathrm{D}}\left(\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3\right)\\ & \times\Big\{\delta_{\mathrm{D}}\left(\mathbf{l}_3+\mathbf{l}_5+\mathbf{l}_6\right)\mathcal{B}(l_1,l_2,l_3)\mathcal{B}(l_3,l_5,l_6)\\ & \quad +\delta_{\mathrm{D}}\left(-\mathbf{l}_3+\mathbf{l}_5+\mathbf{l}_6\right)\mathcal{T}(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_5,\mathbf{l}_6)\mathcal{P}(l_3)\Big\} + 8\text{ perm.},\end{aligned} \tag{4.54}$$

where the symmetry factor $D_{\bar{l}_1\ldots\bar{l}_6}$ is non-zero only for diagonal elements of the covariance ($\{\bar{l}_1,\bar{l}_2,\bar{l}_3\} = \{\bar{l}_4,\bar{l}_5,\bar{l}_6\}$): $D_{\bar{l}_1\ldots\bar{l}_6} = 1$, 2, or 6 for scalene, isosceles, or equilateral triangles respectively. If $\bar{l}_1,\bar{l}_2,\bar{l}_3$ form a triangle, then $\Lambda^{-1}(\bar{l}_1,\bar{l}_2,\bar{l}_3)$ is the area of this triangle, otherwise $\Lambda^{-1} = 0$. The first term is the Gaussian (diagonal) part of the covariance which is proportional to the product of three power spectra which have been modified to account for Gaussian shot noise. The second/ third terms are non-Gaussian contributions from triangle pairs which have at least one common side so that the pentaspectrum can be factorized into two bispectra/a trispectrum and a power spectrum. We have dropped a term which is proportional to the general connected pentaspectrum.

4.4.2 Biased Parameter Estimates for Galaxy Bias Parameters

Having set up a model for the observable data and their covariances, we can now quantify the power of our fictitious survey at constraining model parameters using the Fisher matrix

$$\mathcal{F}_{\alpha\beta} = \frac{\partial \vec{\mathcal{P}}^t}{\partial p_\alpha} \mathrm{Cov}^{-1}\left(\vec{\mathcal{P}}, \vec{\mathcal{P}}\right) \frac{\partial \vec{\mathcal{P}}}{\partial p_\beta} + \frac{\partial \vec{\mathcal{B}}^t}{\partial p_\alpha} \mathrm{Cov}^{-1}\left(\vec{\mathcal{B}}, \vec{\mathcal{B}}\right) \frac{\partial \vec{\mathcal{B}}}{\partial p_\beta} , \tag{4.55}$$

where the $\vec{\mathcal{P}}$ and $\vec{\mathcal{B}}$ are data vectors with the binned angular *galaxy* power spectrum and bispectrum as data points. The data vectors and their covariances depend explicitly on the bias parameters through equations (4.7, 4.6, B.4). Note that we do not include cross-correlations between power spectrum and bispectrum, both for simplicity and because they are small in the weakly nonlinear regime (but see Sefusatti et al. 2006 for their constraining power in the weakly nonlinear regime). The parameters of interest here are the linear and quadratic galaxy bias and we marginalize over the normalization of the matter power spectrum σ_8, i.e., $\mathbf{p} = (b_1, b_2, \sigma_8)$. Our fiducial model assumes $\sigma_8 = 0.8$, no intrinsic alignment contamination, and covers a range of bias parameters, while all other cosmological parameters are fixed to their best-fit WMAP 7 values.

The inverse Fisher matrix serves as a lower limit on the marginalized covariance of statistical parameter errors

$$\left\langle \delta p_\alpha \delta p_\beta \right\rangle = \left(\mathcal{F}^{-1}\right)_{\alpha\beta} . \tag{4.56}$$

Hence the statistical error on the inferred parameters is inversely proportional to $\sqrt{\Omega}$, as can be seen from the expressions (equations4.51, 4.54) for the data covariances. The presence of a systematic error $\Delta\vec{\mathcal{B}}$, $\Delta\vec{\mathcal{P}}$ in the data which is not included in the model induces a bias in the parameter estimate compared to its fiducial values. To first order it is given by (e.g., Huterer et al. 2006, Amara and Réfrégier 2008)

$$\Delta p_\alpha = \langle \hat{p}_\alpha \rangle - p_\alpha^{\mathrm{fid}} = \left(\mathcal{F}^{-1}\right)_{\alpha\beta} \left[\Delta\vec{\mathcal{P}}^t \, \mathrm{Cov}^{-1}\left(\vec{\mathcal{P}}, \vec{\mathcal{P}}\right) \frac{\partial \vec{\mathcal{P}}}{\partial p_\beta} + \Delta\vec{\mathcal{B}}^t \, \mathrm{Cov}^{-1}\left(\vec{\mathcal{B}}, \vec{\mathcal{B}}\right) \frac{\partial \vec{\mathcal{B}}}{\partial p_\beta} \right] , \tag{4.57}$$

where the data vectors and covariances are evaluated at the fiducial model.

This systematic bias is independent of the survey area, but it is influenced by our choice of survey parameters through the selection function (equation 4.47) and data binning scheme. It also depends on projected number density of the galaxy population of interest as $\bar{n}$ determines the impor-

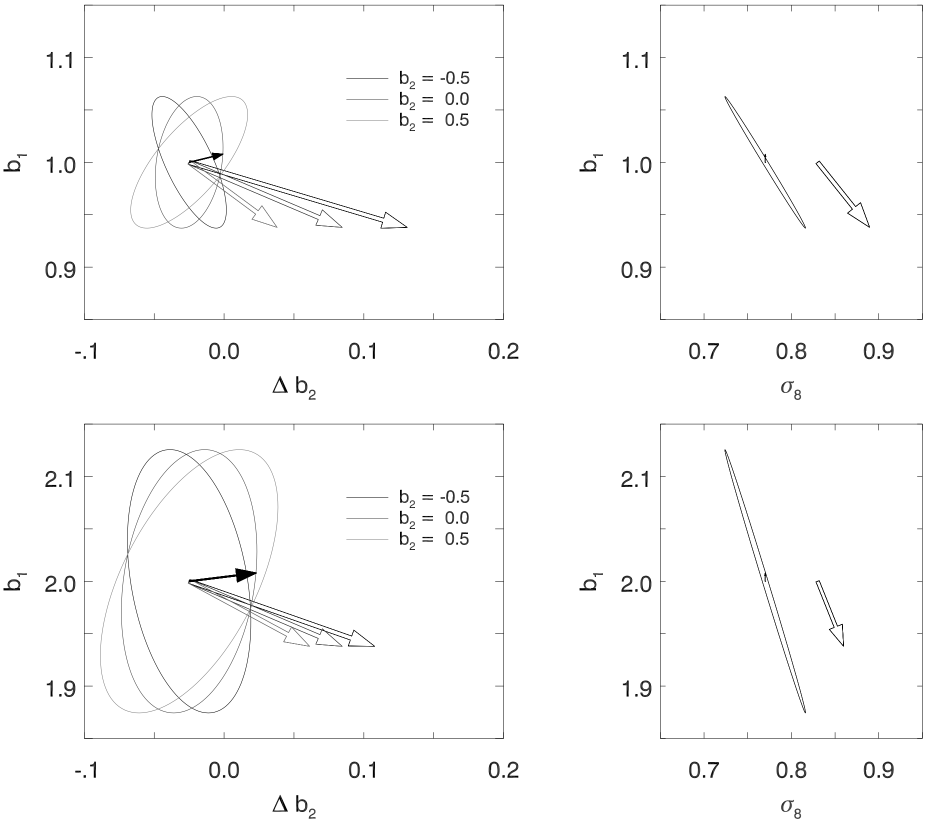


Figure 4.3 Systematic errors induced by intrinsic alignment. Ellipses show 95% C.L. statistical errors on parameter estimates in a DES-like surveys for a fiducial model with $\sigma_8 = 0.8$, for a galaxy population with $b_1 = 1$ (top panels) or $b_1 = 2$ (bottom panel) and $b_2 \in \{-0.5, 0, 0.5\}$. Open/ filled arrows illustrate the systematic parameter shift induced by a quadratic/ linear intrinsic alignment contamination.

tance of shot noise. We adopt a uniform sampling with 20 equidistant bins in all angular frequencies (l_1, l_2, l_3) corresponding to equation (4.49) and assume a projected density of $\bar{n} = 1/\text{arcmin}^2$ for a galaxy population in the redshift range $0.4 \leq z \leq 0.6$.

The systematic error on the bispectrum, $\Delta\vec{\mathcal{B}}$, due to linear or quadratic alignment is modeled by the line of sight projection (equation 4.48) of the tidal alignment contaminations (equations4.29, 4.39) calculated in section 4.3. We set $\Delta\vec{\mathcal{P}} = 0$ for the quadratic alignment model as the first correction to the power spectrum is third-order in the density contrast. In agreement with our findings from equation (4.29), the systematic error induced by linear alignment on the galaxy power spectrum is given by (cf. Hirata 2009)

$$\Delta P_{\mathrm{g}}^{\mathrm{LA}}(k_\perp) = \left[\left(b_1 - \frac{A_1}{3}\right)^2 - b_1^2\right] P_{\mathrm{g}}(k_\perp), \tag{4.58}$$

where we have restricted **k** to be orthogonal to the line of sight as only these modes survive the Limber approximation.

Figure 4.3 shows the marginalized Fisher matrix estimates of statistical parameter errors (95% C.L.) obtained with our fictitious survey in the absence of an intrinsic alignment, and the systematic bias induced by a linear or quadratic alignment contamination.

The systematic bias induced by a linear alignment contamination (solid arrows) we find through the Fisher matrix analysis (equation 4.57) is in agreement with the analytic result (equation 4.29). The parameter bias on b_1 is independent of the value of b_2 assumed in the fiducial model and the solid arrows of different color are indistinguishable. Assuming a normalization of $A_1 = -0.024$ as discussed in section 4.3, the systematic error on b_2 is comparable to the 95% C.L. statistical error for b_2 in our survey. The systematic error on b_1 caused by the linear alignment model is smaller, but may still be important if many photo-z slices are used in the parameter analysis. In the limit of our toy model, the effect of linear alignment on the angular galaxy power spectrum and bispectrum is fully described by a systematic error in the linear and nonlinear bias parameter (equation 4.29) and it has no effect on measurements of σ_8.

The strength of the quadratic alignment contamination depends on triangle shape and size; it is *not* well described by a rescaling of the galaxy bias parameters. Hence the Fisher matrix estimates for the systematic parameter errors depend on the binning scheme and range of scales adopted in the analysis. For our choice of 20 equidistant bins per angular frequency, and with the range of scales of 0.04–0.2$h\,\mathrm{Mpc}^{-1}$, we a systematic shift toward larger nonlinear bias b_2 and smaller b_1. The latter is degenerate between b_1 and σ_8. The plot illustrates a quadratic alignment contamination with normalization $A_2 = 1$. As can been seen from equations (4.39, 4.57), the systematic bias is linear in A_2, and it reverses sign if $A_2 < 0$. While exact form of the systematic error caused by the toy model for quadratic alignment depends on a number of parameters, it may cause a significant contamination in our fictitious survey if $|A_2| \gtrsim 0.5$, or if (as we expect) multiple photo-z slices are used to reduce statistical errors.

4.5 Discussion

Using simple toy models for intrinsic alignment and the local bias approximation we have analyzed the effect of tidal alignment on the galaxy bispectrum. If the orientation of galaxies depends on the surrounding tidal field, and if the detection probability for galaxies is orientation dependent, the observed clustering of galaxies is modified by tidal alignments. This astrophysical contaminant can introduce systematic errors to parameters derived from the bispectrum.

A toy model for linear alignments (Catelan, Kamionkowski and Blandford 2001), which is based on the assumption that tidal fields elongate/compress haloes and thus determine galaxy shapes, results in a rescaling of linear and nonlinear galaxy bias parameters that is proportional to the strength of the halo shape distortion. The presence of this systematic error in the observed galaxy bias measurements cannot be detected from projected clustering data as the strength of the alignment contamination is completely degenerate with the unobservable true bias parameters and outside information will be necessary to remove it. Normalizing the strength of the linear tidal alignment toy model to measurements of intrinsic alignments in weak lensing observations, we find that linear alignment may introduce systematic errors to galaxy bias measurements at the percent level (again using only the real-space observables), and thus will likely not be significant.

Using a simple model for quadratic alignment based on galaxy spin correlations in linear tidal torque theory we calculate a systematic contamination which modifies the shape of the galaxy bispectrum. Depending on survey characteristics, we find that quadratic alignment may introduce significant systematic errors to the galaxy bias parameters and the normalization of the power spectrum derived from the angular galaxy bispectrum. As the quadratic alignment contamination has different shape than the galaxy bispectrum, one can include a model for the contamination in the analysis and marginalize over its normalization. Figure 4.4 illustrates how such a marginalization may remove the systematic bias at the cost of larger statistical errors. The biased data points and contour levels (dashed lines) are taken from figure 4.3 for a fiducial model with $b_1 = 1$ and $b_2 = 0$. The new statistical errors including marginalization over A_2 are calculated by adding A_2 as a nuisance parameter and including the contamination signal in the fiducial model of the Fisher matrix analysis ($\vec{\mathcal{B}} \rightarrow \vec{\mathcal{B}} + \Delta\vec{\mathcal{B}}$ in equation (4.55)).

This analysis lives in the weakly nonlinear regime to enable the use of simple models for linear and quadratic alignment. As the information content of the bispectrum increases dramatically with the maximal spatial frequency that is included in an analysis, any realistic analysis will have extend well into the quasilinear regime. While models from the redshift space bispectrum on these scales (Smith, Sheth and Scoccimarro 2008) approach the required accuracy for such analyses, the treatment of tidal alignments including the non-Gaussian nature of the angular moment distribution and nonlinear stages of galaxy formation requires further work.

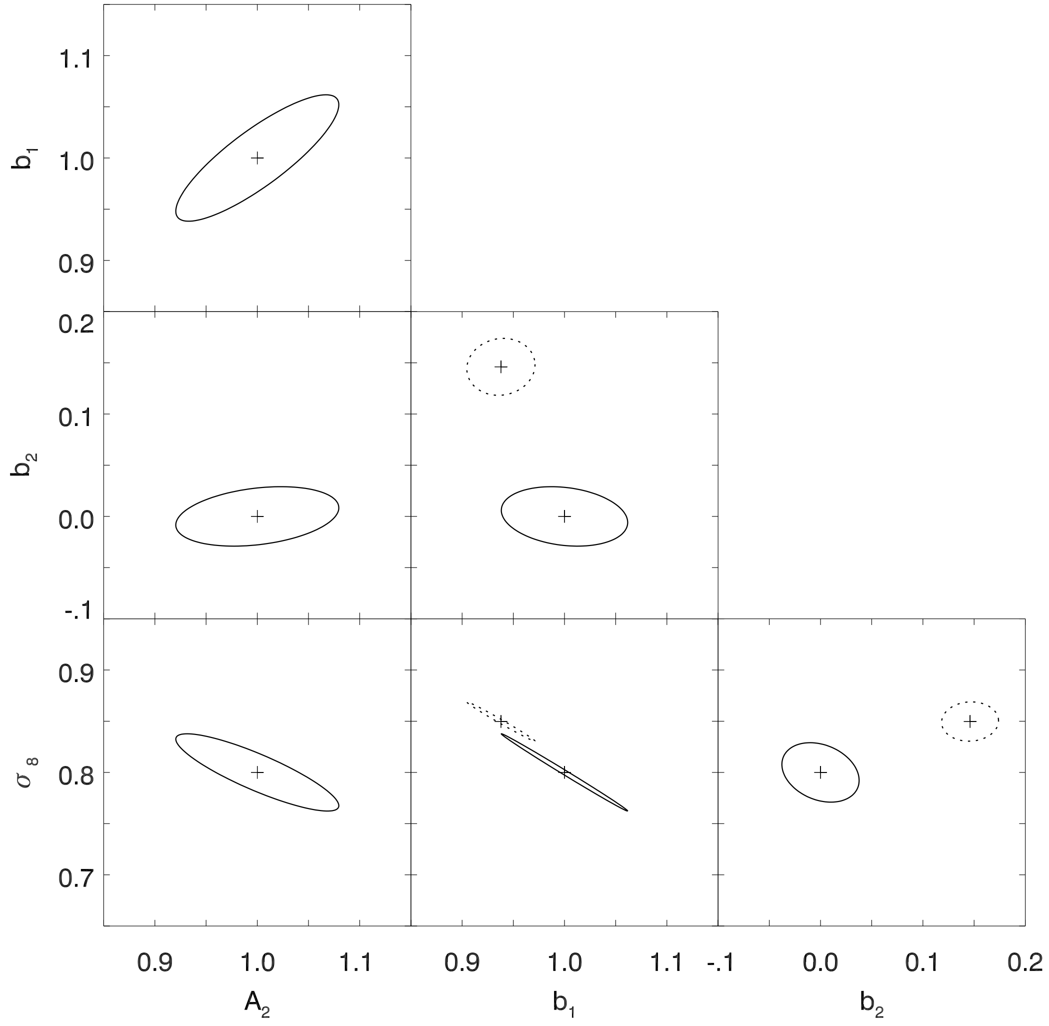


Figure 4.4 Removal of quadratic alignment bias. Dotted ellipses show the biased parameter estimates and their 95% contour regions in the presence of quadratic alignment contamination with $A_2 = 1$ which is unaccounted for in the analysis. The solid ellipses illustrate the 95% contour regions of the unbiased parameter estimates in an analysis which includes a quadratic alignment contamination and marginalizes over A_2.

Chapter 5

Clustering of Green Valley Galaxies

5.1 Introduction

Most nearby galaxies fall into one of two well-known and well-characterized categories. They are either passively evolving elliptical galaxies with old stellar populations, red in color and typically living in high-density regions, or they are actively star-forming spiral galaxies with blue color. The latter often are field galaxies or reside in other low-density regions like cluster outskirts.

This blue/red galaxy color bimodality has been observed to be in place already around $z \sim 1$. The fraction of red galaxies increases with time (e.g., Faber et al. 2007) and therefore galaxies must transition from blue to red. Galaxies in this transitional stage characteristically show low levels of recent star formation. As UV emission is a sensitive tracer of recent star formation, these transition galaxies are easily identified in a $(NUV - r)$–M_r color–magnitude diagram where they populate a "green valley" between well-localized red and blue sequences (Wyder et al. 2007).

The relation between galaxy color and environment density also evolves with redshift, such that the fraction of red galaxies increases with time in dense environments but stays nearly constant for field galaxies (e.g., Cooper et al. 2007, and references therein). This indicates the transition from blue to red galaxies may be driven by environmental processes, associated with the infall of a galaxy into a larger halo ("cluster"). Proposed mechanisms broadly fall into one of the following categories: galaxy–galaxy interactions (galaxy mergers and merger driven AGN activity, high speed galaxy interactions), galaxy–intra cluster medium interactions (e.g., ram pressure stripping or thermal evaporation), and interactions between an infaling galaxy and the cluster potential (e.g., truncation through tidal forces). Observationally these are disentangled through their characteristic timescales, the dependence of their respective efficiencies on halo mass, and position within the cluster (Treu et al. 2003, Cooper et al. 2006, Moran et al. 2007); for example, galaxy mergers are expected to be one of the dominant processes in group-scale halos and in the outskirts of massive clusters.

In the framework of ΛCDM cosmology, the evolution and spatial distribution of dark matter halos is relatively well understood. A common technique for inferring the masses of halos hosting different galaxy populations is to measure the angular or spatial clustering of galaxies and relate it to the predicted clustering and abundance of dark matter halos. While the relation between galaxy and dark matter clustering on large scales can be approximately described by scale-independent biasing, the situation is more complicated – and more informative about the physical processes at work – on small scales: At the level of individual halos, so-called halo-occupation distribution (HOD) models (e.g., Berlind and Weinberg 2002) describe the relation between galaxies and mass in terms of the probability that a halo of given mass contains N galaxies of a given type. Then galaxy clustering, for example the two-point correlation function, is modeled as the sum of contributions from galaxy pairs residing in the same halo and from galaxy pairs living in different halos.

This method of interpreting galaxy correlation functions has been used extensively: For example, Zehavi et al. (2010, see references therein for previous/high-z studies) analyze the completed (DR7) SDSS redshift survey, and find, in agreement with previous results, that at the amplitude of the correlation function increases with luminosity, and that at fixed luminosity redder galaxies are more strongly clustered, due to redder galaxies being satellites in more massive (and thus more biased) halos. Based on correlation function measurements over the redshift range $0.2 < z < 1.2$ from the Canada-France-Hawaii Telescope Legacy Survey, Coupon et al. (2011) also find red central galaxies to reside in more massive halos than average central galaxies in the same luminosity sample.

The clustering of $(NUV-r)$ color selected galaxies from the *Galaxy Evolution Explorer (GALEX)* survey has previously been studied by Heinis et al. (2007), who measure the angular correlation function; Heinis et al. (2009) and Loh et al. (2010) analyze spatial clustering as a function of star formation history and color respectively. These authors find the clustering of green galaxies to have intermediate strength compared to blue and red galaxies and to have a scale dependence closer to that of red galaxies. At small scales their analysis is strongly limited by statistics due to the small number density of green valley galaxies, limiting their ability to constrain the 1-halo term.

In this project we extend the HOD formalism to simultaneously model the cross-correlation functions (CCF) of a sparse luminosity bin galaxy sample with multiple more abundant galaxy populations to study the environment of local green valley galaxies. We consider luminosity bin samples of green valley galaxies as the physical mechanisms populating the green valley, i.e., quenching star formation in blue galaxies or rejuvenating red galaxies, may depend on halo mass and thus vary with galaxy luminosity. Compared to an autocorrelation function based clustering analysis, measuring the CCF between (sparse) *GALEX* selected galaxies and more abundant samples of SDSS galaxies reduces the shot noise contribution to our measurements, and also increases the effective volume probed beyond the combined *GALEX*-SDSS footprint. Extending previous work on HOD models for CCFs (e.g., Krumpe, Miyaji and Coil 2010) to simultaneously fit the clustering of the galaxy sample of interest with respect to multiple tracer populations is particularly helpful for analyzing the clustering of luminosity bin samples, which are harder to constrain than the more frequently used luminosity threshold samples.This allows us to put the separate piece of information found by Heinis et al. (2009) and Loh et al. (2010) into a coherent analysis including HOD modeling, and improve the statistics due to the larger survey area included in the newest data release.

Throughout this analysis we assume a flat ΛCDM cosmology with $\Omega_{\rm m} = 0.3$ and $\sigma_8 = 0.8$. Unless specified otherwise, all distances are coming and quoted in Mpc/h, and all absolute magnitude are given in $h = 1$ units.

5.2 Data

5.2.1 SDSS

Data release 7 (DR7 Abazajian et al. 2009) of the Sloan Digital SkySurvey (York et al. 2000, SDSS) spectroscopic sample provides (u'g'r'i'z')-photometry (Fukugita et al. 1996, Smith et al. 2002) and spectra for nearly 900000 galaxies with $m_{\rm r} < 17.77$ over 8000 square degrees. These galaxies were selected from the photometric survey for spectroscopic follow-up using specific algorithms for the main galaxy sample (Strauss et al. 2002) and luminous red galaxies (Eisenstein et al. 2001). The main spectroscopic galaxy sample is nearly complete to $r < 17.77$ and has a median redshift of $z \sim 0.1$. Based on these observations, the NYU Value Added Galaxy Catalog (VAGC Blanton et al. 2005) contains galaxy samples which have been constructed for large-scale structure studies: all magnitudes are re-calibrated and K-corrected (Blanton et al. 2003a), and the radial selection function and angular completeness are carefully determined from the data. We restrict this sample to $m_{\rm r} < 17.6$ to ensure uniform completeness of faint galaxies across the survey area.

Due to fiber placement in the SDSS spectrograph (Blanton et al. 2003b), galaxies closer than 55″ cannot be observed on the same spectroscopic plate, and hence no redshifts have been measured for about 7% of all targeted galaxies . The lack of observed close galaxy pairs affects the measured correlation functions on small scales. While it is possible to correct for fiber collisions down to $0.01\,\mathrm{Mpc}/h$ (Li et al. 2006), the number density of green valley galaxies is too small to obtain correlation function measurements at such small separations and we simply assign galaxies with missing spectra the redshift of its nearest neighbor. This method has been shown to work well for projected correlation functions above the scale corresponding to $55''$ (Zehavi et al. 2005). For the most distant galaxies in our sample the fiber collision scale is 0.07 comoving Mpc/h and we measure correlation functions only on perpendicular scales $r_\mathrm{p} \geq 0.1\,\mathrm{Mpc}/h$.

Spectral line measurements and mass estimates for these galaxies are taken from the MPA-JHU catalog.[1] We use the former to classify the $(NUV - r)$ selected transitional galaxies with emission line diagrams (figure 5.16) and to compare $(NUV - r)$ color selection with spectroscopic separation of active and quenched galaxies based on $D_\mathrm{n}4000$ (figure 5.7). Note that these quantities are estimated from a fiber size of $3''$, and due to low redshift of our galaxy sample these measurements may not be representative of the luminosity averaged properties of a galaxy but rather be dominated by central (bulge dominated) regions.

5.2.2 GALEX

NUV photometry for this project is taken from the GALEX Medium Imaging Survey Source Catalog (GMSC, Seibert et al. in prep.) derived from the GALEX GR6 data release, which provides unique measurements of point and extended sources up to 1 arcminute diameter in the GALEX bands (Seibert at al., in prep.). The NUV source catalog covers 4827 square degree at $\lambda_\mathrm{eff} = 2316$ with a resolution of $5.3''$ and reaching a depth ≈ 23 mag.

GALEX has a circular field of view of 1.2° which is sampled at $1.5''$. Each field targets a predefined position on the sky, resulting in a hexagonal tiling of the survey. These angular selection parameters are contained in exposure time, coverage and flag maps in `HEALpix` (Górski et al. 2005) format accompanying the GMSC, which we use to define the combined footprint and select our galaxy sample as detailed in section 5.2.3.

---

[1] http://www.strw.leidenuniv.nl/ jarle/SDSS/

5.2.3 SDSS-MIS Cross-Match

In order to match the VAGC with NUV detections, we first construct the combined footprint of these two surveys. This is done by converting the VAGC angular selection function, which is given in terms of `Mangle` polygons (Hamilton and Tegmark 2004), into the pixelized `HEALpix` format (Swanson et al. 2008). Then we multiply the angular selection functions of the VAGC and MIS in each pixel (at resolution $N_{\rm side} = 2048$) and restrict the overlap region to pixels where the angular completeness fraction of both surveys is larger than 0.7. This results in a combined survey with an effective area of 2708 square degrees. Furthermore, we require tiles to have NUV exposure times $t > 1000$ s, which reduces the combined effective area to 1945 square degrees. This final overlap region is shown in black in figure 5.1.

We cross-match all galaxies in the VAGC within this overlap area with NUV detections using a search radius of $4''$. In order to construct a complete statistical sample, we then restrict the cross-match with various cuts summarized in table 5.1. Due to deblending and centering issues for nearby or very bright objects, the NUV and r band photometry pipelines may report positions for these objects that are farther separated than the matching radius, leading to spurious non-detections. Furthermore, the astrometric and photometric precision of the GALEX detections declines toward the edges of each tile, and near light echos and other imaging artifacts and we exclude this regions as detailed in table 5.1. The color–apparent magnitude distribution and completeness of the final cross-match sample is shown in figure 5.2. For apparently bright galaxies ($m_{\rm r} \lesssim 16$) the blue sequence (around $(NUV - r) \approx 2 - 3$) and the red sequence (around $(NUV - r) \approx 5 - 6$) are clearly visible. No galaxies are found with $(NUV - r) \gtrsim 6.5$ though these should well be within the GALEX detection limit (indicated by the inclined line) at these magnitudes if they existed. For these bright galaxies far from the NUV detection limit the cross-match completeness is around 90%, it decreases for fainter objects as the NUV detection limit moves into the color-magnitude space occupied by red galaxies. In order to retain a nearly complete sample of green valley galaxies we cut the cross-match sample at $m_{\rm r} < 17.1$.

Finally, we use `kcorrectv4.2` (Blanton and Roweis 2007) to calculate absolute $\rm NUV_{0.1}$ magnitudes of the cross-match galaxies k-corrected to $z = 0.1$. As the redshift evolution in the NUV is not very well constrained, we do not attempt to apply evolution corrections to the NUV nor optical magnitudes. Similarly, we do not attempt to correct the $(NUV - r)$ colors for intrinsic extinction. To isolate transitional galaxies and avoid identifying dusty (edge on) spiral galaxies as green valley

Table 5.1. Cross-match sample definition

Parameter	Limits
r-band magnitude	$14.1 < r < 17.1$
redshift	$0.02 < z < 0.2$
GALEX field radius	$fov\_radius < 0^\circ.55$
GALEX exposure time	$t > 1000$ s
NUV flag	$nuv\_artifact \leq 1$
NUV magnitude	$16.0 < \mathrm{NUV} < 23.0$
SDSS/ NUV angular completeness	$f_{\mathrm{comp}} > 0.7$

Note. — The parent catalog is the NYU VAGC `dr72bright`.

objects, we only consider objects with r-band isophotal axis ratio $b/a > 0.5$.

5.3 Sample Definition

In order to work with well-defined galaxy populations, we construct a number of volume-limited samples. As the properties of green valley galaxies may vary with luminosity, we define samples of width 0.5 in absolute magnitude, and find the redshift range over which all galaxies in this sample have apparent magnitudes $14.1 < m_{\mathrm{r}} < 17.1$ (the magnitude range of the cross-matched catalog), c.f. figure 5.3. The VAGC has less stringent apparent magnitude requirements ($10 < m_{\mathrm{r}} < 17.6$), and we define two samples of SDSS galaxies occupying the same volume as each luminosity bin sample of NUV detected objects, which are used for the cross-correlation analysis. These samples are described in detail in table 5.2. Specifically, for the luminosity bin $[M_{\mathrm{r,min}}, M_{\mathrm{r,max}}]$ we define the "bright" sample of SDSS galaxies to contain all galaxies in the same redshift range brighter than $M_{\mathrm{r,max}}$, and the "faint" sample to consist of the volume-limited sample $[M_{\mathrm{r,min}} + 0.5, M_{\mathrm{r,max}}]$. We refer to the union of these two samples, which is a luminosity threshold sample with threshold $M_{\mathrm{r,min}} + 0.5$, as the SDSS "all" sample.

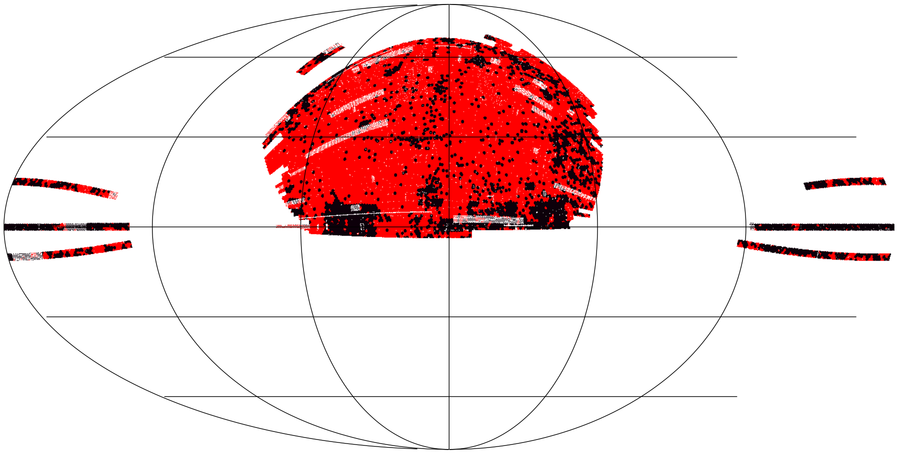


Figure 5.1 Combined SDSS + GALEX MIS footprint. The area covered by the VAGC at an angular completeness $f_{\mathrm{comp}} > 0.7$ is shown in red, the final overlap area of 1945 square degrees between VAGC and MIS, as detailed in section 5.2.3, is shown in black.

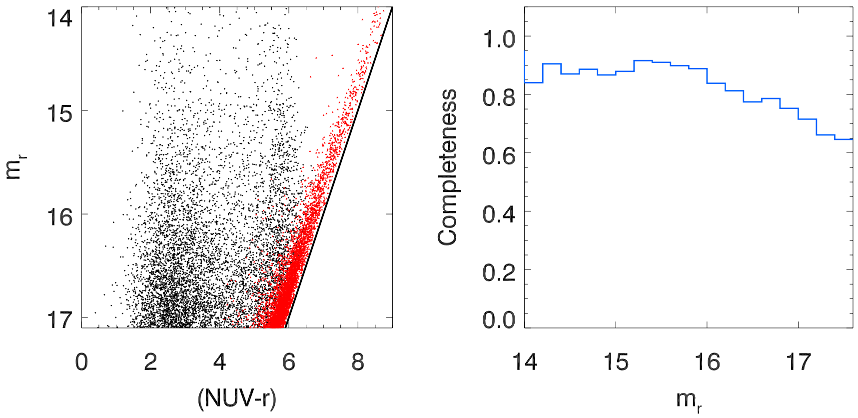


Figure 5.2 Completeness of the cross-match sample.*Left:* Apparent magnitude–$(NUV - r)$ color diagram. Black dots show a random subset of VAGC galaxies with NUV cross-match. Red dots indicate VAGC galaxies without NUV detections, which have been placed at the detection limit $NUV = 23$ and corrected for position dependent galactic extinction.
Right: Completeness of the NUV cross-match as a function of apparent magnitude.

Table 5.2. Volume-limited galaxy samples

Green Valley sample				SDSS samples	
M_r	$\langle z\rangle$	N_G	$\bar{n}_G$	N_f	N_b
$[-18, -18.5]$	0.031	285	1.09	15714	22177
$[-18.5, -19]$	0.036	595	1.19	24725	28488
$[-19, -19.5]$	0.044	869	0.92	38537	33041
$[-19.5, -20]$	0.055	1191	0.67	62193	37310
$[-20, -20.5]$	0.068	1746	0.54	95204	36561
$[-20.5, -21]$	0.083	2028	0.35	109490	23586
$[-21, -21.5]$	0.102	1383	0.13	112647	12073
$[-21.5, -22]$	0.128	775	0.04	87676	4458

Note. — The first two columns give the magnitude range $[M_{\rm r,min}, M_{\rm r,max}]$ and mean redshift of the green valley galaxy samples illustrated in figure 5.3. $N_{\rm G}$ is the number of green valley galaxies in this sample, and $\bar{n}_{\rm G}$ their mean comoving density per $10^{-3}\,({\rm Mpc}/h)^3$. $N_{\rm f}$ and $N_{\rm b}$ are the number of SDSS galaxies in the faint and bright sample in the same volume; the bright sample consists of galaxies in the same volume that are brighter than $M_{\rm r,max}$, and the faint sample contains galaxies in the magnitude range $[M_{\rm r,min} + 0.5, M_{\rm r,max}]$.

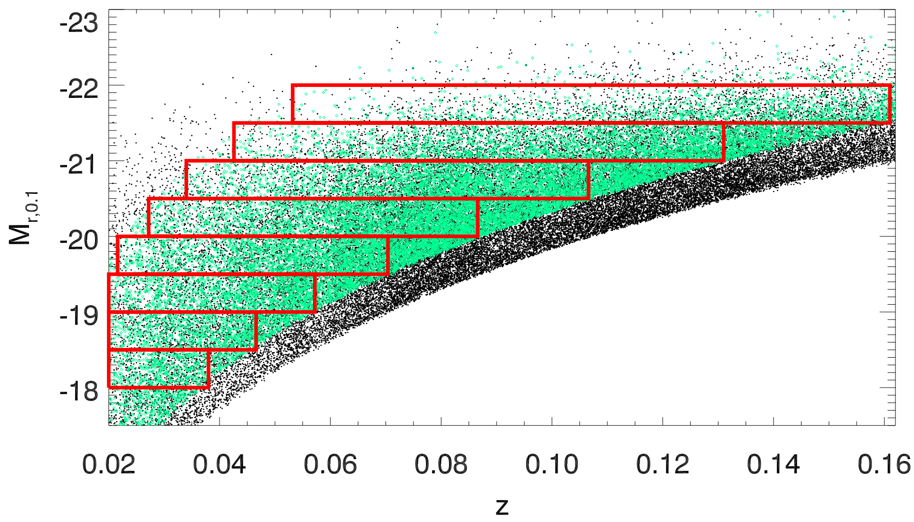


Figure 5.3 volume-limited galaxy samples. Black dots show a random subsample of VAGC galaxies with $m_{\mathrm{r}} < 17.6$, subsampled by a factor 10. Green symbols indicate green valley galaxies identified based on their $(NUV - r)$ color, which are restricted to $14.1 < m_{\mathrm{r}} < 17.1$ to ensure (near) completeness of the cross-matched sample.

5.3.1 Finding the Green Valley

We define the location of the green valley in $(NUV - r)$ color–magnitude space by fitting blue and red sequences to the color distribution of each volume-limited sample. We include galaxies without NUV detections, which otherwise meet all cross-match criteria and are optically red ($(g - r) > 0.8$), by placing them at the NUV detection threshold, correcting for position dependent galactic extinction and assigning the mean k-correction of cross-matched galaxies which are within $\Delta(NUV - r) = \pm 0.1$ mag, $\Delta M_r = \pm 0.1$ mag, and $\Delta z = \pm 0.02$ of the unmatched galaxy. We the find the center and scatter of the color sequences by fitting each sequence with a Gaussian. Initially, we cut the distribution at $(NUV - r) = 4.2$ and fit a Gaussian to each side. We then iteratively adjust the fitting range to include the galaxies within 1σ of the peak location on the ridge toward the Green valley. The best-fit parameters for each sample are shown in figure 5.5 along with fits to the blue and red sequence obtained by Wyder et al. (2007), which are based on a different fitting scheme and one continuous galaxy sample weighted by the v_{max} method instead of using disjunct volume-limited samples. As we include NUV non-detections, which are unaccounted for by Wyder et al. (2007), our red sequence is slightly redder for faint galaxies, but otherwise these results agree very well.

The black error bars in figure 5.4 illustrate the mean photometric uncertainty in the $(NUV - r)$

color of blue/red galaxies, suggesting that asymmetric scatter into the green valley due to photometric uncertainties is small compared to the intrinsic scatter of the red sequence.

5.3.2 Sample Properties

In order to facilitate the comparison with other studies of transitional galaxies based on optical criteria, we characterize the $(NUV - r)$ selected galaxies in other parameter spaces. Figure 5.6 and figure 5.7 show the distribution of $(NUV - r)$ selected galaxies in $(g - r)$ color space and as a function of the Balmer break index $\mathrm{D_n}4000$. Here the red sample again includes NUV non-detections as described in section 5.3.1. The vertical lines indicate the transition between blue/red and star forming/quenched galaxies based on $(g-r)$ and $\mathrm{D_n}4000$ respectively. Most faint $(NUV-r)$ selected green valley galaxies are optically blue and and would be classified as star forming by both of these criteria. On the other end, a large fraction of luminous, $(NUV - r)$ selected transitional galaxies would be classified as red/quenched by both of these criteria. Furthermore, figure 5.8 shows the distribution of stellar masses as a function of $(NUV - r)$ color. The stellar masses are taken from the MPA-JHU catalog and are based on Kauffmann et al. (2003). At fixed luminosity, green valley galaxies and red sequence galaxies have similar stellar masses.

5.4 Clustering Analysis

5.4.1 Projected Correlation Functions

To separate spatial clustering from redshift space distortions, we first measure the correlation functions in radial direction π and perpendicular direction r_p and then project out redshift space distortions. Specifically, we measure the (cross-)correlation function of galaxy samples $D_{X,Y}$ using the Landy and Szalay (1993) estimator

$$\xi_{XY}(r_\mathrm{p},\pi) = \left[\frac{D_X D_Y - D_X R_Y - D_Y R_X + R_X R_Y}{R_X R_Y}\right]\left(r_\mathrm{p},\pi\right), \tag{5.1}$$

on a two-dimensional grid. Here $R_{X,Y}$ are associated random catalogs, $DD\left(r_\mathrm{p},\pi\right)$, $DR\left(r_\mathrm{p},\pi\right)$ and $RR\left(r_\mathrm{p},\pi\right)$ are the (normalized) number of data-data, data-random, and random-random pairs at separation $\left(r_\mathrm{p},\pi\right)$. We adopt linear binning in the radial component, logarithmic bins in perpendicular

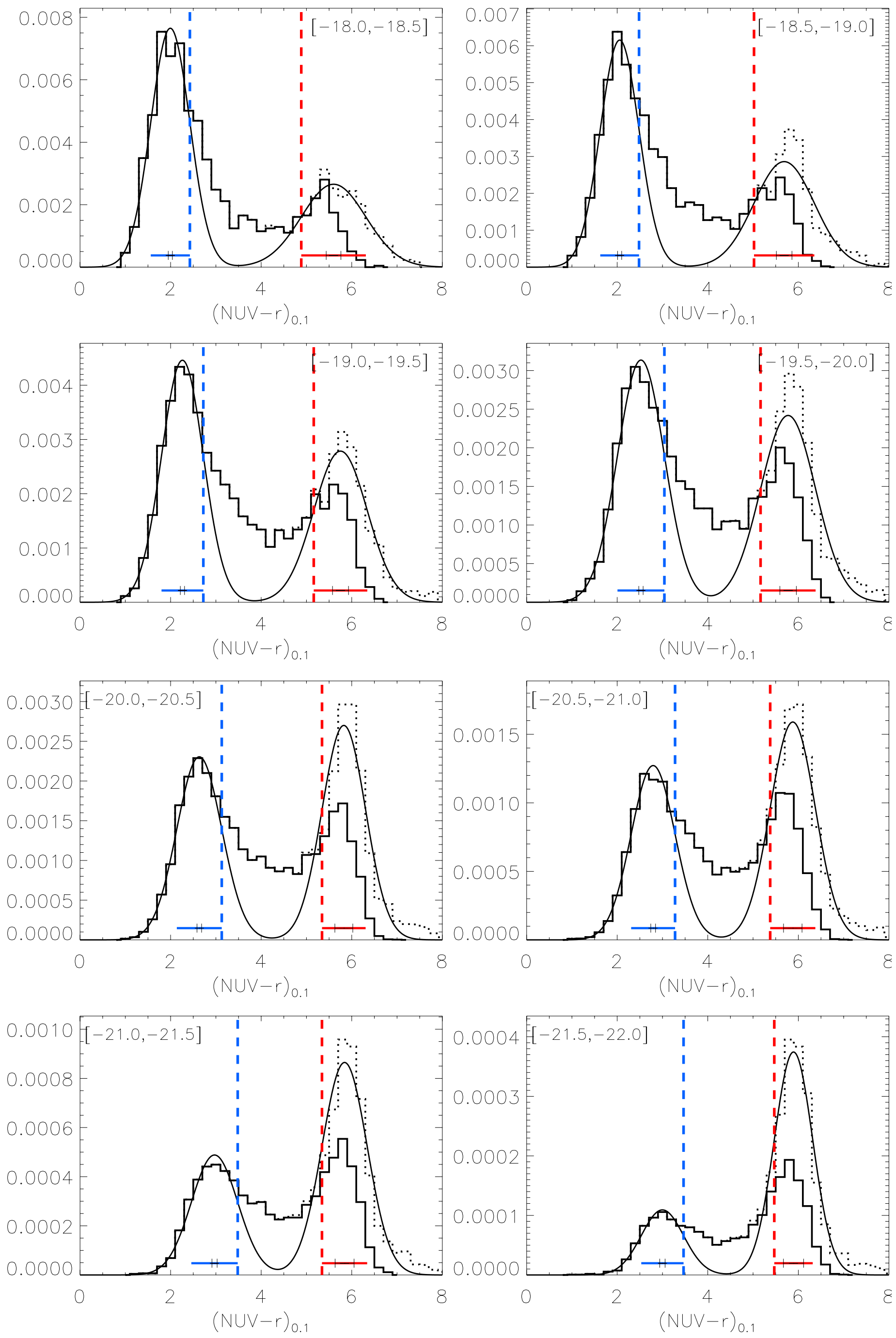


Figure 5.4 Comoving density of the volume-limited galaxy samples as a function of $(NUV - r)$ color. Solid histograms show all NUV detected galaxies. The dotted histograms include NUV non-detections, which otherwise meet all cross-match criteria and are optically red ($(g - r) > 0.8$), placed at the NUV detection threshold, corrected for position dependent galactic extinction and assigned the mean k-correction of cross matched galaxies which are within $\Delta(NUV - r) = \pm 0.1$ mag, $\Delta M_r = \pm 0.1$ mag, and $\Delta z = \pm 0.02$ of the unmatched galaxy. The solid line shows the double Gaussian fit to the blue side of the blue sequence and the red side of the red sequence, as described in 5.3.1, and the vertical blue and red lines show the 1σ ridge of the color sequences derived from these fits. The colored error bars also indicate the 1σ scatter of the color sequences centered on their respective peak. The black error bars illustrate the mean photometric uncertainty in the $(NUV - r)$ color of blue/red galaxies.

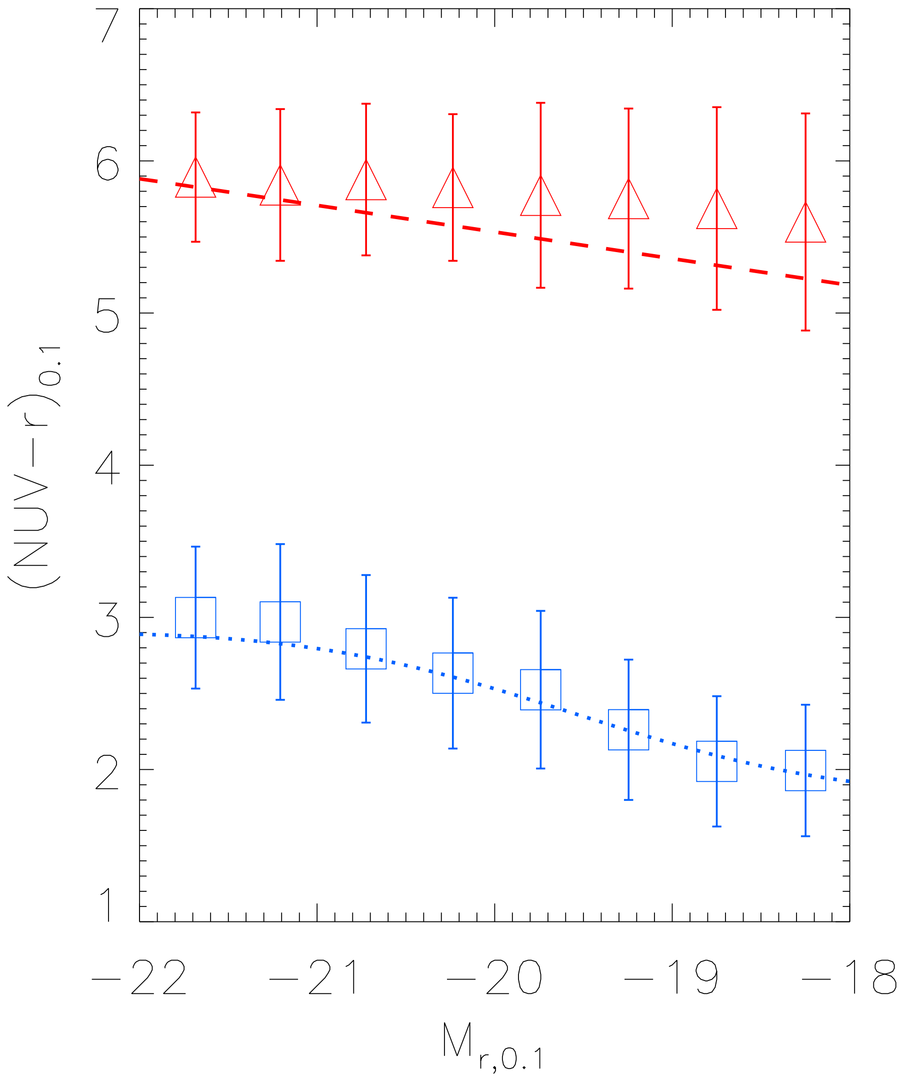


Figure 5.5 Defining the green valley: Symbols and error bars show the location and scatter of the blue and red sequence from the fits in figure 5.4. Lines show the best-fit sequences from Wyder et al. (2007) transformed to our magnitude units.

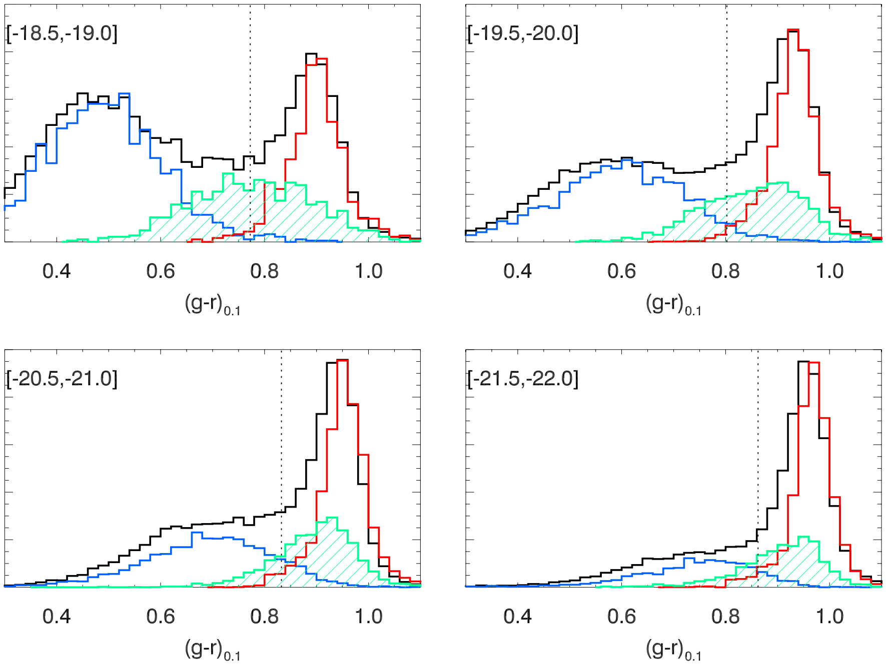


Figure 5.6 Colored histograms show the distribution of $(NUV - r)$ selected blue/green/red galaxies in $(g - r)$ space. The black histogram shows the distribution of all SDSS galaxies in the volume-limited sample, but not restricted to the combined footprint. The vertical line shows the color cut separating blue and red galaxies from Zehavi et al. (2010).

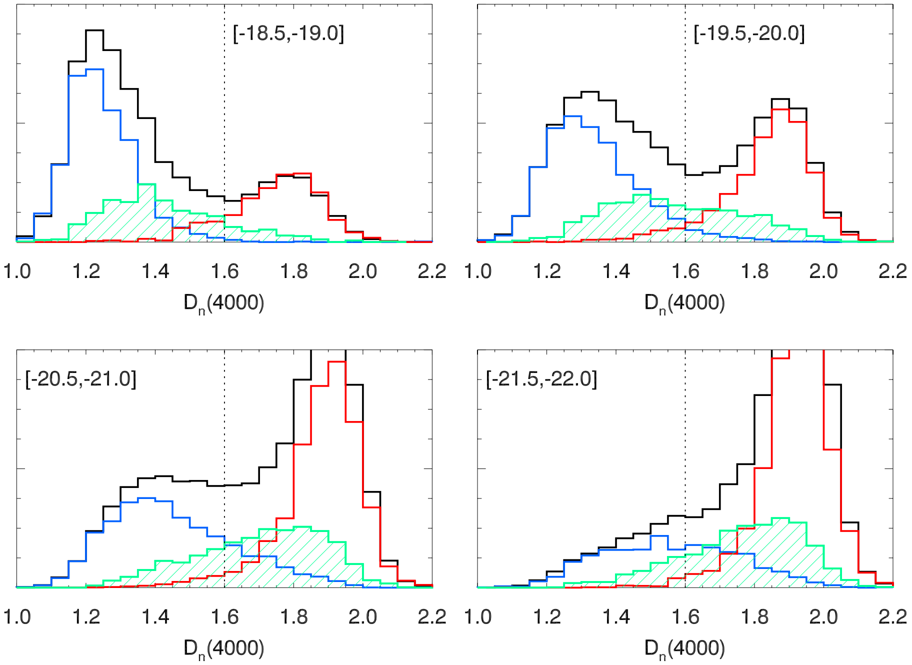


Figure 5.7 Same as figure 5.6 but for D_n4000. The vertical line shows the separation between quenched ($D_n4000 > 1.6$ and star forming galaxies used in Tinker, Wetzel and Conroy (2011).

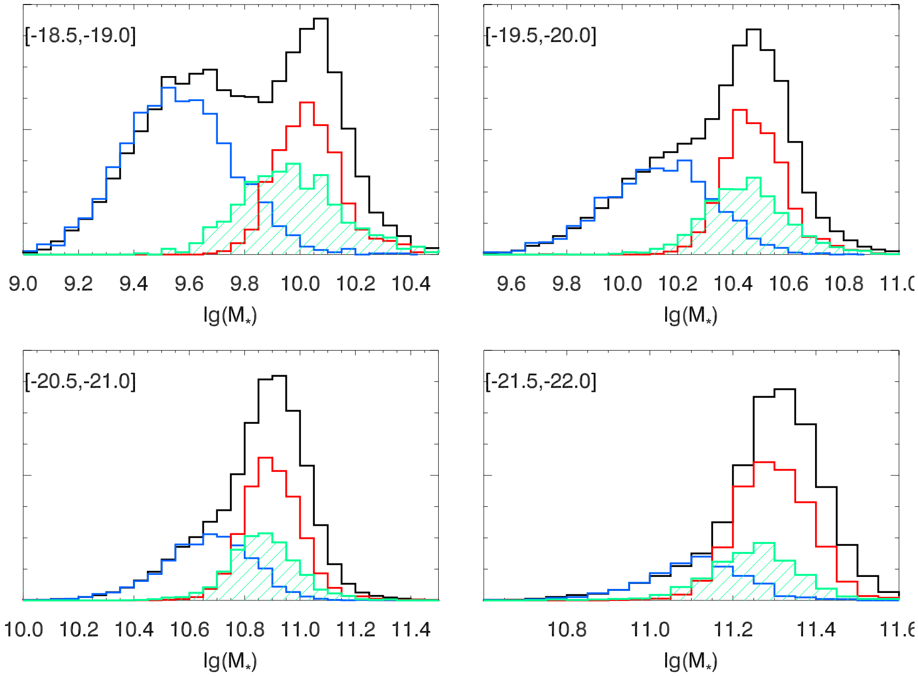


Figure 5.8 Same as figure 5.6 but for stellar mass.

distance and measure the projected (cross-)correlation function as

$$w_{XY}(r_{\mathrm{p}}) = 2\int_0^{\pi_{\max}} d\pi\, \xi_{XY}(r_{\mathrm{p}}, \pi)\,; \tag{5.2}$$

with $\pi_{\max} = 50\,\mathrm{Mpc}/h$.

5.4.2 Measurements

We generate random catalogs with the SDSS angular selection function and the angular selection function of the GALEX-SDSS cross-matchcatalog. As we have constructed volume-limited galaxy samples, the random catalogs have uniform comoving density and do not need to account for the radial selection function. We oversample the random catalogs compared to the galaxy catalogs by a factor 25 for SDSS samples, and by a factor 100 for the sparser $(NUV - r)$ selected samples.

Figure 5.9 demonstrates that we can have characterized the combined survey geometry sufficiently well to measure correlation functions in this patchy survey geometry. Here we show the correlation function between a galaxy sample in the full SDSS footprint in the magnitude bin $[-19.5, -20]$ and blue color $((g - r) < 0.8)$ with different subsets of itself: The dashed line shows its auto correlation function. Next we consider the cross-correlation between this sample and its restriction to the footprint of the SDSS + GALEX combined catalog, which is shown by the dotted line. Compared to the full auto correlation function, this cross-correlation function may be affected by boundary effects associated with the correlation function estimator or finite volume effects, as we have reduced the volume probed by of one copy of the galaxy catalog by a factor of four. Note that in this case the angular selection function in the combined survey area is still given by the SDSS angular selection function. Next we further restrict one copy of the galaxy catalog to galaxies with NUV detections, shown by the solid line. As the galaxy sample consists only of blue galaxies, these should all have NUV detections, and any significant differences between the dotted and solid line would indicate a mis-characterization of the combined angular selection function. One copy of the galaxy catalog stays the same throughout the process, so that we measure the cross-correlation between samples with different footprints, with leads to better statistics and smaller finite volume effects than restricting the SDSS data to the combined footprint region as well.

As described in detail in Zehavi et al. (2010), the clustering of the faintest SDSS luminosity threshold samples is subject to substantial sample variance effects due to the small volume probed by these low-redshift samples. As we are interested in a sparse subpopulation of these samples and are furthermore restricted to one fourth of the SDSS footprint area, these sampling effect are even more severe in our analysis. After reproducing their sub-volume tests, we find that the magnitude bin $[-19.5, -20]$ is the smallest sample for which we can obtain robust correlation function measurements. Examples of measured auto- and cross-correlation functions for SDSS galaxy samples and green valley galaxies are shown in figure 5.10.

We estimate the covariance of our correlation function measurements using bootstrapping with oversampling of subvolumes. Norberg et al. (2009) find that this method gives robust error estimates that are in agreement with external estimates from mock catalogs. For correlation functions between two SDSS galaxy samples, we divide the SDSS footprint into 150 subsets of equal area. For correlation functions between one SDSS galaxy sample and one sample restricted to the combined footprint area, the division into equal area subsets is not clearly defined, and we choose subsets which contain equal number of random-random pairs at angular separation of 2° in order to evenly sample the cross-correlation function on scales of a few Mpc$/h$. Due to the smaller effective area of this restricted geometry, we only have 50 such subareas. Examples for both types of covariances are shown in figure 5.11. As noted by Hartlap, Simon and Schneider (2007), estimated covariances are a biased estimate of the inverse covariance with the bias depending on the number of data points and the number of independent data sets. In order correct the inverse covariances for this bias, we find

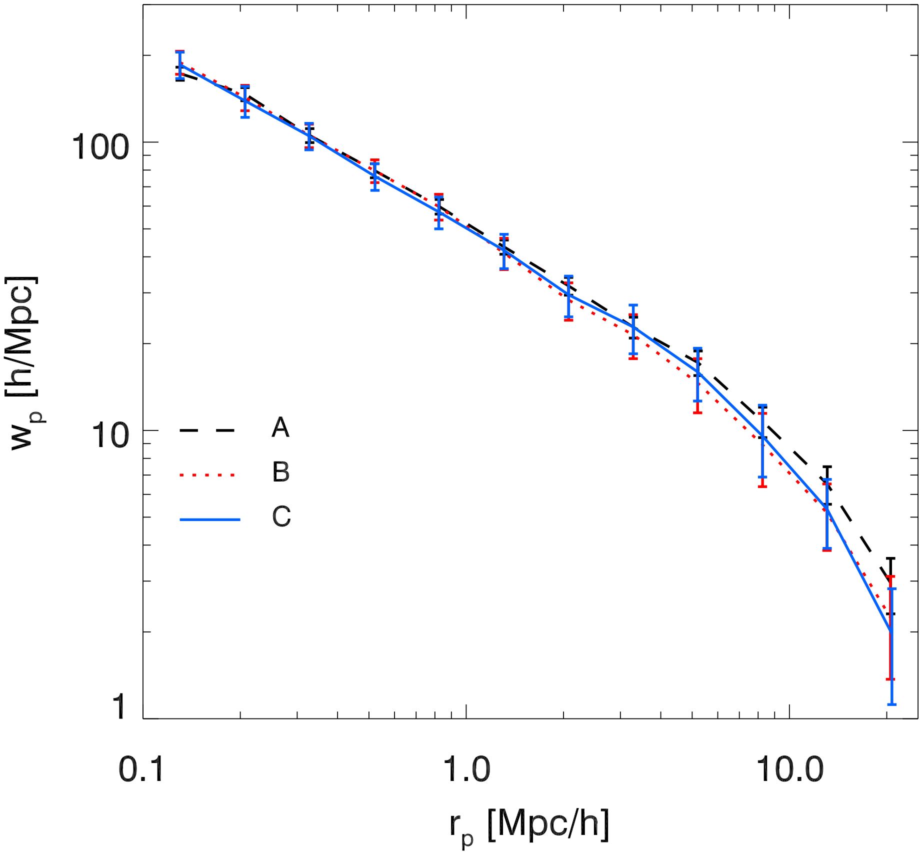


Figure 5.9 Test of survey geometry effects on measured correlation functions. Different lines show the projected cross-correlation function between galaxies in the full SDSS footprint in the magnitude bin [−19.5, −20] and with $(g - r) < 0.8$ (A) with he same sample, (B) with the sample restricted to the combined survey area, (C) with GALEX detected galaxies in the same magnitude and color bin.

the appropriate calibration factor by varying the number of bootstrap realizations as the bootstrap realizations are not independent.

We were unable to obtain stable, invertible covariances for the most luminous green valley galaxy sample. Hence we restrict our analysis of this sample to large scales (section 5.4.3) where it was possible to measure converged and invertible covariances.

5.4.3 Results: Large-Scale Bias

Based on the correlation function measurements described in the previous section, we can measure the large-scale galaxy bias by fitting the projected correlation functions with theoretical matter correlation functions times a linear bias factor. Specifically, we fit measured correlation functions over

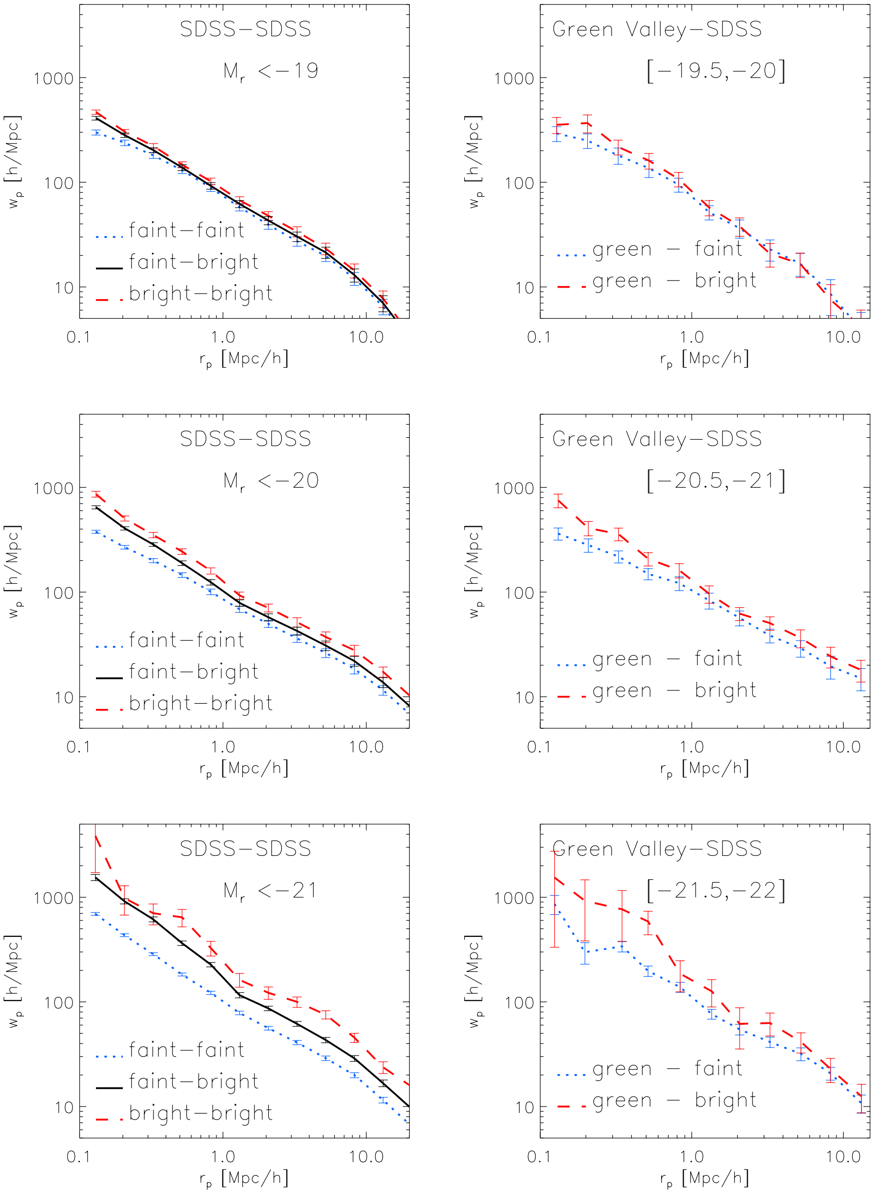
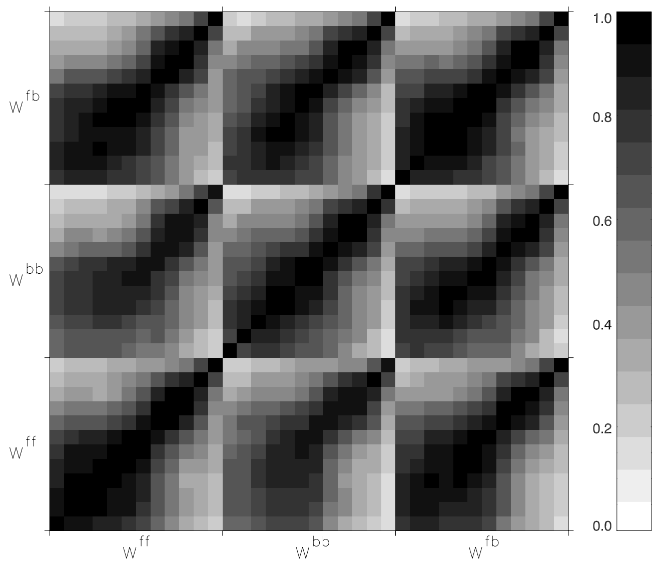


Figure 5.10 Examples of measured correlation functions.



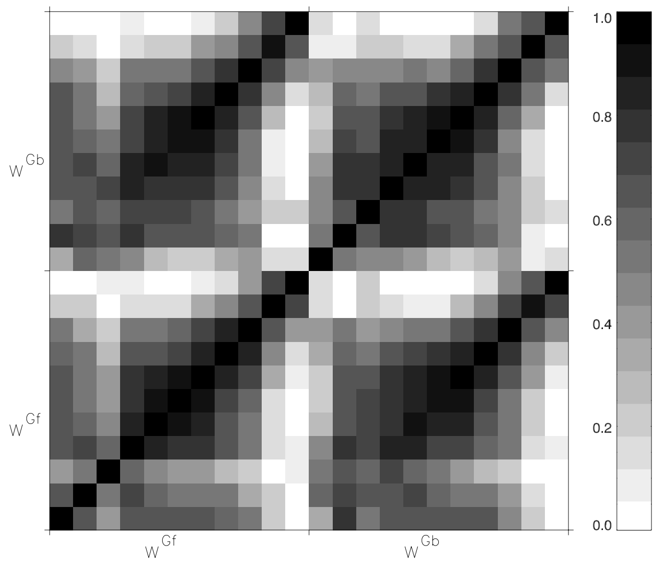


Figure 5.11 Sample covariances. Top: Covariance between the different auto- and cross-correlation functions of the SDSS faint and bright sample associated with the magnitude bin $[-19.5, -20]$. Bottom: Covariance of the cross-correlation function between the $[-19.5, -20]$ green valley sample and the corresponding SDSS faint and bright samples.

In each block of these covariances perpendicular scales increase from left to right and bottom to top.

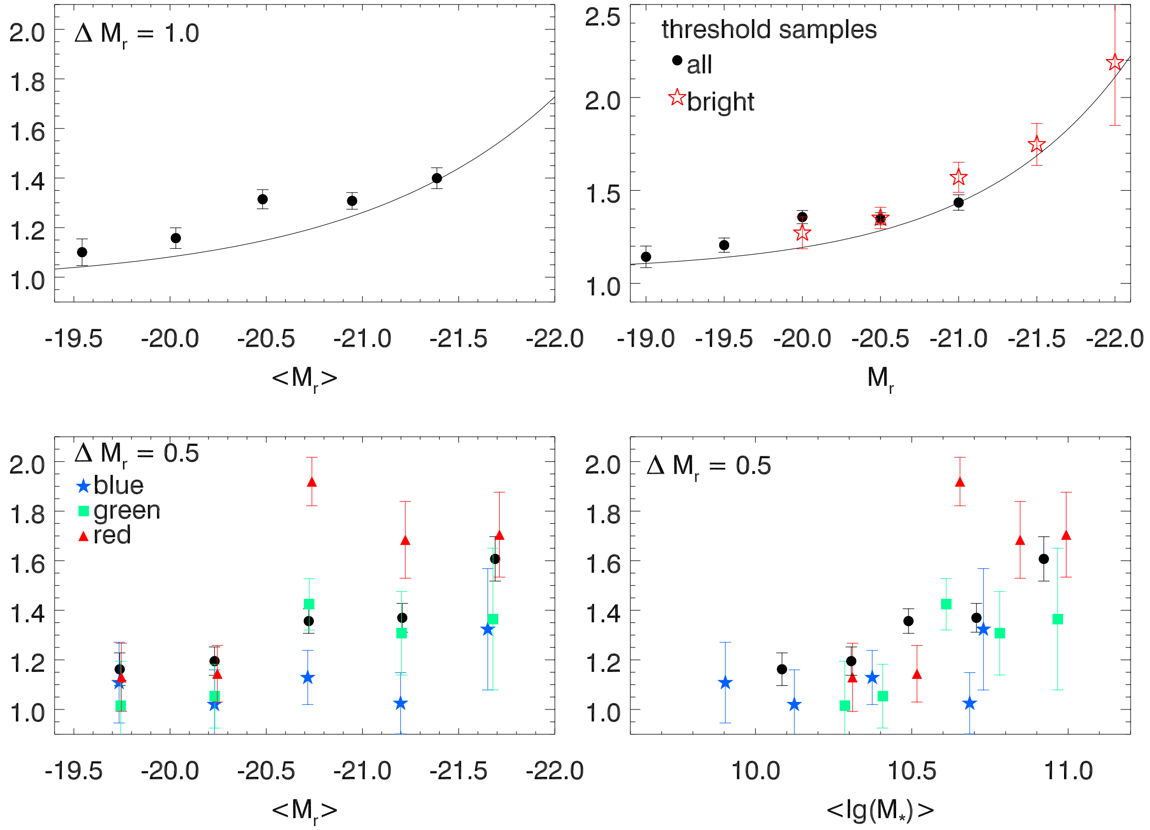


Figure 5.12 Linear galaxy bias measurements obtained from fits to the large-scale correlation function. Top: Linear bias as a function of luminosity of different luminosity bin (left) and threshold samples (right). The lines show best-fit relations from Zehavi et al. (2010).
Bottom: Linear bias as a function of $(NUV - r)$ color and luminosity (left) or stellar mass (right).

the range $3 - 25$ Mpc$/h$ to the theoretical predictions for the projected matter correlation function, including the full data covariance. Figure 5.12 shows the resulting luminosity bias relation. The top two plots are for binned and threshold samples of SDSS galaxies, and the lines are fits from the analysis of galaxy clustering in SDSS DR7 by Zehavi et al. (2010). Overall, we find good agreement with their results. The $M_{\rm r} < -20$ galaxy threshold sample and it subsamples deviate from the best-fit bias relation. As detailed in Tab. 5.2, these samples are centered around the redshift of the Sloan Great Wall, which leads to excess clustering in this and neighboring samples.[2] This effect is enhanced in the lower plots, which show bias as a function of $(NUV - r)$ color and luminosity or mean stellar mass. Here the clustering of red galaxies is strongly enhanced in the Sloan Great Wall.

---

[2] This was also noted by Zehavi et al. (2010) who exclude the redshift range of the Sloan Great Wall from their analysis of luminosity bin galaxy samples

5.5 Halo-Occupation Distribution Modeling

At the level of individual halos, a halo-occupation distribution (HOD) model (e.g., Berlind and Weinberg 2002) describes the relation between galaxies and halo mass in terms of the probability $P(N, M_{\rm h})$ that a halo of given mass $M_{\rm h}$ contains N galaxies. To describe the two-point clustering of galaxies, we need models for first and second moment of the HOD, $\langle N|M_{\rm h}\rangle$ and $\langle N(N-1)|M_{\rm h}\rangle$. Following Zheng et al. (2005), we separate galaxies into central and satellite galaxies. By definition, a halo contains either zero or one central galaxy, and it may host satellite galaxy only if it contains a central galaxy, which motivates the form

$$\langle N(M_{\rm h})\rangle = \langle N_{\rm c}|M_{\rm h}\rangle \left(1 + \langle N_{\rm s}|M_{\rm h}\rangle\right), \tag{5.3}$$

with $\langle N_{\rm c/s}|M_{\rm h}\rangle$ the average number of central/satellite galaxies in a halo of mass $M_{\rm h}$.

5.5.1 HOD Parameterization

While the assumptions in a HOD model describing the properties of dark matter halos are generally agreed upon (see section 5.5.2 for details), the form of the relation between galaxies and halos (equation (5.4)) is less well constrained and leaves more room for experiments. We motivate the details our implementation next.

5.5.1.1 SDSS Samples

We base our model for SDSS galaxy samples on the HOD parameterization of Zehavi et al. (2010) for luminosity thresholds samples with absolute r-band magnitude $M_r < M_r^t$,

$$\left\langle N(M_{\rm h}|M_r^t)\right\rangle = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log M_{\rm h} - \log M^{\rm t}_{\rm min}}{\sigma^{\rm t}_{\log M}}\right)\right]\left[1 + \left(\frac{M_{\rm h} - M^{\rm t}_0}{M'^{\rm t}_1}\right)^{\alpha_t}\right], \tag{5.4}$$

with model parameters $M^{\rm t}_{\rm min}, M^{\rm t}_0, M'^{\rm t}_1, \sigma^{\rm t}_{\log M}, \alpha_t$. The central galaxy occupation function is a softened step function with transition mass scale $M^{\rm t}_{\rm min}$, which is the halo mass in which the median central galaxy luminosity corresponds to the luminosity threshold, and softening parameter $\sigma^{\rm t}_{\log M}$ which is related to the scatter between galaxy luminosity and halo mass. The normalization of the satellite occupation function, $M'^{\rm t}_1$, and cut-off scale $M^{\rm t}_0$ are related to M_1, the mass scale at which a halo hosts at least on satellite galaxy ($N_{\rm s}(M_1) = 1$)); finally α_t is the high-mass slope of the satellite occupation function. This parametrization was found to reproduce the clustering of SDSS and CFHTLS galaxies (Coupon et al. 2012) well over a large range of luminosity thresholds and redshifts.

The HOD model for a binned galaxy sample with $M_r^{t_2} < M_r < M_r^{t_1}$ is typically calculated from model fits to luminosity threshold samples as

$$\left\langle N(M_{\rm h}|M_r^{\rm t_1}, M_r^{\rm t_2})\right\rangle = \left\langle N(M_{\rm h}|M_r^{\rm t_1})\right\rangle - \left\langle N(M_{\rm h}|M_r^{\rm t_2})\right\rangle \,. \tag{5.5}$$

While we note that the results of Zehavi et al. (2010) favor a somewhat steeper slope of the satellite distribution for the most luminous galaxy samples in our analysis, we set $\alpha = 1$ for all SDSS galaxy samples. This is in overall agreement with previous results for the luminosity range of interest, and makes differencing the HOD of neighboring samples numerically more stable. Hence our model has 4 free parameters for a luminosity threshold sample, and 8 free parameters for a luminosity bin sample. Without further constraints, such a parameterization of luminosity bin samples has too many degrees of freedom for general applications. However, it has the advantage that the HODs of neighboring luminosity bins are consistent with each other, and we use this parameterization to fit the different correlation functions among our SDSS faint and bright samples, resulting in 8 parameters for the SDSS HODs in each volume-limited sample.

5.5.1.2 Luminosity and Color bin Samples

For a $(NUV - r)$ selected galaxy sample (X), which is measured in one narrow 0.5 mag bin per sample volume, we need a more compact description of the HOD and we model the central galaxy term as a clipped Gaussian,

$$\langle N_{\rm c}(M_{\rm h}, X)\rangle = \min(\frac{A_X}{\sigma_X}\sqrt{2\pi}\exp\left(\frac{-(\log M_{\rm h} - \log M_{\rm c}^X)^2}{2\sigma_X^2}\right), 1)\,, \tag{5.6}$$

with free parameters A_X, σ_X and $M_{\rm c}^X$. Here the clipping enforces that a halo does not have more than one central galaxy.

For simplicity we assume the satellite occupation function to have linear slope, and the color of satellite galaxies to be independent of the color of their central galaxy. The latter assumption allows us to write the condition that halo has to contain a central galaxy in order to host satellite galaxies in terms of central galaxy occupation function of the full (color independent) luminosity threshold sample with luminosity threshold t_X equal to the minimum luminosity of the luminosity bin under consideration

$$\langle N_{\mathrm{s}}(M_{\mathrm{h}}, X)\rangle = A_X \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log M_{\mathrm{h}} - \log M_{\mathrm{min}}^{\mathrm{t}_X}}{\sigma_{\log M}^{\mathrm{t}_X}}\right)\right]\left(\frac{M_{\mathrm{h}}}{M_1^X}\right) \tag{5.7}$$

which is characterized by one free parameter, M_1^X.

Note that the correlation function of a binned sample is independent of the normalization parameter A_X, which is determined by the galaxy number density.

5.5.2 Relation to Correlation Functions

The halo model prediction for the real-space correlation function takes the form

$$1 + \xi(r) = \left(1 + \xi^{\mathrm{1h}}(r)\right) + \left(1 + \xi^{\mathrm{2h}}(r)\right), \tag{5.8}$$

where $(1 + \xi^{\mathrm{1h}})$ is proportional to the number of galaxy pairs residing in the same halo (*one-halo term*), and the *two-halo term* $(1 + \xi^{\mathrm{2h}})$ is proportional to the number of galaxy pairs occupying different halos. The model real-space correlation function is related to the projected correlation function as

$$w(r_{\mathrm{p}}) = 2\int_0^{\pi_{\max}} d\pi\, \xi\left(\sqrt{r_{\mathrm{p}}^2 + \pi^2}\right). \tag{5.9}$$

We will now describe the computation of these terms in detail. In order to evaluate these expressions numerically, we define halos to enclose a spherical overdensity of 200 times the mean background density and assume that their density distribution follows a NWF profile (Navarro, Frenk and White 1997) with the halo mass–concentration relation of Bhattacharya, Habib and Heitmann (2011); furthermore we use the fitting functions of Tinker et al. (2008) and Tinker et al. (2010) for the halo mass function and halo bias relation. Unless stated otherwise, we assume that the galaxy distribution follows the halo density profile.

5.5.2.1 One-Halo Term

We split the computation of the one-halo term into then clustering of central and satellite galaxy $\xi^{1,c-s}$ and satellite-satellite clustering $\xi^{1,s-s}$ within the same halo. The central-satellite term is given by

$$1+\xi_{XY}^{1,c-s}(r) = \frac{1}{\bar{n}_X\bar{n}_Y}\int_{M_{\mathrm{vir}(r)}}^{\infty} dM_{\mathrm{h}}\,\frac{dn}{dM_{\mathrm{h}}}\Big(\langle N_c(M_{\mathrm{h}},X)N_s(M_{\mathrm{h}},Y)\rangle\,\rho_Y(r|M_{\mathrm{h}})+\langle N_c(M_{\mathrm{h},}Y)N_s(M_{\mathrm{h}},X)\rangle\,\rho_X(r|M_{\mathrm{h}})\Big), \tag{5.10}$$

where dn/dM_{h} denotes the halo mass function, with $\rho_X(r|M_{\mathrm{h}})$ the normalized radial distribution of galaxy population X within the halo, and with

$$\bar{n}_X = \int_0^{\infty} dM_{\mathrm{h}}\,\frac{dn}{dM_{\mathrm{h}}}\,\langle N(M_{\mathrm{h}}|X)\rangle\ . \tag{5.11}$$

The term $\langle N_c(M_{\mathrm{h}},X)N_s(M_{\mathrm{h}},Y)\rangle$ in equation (5.10) is equal to the average number galaxy pairs with a central galaxy from sample X and a satellite galaxy from sample Y in a halo of mass M_{h}. From the definition of satellite galaxy this term evaluates to $\langle N_c(M_{\mathrm{h}},M_r^{\mathrm{t}})N_s(M_{\mathrm{h}},M_r^{\mathrm{t}})\rangle = \langle N_s|M_{\mathrm{h}},M_r^{\mathrm{t}})\rangle$ for the auto correlation of a luminosity threshold sample (Zheng et al. 2005). However, when considering binned samples or cross-correlations between different samples, the central galaxy of a halo hosting satellite galaxies from the sample Y need not be from sample X, and we use $\langle N_c(M_{\mathrm{h}},X)N_s(M_{\mathrm{h}},Y)\rangle = \langle N_c|M_{\mathrm{h}},X\rangle\,\langle N_s|M_{\mathrm{h}},Y\rangle$ Miyaji et al. (2011).

If samples X and Y are disjunct, the satellite-satellite term is given by

$$1+\xi_{XY}^{1,s-s}(r) = \frac{1}{\bar{n}_X\bar{n}_Y}\int_{M_{\mathrm{vir}(r)}}^{\infty} dM_{\mathrm{h}}\frac{dn}{dM_{\mathrm{h}}}\,\langle N_s(M_{\mathrm{h}},X)N_s(M_{\mathrm{h}},Y)\rangle\,(\rho_X*\rho_Y)\,(r|M_{\mathrm{h}})\,, \tag{5.12}$$

where $(\rho_X*\rho_Y)\,(r|M_{\mathrm{h}})$ denotes the convolution of radial galaxy distributions ρ_X and ρ_Y, and where the average number of satellite pairs is given by $\langle N_s(M_{\mathrm{h}},X)N_s(M_{\mathrm{h}},Y)\rangle = \langle N_{\mathrm{s}}|M_{\mathrm{h}},X\rangle\,\langle N_{\mathrm{s}}|M_{\mathrm{h}},Y\rangle$.

To model auto correlations function, the number of galaxy pairs is modified to

$$1+\xi_{XX}^{1,s-s}(r) = \frac{2}{\bar{n}_X\bar{n}_X}\int_{M_{\mathrm{vir}(r)}}^{\infty} dM_{\mathrm{h}}\frac{dn}{dM_{\mathrm{h}}}\frac{\langle N_s(M_{\mathrm{h}},X)(N_s(M_{\mathrm{h}},X)-1)\rangle}{2}\,(\rho_X*\rho_X)\,(r|M_{\mathrm{h}})\,. \tag{5.13}$$

Assuming that satellite galaxies are Poisson distributed, the number of pairs evaluates to $\langle N_{\mathrm{s}}(N_{\mathrm{s}}-1)\rangle = \langle N_{\mathrm{s}}\rangle^2$.

5.5.2.2 Two-Halo Term

On scales above $\sim 5\mathrm{Mpc}/h$, the clustering of galaxies follows the large-scale clustering of dark matter halos, and it is modeled as function of the dark matter correlation function ξ^{mm},

$$\xi^{\mathrm{2h}}_{XY}(r) \approx b_X b_Y \xi^{\mathrm{mm}}(r)\,. \tag{5.14}$$

Here b_X denotes the bias parameter of galaxy sample X, which we calculate as

$$b_X = \frac{1}{\bar{n}_X}\int_0^\infty dM_{\mathrm{h}}\,\frac{dn}{dM_{\mathrm{h}}} b_{\mathrm{h}}(M_{\mathrm{h}})\,\langle N(M_{\mathrm{h}}|X)\rangle\ , \tag{5.15}$$

where b_{h} is the halo bias parameter.

On intermediate scales one needs to account for the distribution of galaxies within different halos and halo exclusion, i.e., the fact that two halos contribution to the two-halo term do not overlap. Following the spherical halo exclusion model of Tinker et al. (2005), we restrict the calculation of the two-halo term at separation r to halos with $R_{\mathrm{vir},1} + R_{\mathrm{vir},2} \leq r$. The effect of the distribution of galaxies within the different halos on the correlation function is given by the convolution of their respective density profiles. As this requires convolving many different halo profiles, we calculate the two-halo term is calculated in Fourier space:

$$\begin{aligned} P^{\mathrm{2h}}_{XY}(k,r) \quad &= \quad P_m(k)\frac{1}{\bar{n}'_X\bar{n}'_Y(r)}\int_{M_{\mathrm{min}}}^{M_{\mathrm{lim},1}(r)} dM_1\,\frac{dn}{dM_1}\,\langle N|M_1,X\rangle\, b_{\mathrm{h}}(M_1)\tilde{\rho}_X(k,M_1) \\ &\quad \times\int_{M_{\mathrm{min}}}^{M_{\mathrm{lim},2}(M_1,r)} dM_2\,\frac{dn}{dM_2}\,\langle N|M_2,Y\rangle\, b_{\mathrm{h}}(M_2)\tilde{\rho}^*_Y(k,M_2)\,, \end{aligned} \tag{5.16}$$

where $M_{\mathrm{lim},1}$ is the maximum halo mass such that $R_{\mathrm{vir}}(M_{\mathrm{lim},1}) = r - R_{\mathrm{vir}}(M_{\mathrm{min}})$ with M_{min} the minimum halo mass of the HOD, where $M_{\mathrm{lim},2}$ is defined by $R_{\mathrm{vir}}(M_{\mathrm{lim},2}) = r - R_{\mathrm{vir}}(M_{\mathrm{lim},1})$, and where $\tilde{\rho}_X$ denotes the Fourier transform of the normalized galaxy distribution ρ_X. $\bar{n}'_X\bar{n}'_Y(r)$ denotes the number density of galaxy pairs restricted to non-overlapping halos at separation r

$$\bar{n}'_X\bar{n}'_Y(r) = \int_{M_{\mathrm{min}}}^{M_{\mathrm{lim},1}(r)} dM_1\,\frac{dn}{dM_1}\,\langle N|M_1,X\rangle \int_{M_{\mathrm{min}}}^{M_{\mathrm{lim},2}(M_1,r)} dM_2\,\frac{dn}{dM_2}\,\langle N|M_2,Y\rangle\ . \tag{5.17}$$

The two-halo correlation function is obtained from the power spectrum by

$$\xi^{\mathrm{2h}\prime}_{XY}(r) = \frac{1}{2\pi^2}\int_0^\infty dk\,k^2\frac{\sin(kr)}{kr}P^{\mathrm{2h}}_{XY}(k,r)\,. \tag{5.18}$$

As $\xi^{2h\prime}_{XY}(r)$ has been obtained from a (radius -) restricted sample of galaxy pairs, it is converted to a probability for the whole sample by

$$1 + \xi^{2h}_{XY}(r) = \frac{\bar{n}'_X \bar{n}'_Y(r)}{\bar{n}_X \bar{n}_Y} \left(1 + \xi^{2h\prime}_{XY}(r)\right) . \tag{5.19}$$

5.5.3 Analysis

As described in section 5.4.2, for each luminosity bin sample of interest we measure the projected auto and cross-correlation functions of the SDSS faint and bright galaxy samples

$$(\mathbf{w}_{\mathrm{ff}}, \mathbf{w}_{\mathrm{fb}}, \mathbf{w}_{\mathrm{bb}}) \equiv \mathbf{W}_{\mathrm{S}} , \tag{5.20}$$

where we have introduced the correlation function data vector $\mathbf{w} = (w(r_{\mathrm{p},1}), w(r_{\mathrm{p},2}), ..., w(r_{\mathrm{p},N_{\mathrm{b}in}}))$, and the cross-correlation between $(NUV - r)$ color selected luminosity bin samples and the two SDS galaxy sample

$$(\mathbf{w}_{X\mathrm{f}}, \mathbf{w}_{X\mathrm{b}}) \equiv \mathbf{W}_X , \tag{5.21}$$

with $X \in$ {blue, green, red}).

Ideally one would fit all these cross-correlation functions simultaneously, however this is not practicable: As the $(NUV - r)$ selected galaxy samples are restricted to GALEX + SDSS overlap area, obtaining a joint covariance for the SDSS reference samples and the color selected sample ($\mathrm{Cov}(\mathbf{w}_{\mathrm{ff}}, \mathbf{w}_{\mathrm{fb}}, \mathbf{w}_{\mathrm{bb}}, \mathbf{w}_{X\mathrm{f}}, \mathbf{w}_{X\mathrm{b}})$) would require restricting the SDSS clustering analysis to the combined SDSS + GALEX footprint, which would discard 75% of the SDSS area.[3]

Instead, we first model the SDSS correlation functions and galaxy number densities with an eight parameter HOD described in section 5.5.1.1, and then fit the color bin sample HOD (section 5.5.1.2) using the model for the SDSS samples obtained in the previous step; using the full (non block diagonal) data covariances (figure 5.11) in each step. This method assumes that the color sample - SDSS sample cross-correlations $(\mathbf{w}_{X\mathrm{f}}, \mathbf{w}_{X\mathrm{b}})$ contain little information on the HOD of the SDSS sample compared to the SDSS internal correlation functions used in the first step of the fitting procedure. This assumption is well motivated by statistical uncertainties as the color selected

---

[3] Also note that even if one was willing to discard most of the SDSS data, obtaining an invertible joint covariance for the five different correlation functions, sampled with N_{bin} radial bins, would require dividing the joint footprint into more than $5 \times N_{\mathrm{bin}}$ equal-area jack knife regions N_{sub}. Additionally, the correction factor required to obtain an unbiased estimate of the inverse covariance scales as the ratio of the number of bins (data vector variables) to the number of data sets (Hartlap, Simon and Schneider 2007), resulting either in very large error bars ($N_{\mathrm{sub}} \sim 5N_{\mathrm{bin}}$) or restricting the analysis to very small scales ($N_{\mathrm{sub}} \gg 5N_{\mathrm{bin}}$).

Table 5.3 best-fit HOD model parameters for SDSS samples

$\lg M_\mathrm{r}$	$\lg M^\mathrm{f}_\mathrm{min}$	$\sigma^\mathrm{f}_{\log M}$	$\lg M^\mathrm{f}_0$	$\lg M'^{,\mathrm{f}}_1$	χ^2/d.o.f
[−19.5, −20.0]	11.55 ± 0.04	0.24 ± 0.11	10.14 ± 0.15	12.80 ± 0.03	1.80
[−20.0, −20.5]	11.64 ± 0.02	0.16 ± 0.08	10.07 ± 0.14	12.92 ± 0.03	3.04
[−20.5, −21.0]	11.98 ± 0.09	0.42 ± 0.13	9.45 ± 0.28	13.12 ± 0.04	3.85
[−21.0, −21.5]	12.20 ± 0.03	0.18 ± 0.07	11.77 ± 0.26	13.45 ± 0.03	3.84
$\lg M_\mathrm{r}$	$\lg M^\mathrm{b}_\mathrm{min}$	$\sigma^\mathrm{b}_{\log M}$	$\lg M^\mathrm{b}_0$	$\lg M'^{,\mathrm{b}}_1$	χ^2/d.o.f
[−19.5, −20.0]	12.01 ± 0.04	0.26 ± 0.10	11.08 ± 0.33	13.27 ± 0.03	1.80
[−20.0, −20.5]	12.26 ± 0.03	0.36 ± 0.14	11.91 ± 0.10	13.46 ± 0.03	3.04
[−20.5, −21.0]	12.97 ± 0.08	0.80 ± 0.26	10.59 ± 0.42	13.81 ± 0.03	3.85
[−21.0, −21.5]	13.41 ± 0.04	0.69 ± 0.14	11.95 ± 0.23	14.41 ± 0.03	3.84

samples are over an order of magnitude smaller than the SDSS reference samples. We propagate correlated uncertainties in the HOD model parameters for the SDSS reference sample to the HOD of the color bin sample by marginalizing over 15 randomly chosen models for the SDSS HOD.

Specifically, we compute the χ^2 as

$$\chi^2 = \left(\mathbf{W}_Y^\mathrm{data} - \mathbf{W}_Y^\mathrm{model}\right) \mathrm{Cov}^{-1}(\mathbf{W}_Y) \left(\mathbf{W}_Y^\mathrm{data} - \mathbf{W}_Y^\mathrm{model}\right) + \left(\mathbf{n}_Y^\mathrm{data} - n_Y^\mathrm{model}\right) \mathrm{Cov}^{-1}(\mathbf{n}_Y) \left(\mathbf{n}_Y^\mathrm{data} - n_Y^\mathrm{model}\right) \tag{5.22}$$

where $Y \in \{S, X\}$, with galaxy number densities $\mathbf{n}_S = (n_\mathrm{f}, n_\mathrm{b})$ or $\mathbf{n}_X = n_X$, and with the statistical error on the number densities $\mathrm{Cov}(\mathbf{n}_Y)$ estimated from field to field variations. The HOD parameter space is explored using a Markov Chain Monte Carlo method with a Gaussian distribution of step sizes in each parameter and flat priors $\{\log_{10} M_{\mathrm{m}in}, \log_{10} M_\mathrm{c}^X, \log_{10} M_1, \log_{10} M_1'\} \in [11, 17]$, $\{\sigma_{\log M}, \sigma_X\} \in [0.05, 1.0]$, and $\log_{10} M_0 \in [8, 15]$. At each step a new set of HOD parameters is always accepted if $\chi^2_\mathrm{new} \le \chi^2_\mathrm{old}$, and it is accepted with probability $\exp(-(\chi^2_\mathrm{new} - \chi^2_\mathrm{old})/2)$ if $\chi^2_\mathrm{new} > \chi^2_\mathrm{old}$. The typical chain length is 20000 and we compare multiple realizations to test for convergence.

5.5.4 Results

Our best-fit HOD model parameters for the SDSS samples and their marginalized 1σ errors are given in Tab. 5.3. Our results agree well with the corresponding luminosity threshold samples in the analysis of Zehavi et al. (2010), and we confirm the overall trend of characteristic halo masses for hosting central and satellite galaxies with luminosity threshold. For a detailed comparison note that these two analyses use different fitting formulae for the halo mass function, halo bias and halo mass –concentration relations.

Based on these HOD models for the SDSS reference samples, we now turn to the color selected galaxy samples. Figure 5.13 shows the measured cross-correlation functions between color samples and the SDSS reference samples, the best-fit model correlation functions, and the best-fit halo occupation distribution. Overall, these models provide acceptable fits to the measured correlation functions, with an exception for the green and red galaxy samples in luminosity bin $[-20.5, -21.0]$. These correlation functions have an unusual flat shape and do not show the characteristic transition from one-halo to two-halo term regime. As discussed in section 5.4.3, the redshift of this luminosity bin is centered on the Sloan Great Wall, which is contained almost completely in the angular mask of the SDSS-GALEX cross-match. Hence the clustering measurements in this luminosity bin may be subject to increased sample variance. For comparison we show the cross-correlation functions of $(g - r)$ color identified red galaxies in this luminosity bin computed over the full SDSS area and the combined survey footprint in figure 5.14. The clustering of $(NUV - r)$ and $(g - r)$ selected red galaxies in the joint survey geometry is nearly indistinguishable, while the cross-correlation function of red galaxies in this luminosity bin over the full SDSS area has the expected shape. It can be fit with a color bin HOD model with reduced $\chi^2 = 5.2$, suggesting that the poor fit in figure 5.13 is indeed caused by increased sample variance due to the Great Wall structure and not a systematic effect in the construction of the $(NUV - r)$ selected galaxy sample.

Figure 5.15 shows marginalized constraints on central galaxy halo mass, satellite fraction, and HOD derived galaxy bias for color and luminosity bin samples based on the parameterization described in section 5.5.1.2. Based on this simple parameterization, we find red central galaxies to occupy more massive halos than the average central galaxy from the same luminosity bin. The halo masses of blue and green central galaxies are indistinguishable within the statistical uncertainty of our galaxy samples. At fixed luminosity, the satellite fraction and HOD derived galaxy bias increases with $(NUV - r)$ color, and the host halo masses of green satellite galaxies are intermediate between those of blue and red satellite galaxies. The latter is consistent with the results of Zehavi et al. (2010) who found the satellite fraction to vary smoothly with $(g - r)$ color at fixed luminosity. Their analysis used a one-parameter family of models based on the HOD of the full luminosity threshold sample with only the normalization of the satellite galaxy occupation function as a free parameters, which is insensitive to shifts in the halo mass for central galaxies. For luminosity bin $[-20.5, -21.0]$ we also show results derived from $(g - r)$ selected red galaxies in the full SDSS area to indicate the impact of the Sloan Great Wall. In the Great Wall the satellite fraction and halo mass of red galaxies is increased compared to a more representative survey volume, as expected from the color-density relation. As the $(NUV - r)$ color selected samples in this luminosity bin are subject

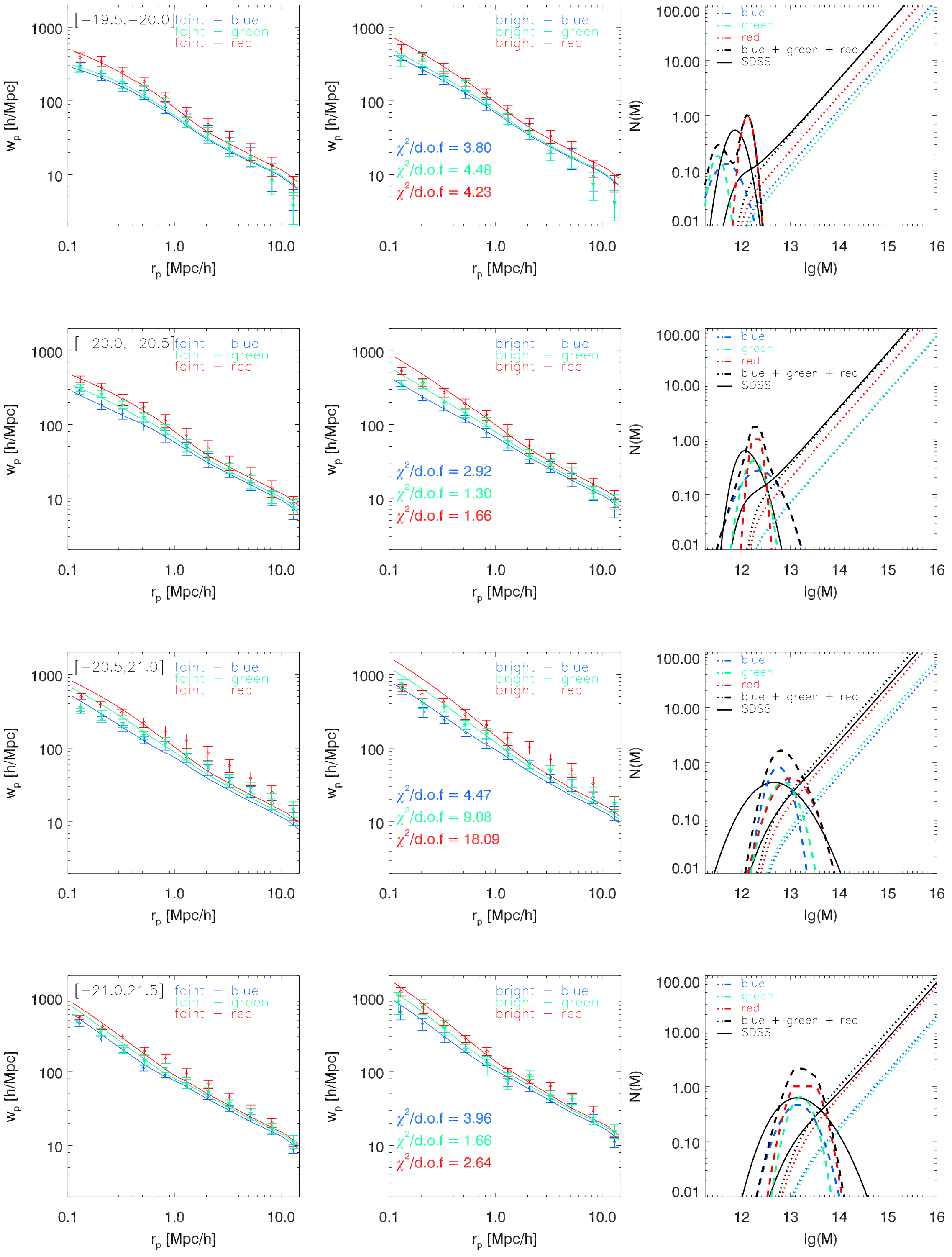


Figure 5.13 Each row shows the measured correlation functions and best-fit HOD of $(NUV - r)$ selected galaxy samples for one luminosity bin. The left/middle panel show the cross-correlation measurements using the faint/bright sample and their joint fit. We list the reduced χ^2 of these fits in the middle panel. The right panel shows the color sample HOD derived from fitting these cross-correlation functions, the sum of all the color samples, and the best-fit HOD of all SDSS galaxies in the same luminosity bin.

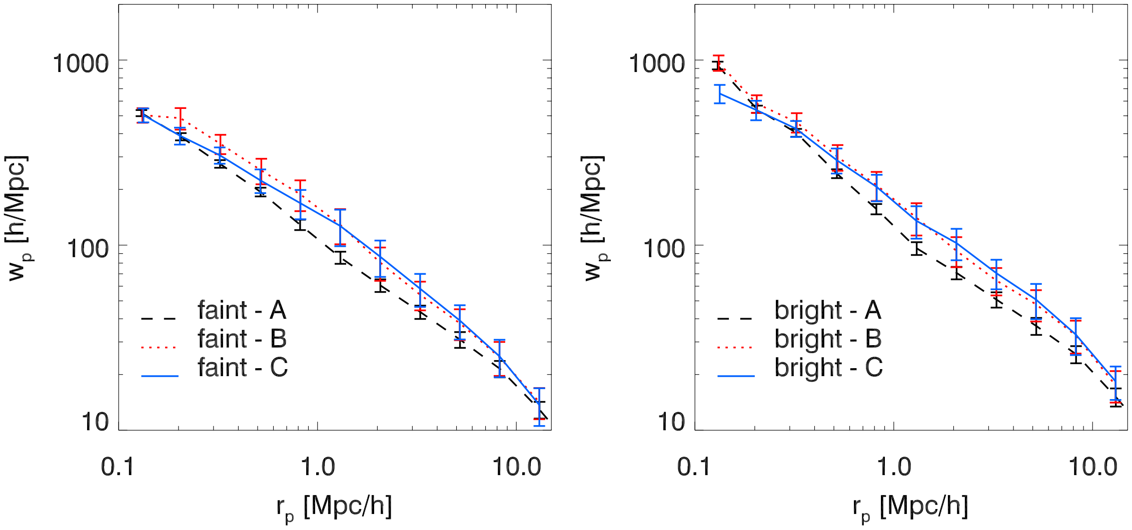


Figure 5.14 cross-correlation functions of red galaxies in luminosity bin [−20.5, −21.0] for different survey areas. The dashed line are the cross-correlation functions with all $(g - r) > 0.85$ galaxies in SDSS in this magnitude bin, the dotted line restricts the SDSS red galaxies to the combined footprint, and the solid line shows the cross-correlation function for $(NUV - r)$ identified red galaxies.

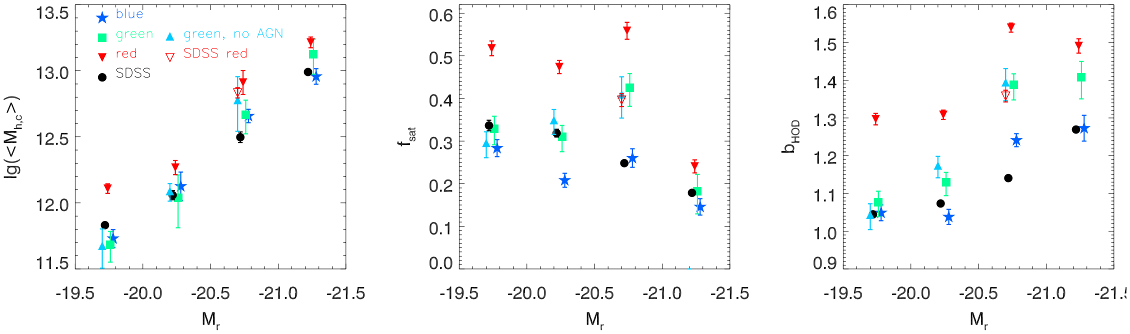


Figure 5.15 Derived HOD parameters for luminosity and color bin samples. Left: Mean halo mass for a halo have a central galaxy from a particular sample. Middle: Satellite fraction as a function of galaxy luminosity and color. Right: Galaxy bias derived from the HOD model fit.

to increased sample variance, the results for blue and green galaxies in this luminosity bin should similarly be interpreted with caution.

As noted by Martin et al. (2007) and Salim et al. (2007), a large fraction of active galactic nuclei (AGN) have green ($NUV-r$) colors. These galaxies may be transitional galaxies with star formation being quenched by AGN feedback (e.g., after undergoing a major merger, Springel, Di Matteo and Hernquist 2005), or red sequence interlopers which appear green due to the NUV AGN continuum emission. We test whether the intermediate clustering of green valley galaxies is caused by AGN, which may be a different population than the non-AGN transitional galaxies. We identify green AGN through emission line diagrams (figure 5.16, Baldwin, Phillips and Terlevich 1981) using the Kewley et al. (2001) extreme starburst classification line. We use the emission line measurements from the MPA-JHU catalog and require a signal-to-noise $S/N \geq 3$ in the emission lines. Our goal is to remove any potential AGN contamination from the green valley galaxy sample, and we remove all galaxies which are classified as AGN in at least one of the three diagrams as this allows us to categorize galaxies which do not meet the S/N threshold for all emission line. After redoing our clustering and HOD analysis for non-AGN green galaxies we find the HOD of green non-AGN galaxies to be indistinguishable of that of green galaxies including AGN. We do not show results derived from HOD fits for the non-AGN green valley galaxies in luminosity bin $[-21.0, -21.5]$ as this sample is too small to obtain stable covariances.

5.6 Summary and Discussion

In this chapter we introduced a new HOD modeling technique for galaxy cross-correlation functions using multiple tracer populations. This approach is particularly useful for interpreting the clustering of sparse and/or luminosity bin selected galaxy samples of interest. It is advantageous for the analysis of sparse galaxy samples as considering the cross-correlation function with more abundant galaxy populations significantly reduces the statistical uncertainty. While the galaxy number density provides strong constraints on the HOD of luminosity threshold samples, the HOD of luminosity bin samples is independent of the galaxy abundance; in this case considering the cross-correlation with multiple tracer populations is particularly useful as it provides an additional mass scale for the calibration of the luminosity bin HOD. This allows us to constrain the central galaxy HOD of color and luminosity bin selected samples for the first time.

We apply this multiple tracer technique to analyze the clustering of $(NUV - r)$ color selected

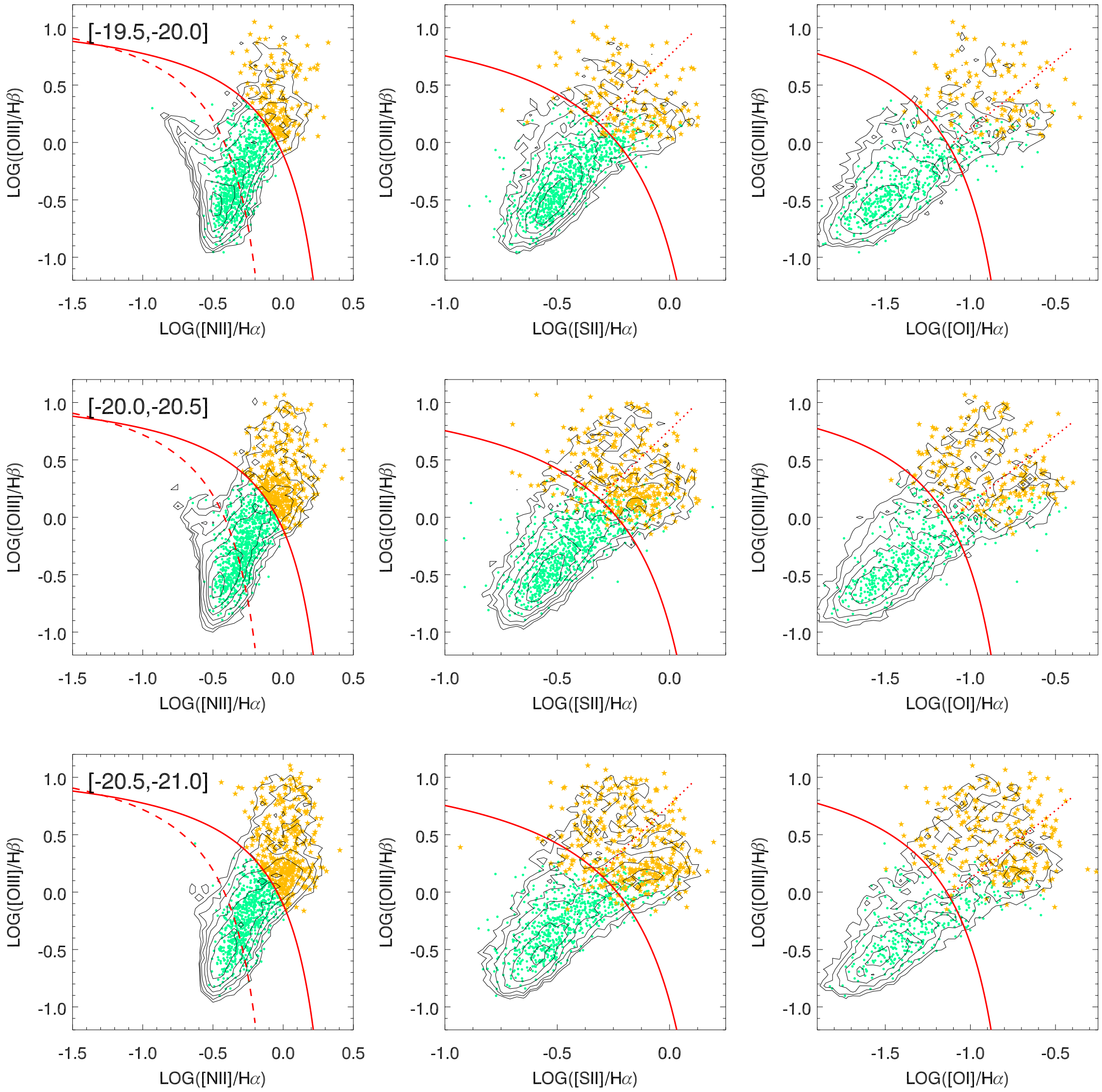


Figure 5.16 BPT emission line diagrams (Baldwin, Phillips and Terlevich 1981) used to identify green AGN galaxies for different luminosity bins. The black contours show the distribution of all SDSS galaxies within the given luminosity and redshift range, symbols show the distribution of green valley galaxies, with orange stars indicating galaxies identified as AGN based on the [NII] – [OIII]diagram. The red solid line shows the Kewley et al. (2001) extreme starburst classification line, the red dashed line the Kauffmann et al. (2003) pure star formation line, and, for completeness, the red dotted lines show the Seyfert/LINER separation lines from Kewley et al. (2006).

green valley galaxy samples. At fixed luminosity we find the satellite fraction and host halo mass to increase with $(NUV - r)$ color. Specifically, we find the halo masses of blue and green central galaxies to be indistinguishable, and less massive than those of red central galaxies, while there is a smooth trend with color for satellite galaxies.

The reduced χ^2 values of the best-fit HODs in our analysis of color selected galaxy samples are relatively large, and our model is particularly insufficient to reproduce the clustering of galaxies in or near the Sloan Great Wall. Overall, it is not surprising that a two parameter HOD model is insufficient to describe the color dependent clustering of galaxies. While the HOD formalism works reasonably well to describe the overall relation between (color independent) galaxies and their halos, it is unlikely that the strong assumptions implicit in the HOD formalism hold for each sub-population. For example, the different color samples follow different radial distributions within a halo (e.g., von der Linden et al. 2010, find the fraction of red galaxies to increase toward the cluster center). Similarly, the efficiency of star formation quenching in satellite galaxies depends on halo mass (e.g., Tinker, Wetzel and Conroy 2011, Wetzel, Tinker and Conroy 2011). Additionally these observational results indicate that the influence of massive halos may extend beyond R_{200}, e.g., through highly eccentric satellite orbits (Benson 2005) and infall related shocks extending beyond the virial radius (e.g., Balogh, Navarro and Morris 2000), which is not easily incorporated in halo models.

We are attempting to address some of these issues by extending of HOD parameterization to incorporate different radial profiles by varying the slope of the satellite galaxy halo-occupation distribution and the normalization of the halo mass–(galaxy density profile) concentration normalization as a function of color and luminosity. Preliminary results favor more concentrated profiles for red galaxies, and more extended galaxy distributions for blue and green galaxies. The slope of the satellite galaxy halo-occupation distribution seems to be poorly constrained, and we leave the details of this analysis to future work.

Finally we note that Behroozi, Conroy and Wechsler (2010), Leauthaud et al. (2011) recently proposed an improved HOD parameterization based on a detailed model for the relation between stellar mass and halo mass. Their results (figure 3 in Leauthaud et al. (2011)) indicate that halo masses derived from the HOD parameterization for luminosity threshold samples adopted in our analysis (equation (5.4)) may be biased by up to 40%, with the main source of this discrepancy being the assumptions of a power-law form and constant scatter for the luminosity-halo mass relation. For luminosity bin samples, however, these assumptions are better justified, and we expect only small discrepancies between different HOD parameterizations.

Acknowledgements

We thank Don Neill, Mark Seibert, and Ted Wyder for early access to the GR6 GMSC and many explanations about GALEX data. We also acknowledge helpful discussions with David Weinberg and Zheng Zheng about HOD modeling of binned samples.

Chapter 6

Merger Induced Scatter and Bias in the Cluster Mass–Sunyaev-Zeldovich Effect Scaling Relation [*]

6.1 Introduction

Clusters of galaxies are the most massive gravitationally bound objects in the universe, which makes them an important tool for cosmology: among other tests, their abundance provides information on the gravitational growth of structures and is regulated by the initial density field, gravity, and the expansion history of the universe, which critically depend on the underlying cosmology. Thus number counts of clusters, for which masses and redshifts are known, can be used to constrain cosmological parameters (see Allen, Evrard and Mantz 2011, for a recent review).

To relate observed number counts to theoretical predictions of the cluster mass function, these experiments need to infer cluster masses from observables. The thermal Sunyaev Zeldovich (SZ) effect, the signature of inverse Compton scattering of cosmic microwave background photons with hot cluster electron, is thought to provide an excellent mass proxy as the SZ signal is proportional to the total thermal energy of a cluster and is thus less affected by physical processes in the cluster core which can largely affect the X-ray luminosity. This is confirmed by simulations (e.g., Nagai 2006, Shaw, Holder and Bode 2008, Battaglia et al. 2010, Sehgal et al. 2010) finding the scatter in the mass–SZ scaling relation to be of order 5% - 10%. Furthermore, the SZ effect is not subject to surface brightness dimming and has a very weak redshift dependence, making it an ideal probe to study galaxy clusters at high redshift.

---

[*] This chapter was adapted from *Merger induced scatter and bias in the cluster mass - Sunyaev-Zeldovich effect scaling relation*, Elisabeth Krause, Elena Pierpaoli, Klaus Dolag, and Stefano Borgani; *MNRAS*, **419**, 1766 (2012). Reproduced here with permission, copyright (2011) by Wiley-Blackwell.

Currently several large surveys are starting to detect hundreds of galaxy clusters through their SZ signal (Vanderlinde et al. 2010, Marriage et al. 2011, Planck Collaboration et al. 2011a) and derive cosmological constraint based on these samples (Andersson et al. 2011, Sehgal et al. 2011, Williamson et al. 2011). To exploit the statistical power of these upcoming cluster samples, the mapping between SZ signal and cluster mass needs to be well understood. Observations find normalization and slope of the scaling relations between SZ signal and lensing derived masses (Marrone et al. 2011), or between SZ signal and X-ray properties (Planck Collaboration et al. 2011b,c) to be consistent with self-similar scaling and predictions from simulations.

Due to the steep slope of the cluster mass function, competitive cosmological constraints from these experiments require information about the distribution and redshift evolution of scatter in the mass scaling relation (e.g., Majumdar and Mohr 2004, Lima and Hu 2005, Shaw, Holder and Dudley 2010). As the true cluster mass and other physical cluster properties which may bias the mass proxy are unobservable, and as the noise and biases in the different mass estimators may be correlated, characterizing the intrinsic scatter in any of these scaling relation is difficult to obtain from observations. Hence the sources and distribution of scatter in different mass estimators are mainly studied through simulations and mock observations (e.g., Rasia et al. 2006, Nagai, Vikhlinin and Kravtsov 2007, Shaw, Holder and Bode 2008, Becker and Kravtsov 2010, Yang, Bhattacharya and Ricker 2010, Fabjan et al. 2011).

In this work we focus on the effect of merging events on the SZ signal of a galaxy cluster. As clusters form through merging of smaller objects, these are frequent and disruptive events, which may alter the physical state of the involved clusters significantly. Hence merging clusters may deviate from the scaling relations observed in relaxed clusters and, as the fraction of morphologically disturbed clusters increases with redshift, cause a redshift dependent scatter or bias in the mass scaling relation. Simulations of binary cluster mergers (Randall, Sarazin and Ricker 2002, Poole et al. 2006, 2007, Wik et al. 2008) find that the X-ray luminosities, temperatures, SZ central Compton parameters and integrated SZ fluxes increase rapidly during the first and second passage of the merging clusters. The clusters temporarily drift away from mass scaling relations and return to their initial scaling relation as the merging system virializes. These transient merger boosts found in binary mergers and some observations (Smith et al. 2003a) can scatter the inferred masses toward higher values and thus bias the derived cosmology toward a higher normalization of the power spectrum, σ_8, and lower matter density (Randall, Sarazin and Ricker 2002, Smith et al. 2003a, Wik et al. 2008, Angrick and Bartelmann 2011). On the other hand, mergers increase the non-thermal pressure support (Rasia et al. 2006, Lau, Kravtsov and Nagai 2009, Battaglia et al. 2010) found in cluster outskirts, and due to partial virialization merging clusters can appear cooler than relaxed clusters of the same mass (e.g., Mathiesen and Evrard 2001). For a cluster sample extracted from cosmological simulations, Kravtsov, Vikhlinin and Nagai (2006) find the X-ray temperatures of morphologically disturbed clusters to be biased, while the X-ray derived SZ-equivalent Y_X shows no significant correlation with cluster structure. Comparing X-ray and SZ to weak lensing derived masses, Okabe et al. (2010) and Marrone et al. (2011) found undisturbed clusters to have of order $\sim 40\%$ higher weak lensing masses than disturbed clusters at fixed T and $Y_{\rm SZ}$, and $\sim 20\%$ higher weak lensing masses at fixed $Y_{\rm X}$.

Our goal is to isolate how mergers in a cosmological context affect the SZ signal of clusters, and if merging cluster can be detected as outliers of scaling relations. This extends previous work, as our analysis includes both multiple mergers with realistic distributions of orbits and mass ratios, and full SPH treatment of gas physics with radiative cooling, star formation and supernova feedback. The simulations and the cluster sample are described in section 6.2. We discussion the best-fit scaling relations and their scatter in section 6.3. The effect of merging events of the clusters SZ signal is quantified and the evolution of merging clusters with respect to the scaling relations is discussed in section 6.4. In section 6.5 we investigate if the dynamical state of clusters can be inferred from the morphology of the SZ signal. We summarize our results and conclude in section 6.6.

6.2 Simulations

This analysis is based on two samples of galaxy clusters extracted from cosmological hydrodynamics simulations. In this section we summarize the simulated physics and describe the derived quantities used in our analysis.

6.2.1 Cluster Samples

Sample A To study the time evolution of the cluster SZ signal we use a sample of 39 galaxy groups and clusters with virial masses above $3\times10^{13}M_\odot/h$ from simulations presented in Dolag et al. (2006, 2009). 25 of these clusters are more massive than $10^{14}M_\odot/h$. These structures were identified as 10 different regions in a $(479{\rm Mpc}/h)^3$ dark-matter-only cosmological simulation (Yoshida et al. 2001), and re-simulated at higher resolution using the Zoomed Initial Conditions method (Tormen, Bouchet and White 1997). The re-simulations, described in detail in Dolag et al. (2006), are carried out with GADGET-2 (Springel, Di Matteo and Hernquist 2005), and include a uniform, evolving UV-background and radiative cooling assuming an optically thin gas of primordial composition. Star formation is included using the two-phase model of the interstellar medium (ISM) by Springel and Hernquist (2003). In this sub-resolution model the ISM is described as cold clouds, providing a reservoir for star formation, embedded in the hot phase of the ISM. Star formation is self-regulated through energy injection from supernovae evaporating the cold phase. Additional feedback is incorporated in the form of galactic winds triggered by supernovae that drive mass outflows (Springel and Hernquist 2003).

The simulation assumes a flat ΛCDM cosmology with $(\Omega_{\rm m}, \Omega_{\rm b}, \sigma_8, h) = (0.3, 0.04, 0.9, 0.7)$. It has a mass resolution of $m_{\rm DM} = 1.1 \times 10^9 M_\odot/h$ and $m_{\rm gas} = 1.7 \times 10^8 M_\odot/h$ and the physical softening length is $\epsilon = 5{\rm kpc}/h$ over the redshift range of interest. Our analysis is based on 52 snapshots covering the redshift range $z = 1$ to $z = 0$ and separated evenly in time with a spacing of 154 Myrs between snapshots.

Sample B The second cluster sample is a volume-limited sample of 117 clusters at $z = 0$ described in Borgani et al. (2004). These clusters are identified in a $(192{\rm Mpc}/h)^3$ cosmological SPH simulation carried out with GADGET-2 and using the same physics as described above. This simulation assumes a flat ΛCDM cosmology with $(\Omega_{\rm m}, \Omega_{\rm b}, \sigma_8, h) = (0.3, 0.04, 0.8, 0.7)$. The mass resolution is $m_{\rm DM} = 4.6 \times 10^9 M_\odot/h$ and $m_{\rm gas} = 6.9 \times 10^8 M_\odot/h$, the physical softening length at $z = 0$ is $\epsilon = 7.5{\rm kpc}/h$.

6.2.2 Masses and Merging Histories

Halos are identified using a friend-of-friends algorithm, and the cluster center is defined by the particle in a halo with the minimum gravitational potential. Cluster radii R_Δ and masses M_Δ are defined through spherical regions around the cluster center within which the average density is Δ times the critical density of the universe,

$$\int_0^{R_\Delta} \rho(r)\, 4\pi r^2\, dr = \frac{4\pi}{3} R_\Delta^3 \Delta \rho_{\rm crit} = M_\Delta \ . \tag{6.1}$$

We identify mergers by a mass jump criteria applied to the mass history of the main progenitor. Motivated by the findings that the average mass accretion history of halos is well described by exponential growth with redshift (Wechsler et al. 2002, McBride, Fakhouri and Ma 2009) and that

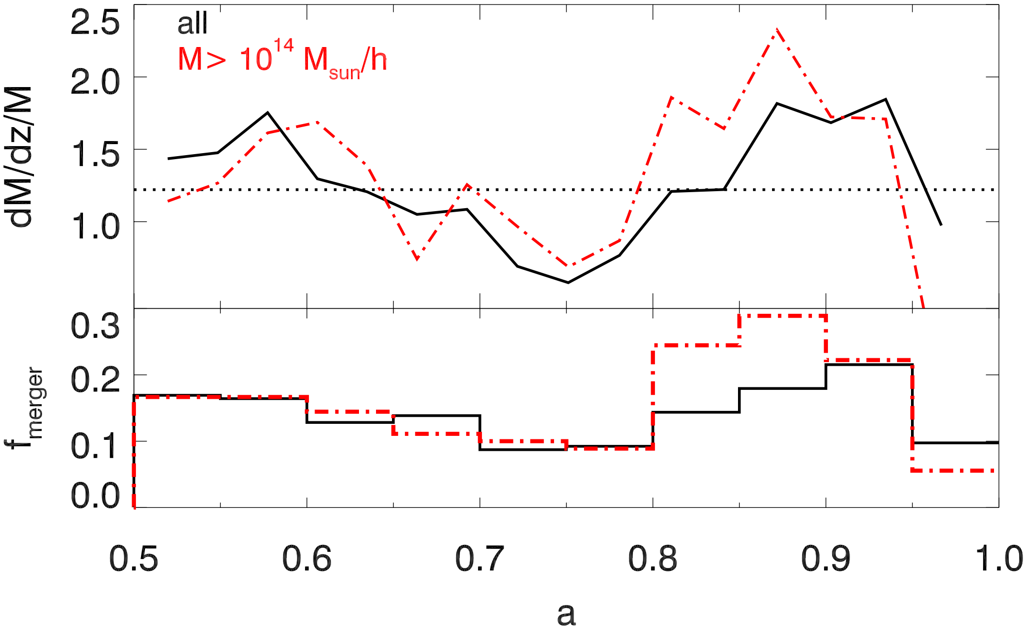


Figure 6.1 Fractional accretion rate per unit redshift (top panel) and merger fraction as a function of scale factor. The solid shows the complete sample A, the dash-dotted line a subsample of massive clusters. The dotted line indicates the overall mean accretion rate. Accretion rate (merger fraction) are averaged over 3 (5) neighboring simulation snapshots to reduce noise.

the average merger rate per halo per unit redshift is nearly constant for a wide range of halo masses and redshifts (Fakhouri and Ma 2008), we select merging events based on a threshold in fractional mass accretion rate per unit redshift $\mathrm{d}M/\mathrm{d}z/M > \zeta_\mathrm{m}$. We choose ζ_m such that halos accrete on average 30% of the mass accreted since its formation redshift z_f, defined as the redshift at which a halo reaches half its present day mass, during mergers. We checked that our results are insensitive to the exact choice of ζ_m: We find similar trends for any merger definition $\zeta_\mathrm{M} \geq \langle \mathrm{d}M/\mathrm{d}z/M \rangle_{z,\mathrm{cluster}}$ that requires the accretion rate $\mathrm{d}M/\mathrm{d}z/M$ during mergers to be larger than the mean accretion rate (cf. discussion of figure 6.6).

Figure 6.1 confirms that this merger definition does not strongly depend on cluster mass or redshift. The top panel shows the mean accretion rate as a function of scale factor for all clusters (solid line) and massive clusters ($M \geq 10^{14} M_\odot/h$, dash dotted line), and the overall mean accretion rate (dotted line). The lower panel shows the fraction of clusters that are merging as a function of scale factor. There is a peak of merging activity around $a = 0.9$, but the accretion rate and merger fraction show no clear trends with cluster mass or redshift.

6.2.2.1 Comparison to the Millenium Run

The 39 cluster and group-scale sized halos in sample A are extracted from 10 re-simulation regions selected from a large simulation box. One of the re-simulated regions hosts a filamentary structure with four massive clusters ($M > 10^{15} M_\odot/h$), and three of the re-simulation regions hosting other massive clusters contain several other smaller clusters. The re-simulation technique allows us to analyze the evolution of these regions of interest in their cosmological context at a higher resolution. As a result of the re-simulation strategy, the mass distribution of this sample does not follow the cluster mass function, and clusters which are not the most massive object in their re-simulation region live in denser regions than an average cluster of the same mass in a volume-limited sample. In the following discussion we refer to the most massive objects in their respective re-simualtion region as primary clusters, and all others as secondary clusters.

Simulations indicate a dependence of halo formation histories on environment with merger being more frequent in dense environments and late-forming massive clusters living in denser environments than earlier forming clusters of the same mass (Gao, Springel and White 2005, Wechsler et al. 2006, Fakhouri and Ma 2009). Hence the merging histories of cluster sample A might not be representative of those of a volume-limited sample. To assess the impact of our sample selection on halo formation histories we compare the formation redshifts of primary and secondary clusters in sample A and halos in the Millenium run simulation (Springel et al. 2005) in figure 6.2.

The symbols show the present day masses and formation redshift $z_{\rm f}$ for all clusters in sample A. Primary clusters are indicated by star symbols. The dashed and dotted lines are a fit to the mean formation time and its 1σ scatter for halos in the Millenium Run from McBride, Fakhouri and Ma (2009). We convert the fitting formula from friend-of-friends halo mass with linking length $b = 0.2$ to M_{200} assuming a constant conversion factor $M_{200} = 0.7 M_{\rm FOF}$. For the mass range of our sample this conversion underestimates M_{200} [1] and biases the fit for $z_{\rm f}$ to more recent times.

Due to the differences in matter density used in simulation A ($\Omega_{\rm M} = 0.3$) and in the Millenium Run ($\Omega_{\rm m} = 0.25$) the average clusters in simulation A forms earlier than a cluster of the same mass in the Millenium Run. Hence formation redshifts for primary clusters in sample A are broadly consistent with the formation history of halos in the Millenium run. Figure 6.2 suggests that secondary clusters in sample A may form somewhat later than primary clusters of the same mass. However, the distribution of formation redshifts at fixed mass is not expected to be symmetric but to have a long tail toward later formation times, and the comparison is limited by the small number objects. Overall, we expect the merging histories analyzed in this study to be similar to those found in a volume-limited sample.

6.2.3 SZ Maps

The amplitude of the thermal SZ effect along a line of sight is proportional to the Compton y parameter

$$y = \frac{k_{\rm B}\sigma_{\rm T}}{m_{\rm e}c^2} \int {\rm d}l\, n_{\rm e} T_{\rm e}\,, \tag{6.2}$$

where $n_{\rm e}$ and $T_{\rm e}$ are the electron density and temperature, $k_{\rm B}$ is the Boltzmann constant, $\sigma_{\rm T}$ the Thomson cross section, $m_{\rm e}$ the electron rest mass, and c the speed of light. For each cluster we analyze Compton y parameter maps obtained from three orthogonal lines of sight. For sample A the projection depth is 8 Mpc and maps are produced using the map making tool Smac (Dolag et al. 2005) and the JobRunner web application.[2] For sample B we use projected maps which include all material with $6R_{\rm vir}$ described in Ameglio et al. (2007). From these maps we measure integrated Y_Δ

---

[1] For equal mass particles, a FOF group with linking length b is bounded by a surface of density $3\Omega_{\rm m}\rho_{\rm crit}/(2\pi b^3)$ (White 2002). Assuming that halos follow NFW-profiles with concentration $c = (4, 7, 10)$, the ration between M_{200} and $M_{\rm FOF}$ with $b = 0.2$ in the Millenium run cosmology is given by $(0.71, 0.80, 0.85)$. In practice however, the conversion between these mass definitions is complicated by deviations from the NFW-profile and spherical symmetry.

[2] Access to the cluster simulations of sample A, including web services allowing to interactively produce various kinds of maps, are publicly available via the web portal at http://www.mpa-garching.mpg.de/HydroSims

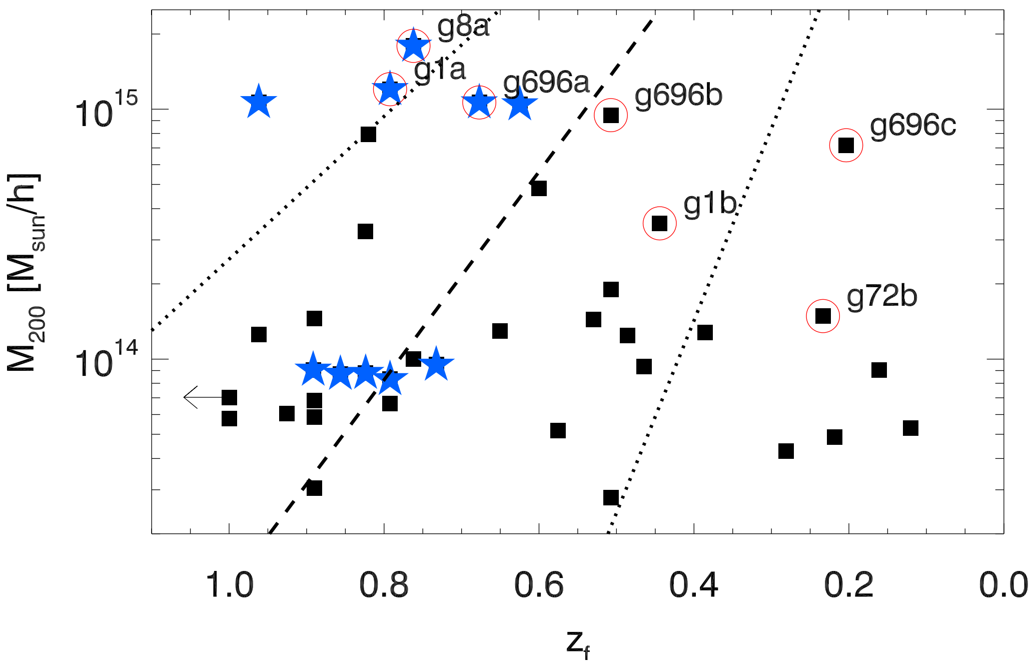


Figure 6.2 Distribution of formation redshifts z_f as a function of cluster mass. The symbols represent clusters from sample A, the most massive cluster in each re-simulation region is marked with a star symbol. Dashed (dotted) lines show a fit to the mean (1σ scatter) formation redshift as a function of friends-of-friends mass found in the Millenium Run (McBride, Fakhouri and Ma 2009), converted to spherical overdensity mass assuming $M_{200} = 0.7 M_{\rm FOF}$ (see text for details). Formation redshift is defined as the redshift at which a halo reaches half its present day mass. One cluster in sample A forms before $z = 1$, indicated by the left arrow. Open circles indicate the clusters shown as examples in subsequent plots, labels indicate the names of these clusters in table 1 of Dolag et al. (2009).

parameters within different overdensity radii ($R_{2500}, R_{500}, R_{200}, R_{\rm vir}$)

$$Y_\Delta = \frac{k_{\rm B}\sigma_{\rm T}}{m_{\rm e}c^2} \int_{V_\Delta} {\rm d}V\, n_{\rm e} T_{\rm e} \tag{6.3}$$

where the integration volume is a cylinder of radius R_Δ and height 8 Mpc (or 12 $R_{\rm vir}$) for sample A (or B). This definition of the integrated Y parameter includes projection effects due to halo triaxiality and nearby structures within the projection cylinder, but does not account for projection effects from uncorrelated large-scale structure along the line of sight.

6.3 Mass Scaling Relations

Self-similar clusters models predict the gas temperature to scale as

$$T \propto (ME(z))^{2/3}\ . \tag{6.4}$$

Hence the self-similar prediction for the relation between integrated Compton Y parameter and mass is

$$Y_\Delta \propto M_{{\rm gas},\Delta} T \propto f_{\rm gas} M_\Delta^{5/3} E^{2/3}(z)\,. \tag{6.5}$$

In this section we determine the best-fit scaling relations for the simulated clusters and discuss the scatter in these relations, focussing on the role of mergers.

6.3.1 Best-fit Scaling Relations

We now determine the best-fit $M_\Delta(Y_\Delta)$ scaling relation

$$M_\Delta(Y_\Delta) = 10^A \left(\frac{Y_\Delta}{\rm kpc^2}\right)^\alpha E^\beta(z)\, 10^{14} M_\odot/h \tag{6.6}$$

and $Y_\Delta(D_\Delta)$ scaling relation

$$Y_\Delta(M_\Delta) = 10^B \left(\frac{M_\Delta}{10^{14} M_\odot/h}\right)^\gamma E^\delta(z)\, {\rm kpc}^2\,, \tag{6.7}$$

where the self-similar predictions are $(\alpha,\beta) = (3/5, -2/5)$ and $(\gamma,\delta) = (5/3, 2/3)$. Specifically we first fit a line to the $\lg(Y_\Delta) - \lg(M_\Delta)$ distribution at each redshift, and then determine the redshift dependence by determining a linear fit in $\lg(E(z))$ to the evolution of the normalization constant

Table 6.1 Best-fit $M_\Delta(Y_\Delta)$ scaling relation parameters (equation (6.6))and logarithmic scatter σ_M at fixed Y, defined analogously to equation (6.8), A*/B* denote samples A/B restricted to clusters at $z = 0$ with $M > 2 \times 10^{14} M_\odot/h$

Sample	Δ	$A(z=0)$	α	β	σ_M
A	200	-0.348 ± 0.007	0.639 ± 0.010	-0.57 ± 0.08	0.063
A*	200	-0.281 ± 0.042	0.588 ± 0.020	-	0.042
B	200	-0.297 ± 0.006	0.617 ± 0.007	-	0.042
B*	200	-0.261 ± 0.014	0.593 ± 0.010	-	0.027
A	500	-0.466 ± 0.001	0.641 ± 0.007	-0.74 ± 0.10	0.089
A*	500	-0.406 ± 0.036	0.607 ± 0.020	-	0.042
B	500	-0.400 ± 0.004	0.626 ± 0.005	-	0.037
B*	500	-0.379 ± 0.011	0.604 ± 0.009	-	0.024

Table 6.2 Best-fit $Y_\Delta(M_\Delta)$ scaling relation parameters (equation (6.7)) and logarithmic scatter σ_Y at fixed mass, A*/B* denote sample A/B restricted to clusters at $z = 0$ with $M > 2 \times 10^{14} M_\odot/h$

Sample	Δ	$B(z=0)$	γ	δ	σ_Y
A	200	0.547 ± 0.003	1.560 ± 0.014	0.85 ± 0.10	0.103
A*	200	0.489 ± 0.052	1.648 ± 0.056	-	0.070
B	200	0.494 ± 0.005	1.555 ± 0.017	-	0.071
B*	200	0.445 ± 0.030	1.668 ± 0.044	-	0.046
A	500	0.714 ± 0.003	1.553 ± 0.017	1.03 ± 0.14	0.136
A*	500	0.697 ± 0.038	1.601 ± 0.051	-	0.068
B	500	0.641 ± 0.003	1.556 ± 0.014	-	0.059
B*	500	0.624 ± 0.013	1.637 ± 0.027	-	0.037

$B(z)$. We find no significant indication for a redshift evolution of the slope α or γ.

The best-fit parameters and the logarithmic scatter at fixed mass,

$$\sigma_Y = \left(\frac{\sum_{i=1}^{N} (\lg(Y_i/Y(M_i)))^2}{N-2} \right)^{1/2} , \tag{6.8}$$

where the sum runs over all Y measurements (three projections of each cluster at each redshift), are given in table 6.1 and table 6.2.

The two scaling relations contain the same information. While the $M(Y)$ scaling relation is the relation of more interest for cosmology and is the relation used in the rest of our analysis, the $Y(M)$ relation is easier to interpret if one is more used to thinking about clusters properties at fixed mass rather than at fixed Y, and we will focus the discussion of the fit results on this relation.

The slope γ of the best-fit relation in samples A and B is below the self-similar value, while other simulations including cooling and star formation find slopes comparable to or steeper than the self-similar predictions (Nagai 2006, Battaglia et al. 2010, Sehgal et al. 2010). We find a slope in agreement with previous results if we only consider massive clusters with $M_{200} > 2 \times 10^{14} M_\odot/h$ ("Sample B*") which is identical to the mass threshold used in Sehgal et al. (2010). Projection effects may account for some of the difference with the results of Nagai (2006) and Battaglia et al. (2010): these authors use spherically averaged Y measurements and do not include projection effects, which effectively boost the integrated Y signal of lower mass clusters[3] and hence lower the slope of the scaling relation.

After accounting for differences in the baryon fractions of different simulations, the normalization B of the best-fit scaling relation for sample B* is consistent with those obtained from other hydrodynamical simulations with similar physics (the csf run in Nagai (2006) and the radiative run in Battaglia et al. (2010)).

The slope and normalization of the scaling relation for a subsample of massive clusters at $z = 0$ from sample A, denoted as A*, are comparable to those found for the sample B*. A direct comparison of these numbers is complicated by the fact that slope and scatter of the scaling relations are mass dependent, and that the mass distribution within sample A does not follow the cluster mass function. Also sample A* consists of only 11 clusters, five of these are the most massive objects in their respective re-simulation region, and it is hard to assess at a precision cosmology level whether

---

[3] Projection effects introduce an additive signal $Y_\mathrm{p} \geq 0$ which scales as $Y_{\mathrm{p},\Delta} \propto R_\Delta^2 \propto M^{2/3}$, and thus the fractional error induced by projection effects decreases with cluster mass.

the non-representative environment of clusters in sample A affects the normalization of their scaling relation.

The redshift evolution of the scaling relation for sample A deviates significantly from self-similar expectations. This deviation may be caused by mergers: As we will discuss in detail in section 6.4 the Y signal of recently merged clusters is suppressed on timescales of order a few Myr. As the merger rate per halo per unit time increases with redshift, the increasing fraction of recently merged clusters reduces the normalization of the scaling relation, causing δ to deviate from the self-similar value.

In the following we will focus on scaling relations within R_{200} as the $M_{200} - Y_{200}$ relation for sample A has less scatter than that within R_{500}. The accretion histories at R_{500} are more erratic than at R_{200}, which complicates the identification of merging events and the interpretation of trajectories in the $M-Y$ plane. At the time resolution of the simulation snapshots, infalling substructures sometimes cross in and out of R_{500} before coalescence, causing a series of mass jumps and mass losses in M_{500}. While it is not clear what the best mass definition is for a merging cluster, the scatter in the $M_{\rm vir} - M_\Delta$ relation illustrates that masses within larger radii are less volatile: fitting M_Δ as a power law in $M_{\rm vir}$ and $E(z)$ we find logarithmic scatter $(\sigma_{M_{200}}, \sigma_{M_{500}}, \sigma_{M_{2500}}) = (0.046, 0.108, 0.326)$.

Figure 6.3 shows the best-fit Y_{200}-M_{200} scaling relation for sample A and the distribution of the $z = 1$ and $z = 0$ clusters, which we plot in the form of the SZ signal scaled for redshift evolution,

$$\tilde{Y}_{200}(z) = Y_{200}(z) E^{\beta/\alpha}(z)\,. \tag{6.9}$$

The right panel shows the distribution of the scatter around the scaling relation,

$$\delta \lg M \equiv \lg\left(M(Y)/M\right)\,, \tag{6.10}$$

for the full sample and subsamples. This scatter definition gives the logarithmic error in the mass inferred from Y measurements, positive scatter corresponds to clusters with Y larger than expected for their actual mass. At all redshifts the distribution deviates from lognormality with a tail at large $\delta \lg M$, causing the distribution to have positive skewness and kurtosis.

The left panel of figure 6.4 shows the M_{200} and Y_{200} data from sample B and the best-fit scaling relation. We checked by visual inspection that the most extreme outliers, which are all in the direction of Y higher than expected for the cluster mass, are indeed projection effects. These clusters have

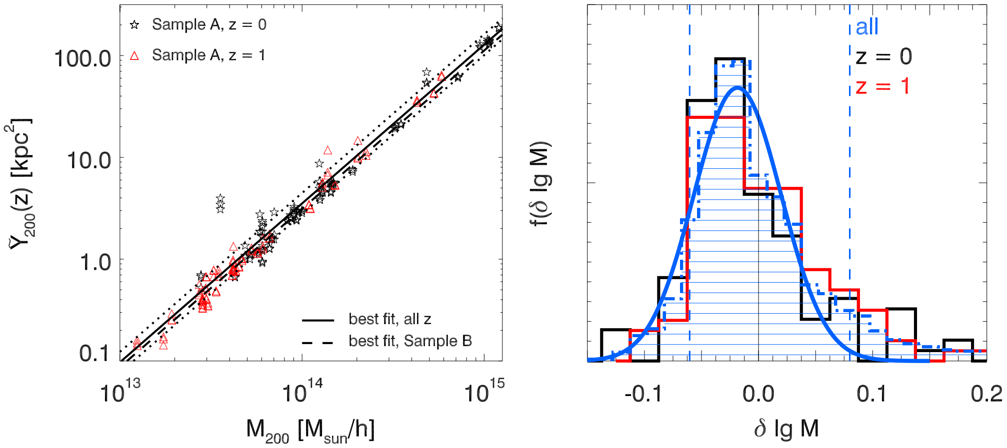


Figure 6.3 Left: Relation between mass M_{200} and integrated Compton Y_{200} parameter for the $z = 0$ (stars) and $z = 1$ clusters (triangles) in sample A. The Compton Y parameter has been scaled to absorb the redshift evolution of the scaling relation in order to show the power law relation $M \propto \tilde{Y}^{\alpha}(z) = (YE^{\beta/\alpha}(z))^{\alpha}$. The solid and dotted lines show the best-fit scaling relation for sample A and its 1σ error. For reference, the dashed line indicates the best-fit scaling relation for sample B.
Right: Distribution of residuals of the best-fit scaling relation for the full sample (filled histogram) and the redshift subsamples (black/red line), and the best-fit Gaussian to the full distribution. The vertical dashed lines illustrate the 10% and 90% quantile for the full sample, illustrating the non-lognormality of the scatter distribution.

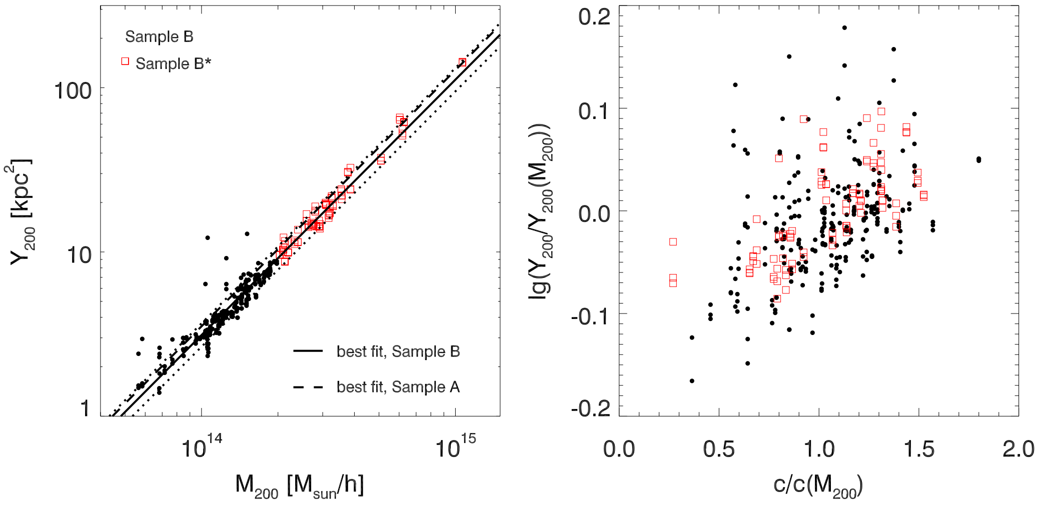


Figure 6.4 Left: Relation between mass M_{200} and integrated Compton Y_{200} parameter for cluster sample B. Massive clusters with $M_{200} > 2 \times 10^{14} M_{\odot}/h$ are shown with filled symbols. The solid and dotted lines show the best-fit scaling relation for sample B and its 1σ error. For reference, the dashed line indicates the best-fit scaling relation for sample A. The strong outliers with boosted Y signal in the low-mass range are visually identified to be caused by projection effects.
Right: Residuals of the Y–M relation at fixed mass vs. scatter in the mass – halo concentration relation at fixed mass. Concentration measurements are from Ameglio et al. (2009), see text for details on the determination of $c/c(M_{200})$.

multiple peaks or appear otherwise distorted in only one or two of the three orthogonal projections, indicating that these are not merging systems (yet).

The intrinsic scatter in the spherically integrated Y parameter of large cluster samples has been found to be close to log-normal (Stanek et al. 2010, Fabjan et al. 2011). However, projection effects due to correlated structures and diffuse large-scale structure have been identified as an non-negligable source of scatter and bias in the mass scaling relation. The non-lognormal, positively skewed distribution of scatter in projected Compton Y parameter in our cluster sample is in good agreement with the results of Hallman et al. (2007) and Yang, Bhattacharya and Ricker (2010), who analyzed light cone/cylindrical projections of the SZ effect, respectively. Based on an Edgeworth expansion of the mass–observable distribution, Shaw, Holder and Dudley (2010) find that the higher-order moments do not significantly impact the observed cluster mass function if the product of the scatter in the scaling relation, σ_M, and the slope of the mass function at the limiting mass of a survey is less than unity. Due to low scatter of the SZ scaling relation, this criterion is met by all upcoming SZ experiments, suggesting that projection effects will be insignificant for cosmological constrains (but see Shaw, Holder and Bode 2008, Erickson, Cunha and Evrard 2011, for additional mitigation strategies).

6.3.2 Influence of Halo Concentration

The scatter in halo concentration at fixed cluster mass has been identified as an important source of scatter in X-ray temperature (Yang, Ricker and Sutter 2009, Ameglio et al. 2009) and SZ signal (Shaw, Holder and Bode 2008, Yang, Bhattacharya and Ricker 2010) of simulated clusters. Understanding the role of halo concentration on these observables is especially important for understanding selection biases and for the comparison to lensing derived cluster masses.

The right panel of figure 6.4 shows the correlation between scatter in halo concentration at fixed mass and scatter in $\lg Y_{200}$ at fixed mass for all clusters in sample B. We use the halo concentration measurements from Ameglio et al. (2009) derived from fitting NFW-profiles to the integrated mass profile over the range $0.05 < r/R_{\mathrm{vir}} < 1$, and model concentration $c(M_{200})$ with a power law in mass. The scatter is positively correlated with more concentrated clusters having higher SZ signals at fixed mass, with a correlation coefficient of 0.30 for the full sample B and 0.68 for the massive subsample B*. This result is in agreement with the positive correlation between scatter in concentration and spectroscopic-like temperature of these clusters reported in Ameglio et al. (2009). Similarly, Shaw, Holder and Bode (2008) find a positive correlation between scatter in concentration and integrated Y-parameter in halos from adiabatic SPH simulations and from N-body simulation in combination with semi-analytic gas models. On the other hand, Yang, Ricker and Sutter (2009), Yang, Bhattacharya and Ricker (2010) find a negative correlation between scatter in concentration[4] and scatter in temperature and integrated SZ signal. As discussed in Yang, Bhattacharya and Ricker (2010), the correlation between halo concentration and temperature at fixed mass depends on the assumed gas physics and the inclusion of radiative cooling, star formation and feedback may change the sign of the correlation.

On the observational side, Comerford, Moustakas and Natarajan (2010) find ΔT anticorrelated with Δc. However this analysis is based on a sample of 8 strong lensing clusters and the authors note that this result vanishes if a different measurement for the concentration of one cluster (MS 2137.3-2353) is used. As strong lensing selected cluster samples are strongly affected by projection effects and are biased toward higher halo concentrations and X-ray luminosities than average clusters (e.g.,

---

[4] These authors use $\lg(R_{200}/R_{500})$ as a proxy for concentration, which for an NFW profile is a monotonically decreasing function to halo concentration. We find correlation coefficients of -0.22 (-0.47) for the scatter in $\lg(R_{200}/R_{500})$ and Y_{200} at fixed mass for sample B (B*), indicating that our result is robust with respect to the definition of halo concentration employed.

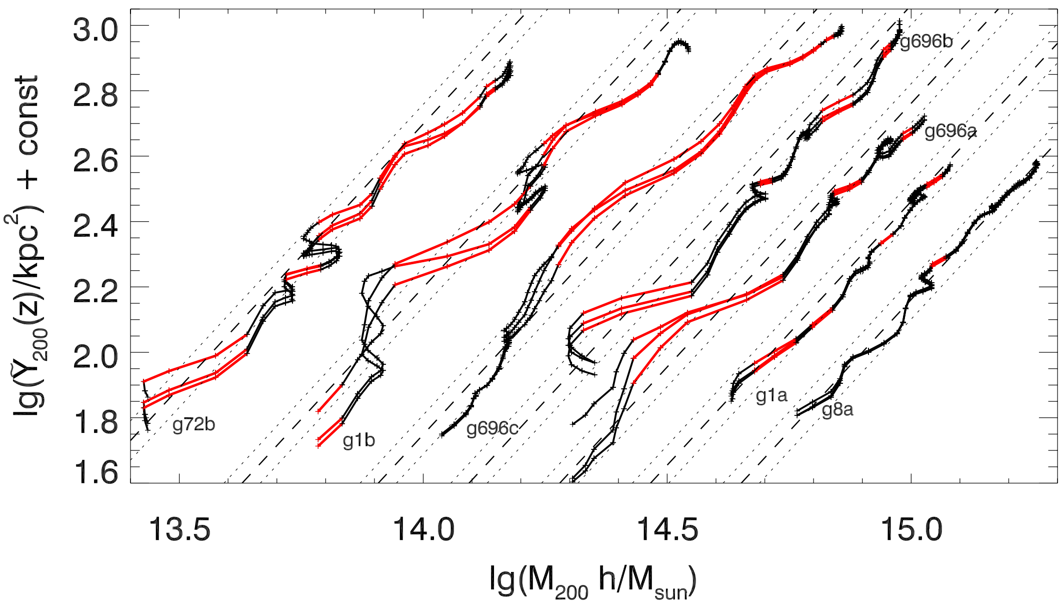


Figure 6.5 Evolution of six massive clusters in mass and $\tilde{Y}_{200}$, the redshift evolution scaled Y_{200}. Offsets are added to show all clusters in one plot. We show three orthogonal projections for each cluster to illustrate the magnitude of projection effects. Phases identified as merging events are shown in red. The dashed and dotted lines show the best-fit scaling relation for sample A and its 1σ error.

Meneghetti et al. 2010, 2011), larger, X-ray selected data sets like the CLASH survey (Postman et al. 2011) will be needed to observationally constrain the the correlation between scatter in temperature and halo concentration.

The scatter in halo concentration at fixed mass is linked to the formation epoch of a halo with more concentrated halos forming earlier (Navarro, Frenk and White 1997), albeit with large scatter (e.g., Neto et al. 2007) which is likely due to enviromental effects (see also Gao and White 2007). Hence the positive correlation between scatter in concentration and SZ signal suggests that clusters with Y biased low formed more recently.

6.4 Scatter Induced by Mergers

We now turn to a detailed analysis of the evolution of merging clusters around the $M(Y)$ scaling relation fit to sample A. Figure 6.5 shows the trajectory of six massive clusters around the best-fit scaling relation in the $M_{200} - \tilde{Y}_{200}$ plane. Phases identified as mergers are shown in red. These examples suggest that the SZ signal lags behind the change in mass during extended merger events moving the merging clusters below the best-fit scaling relation. This is similar to the findings of

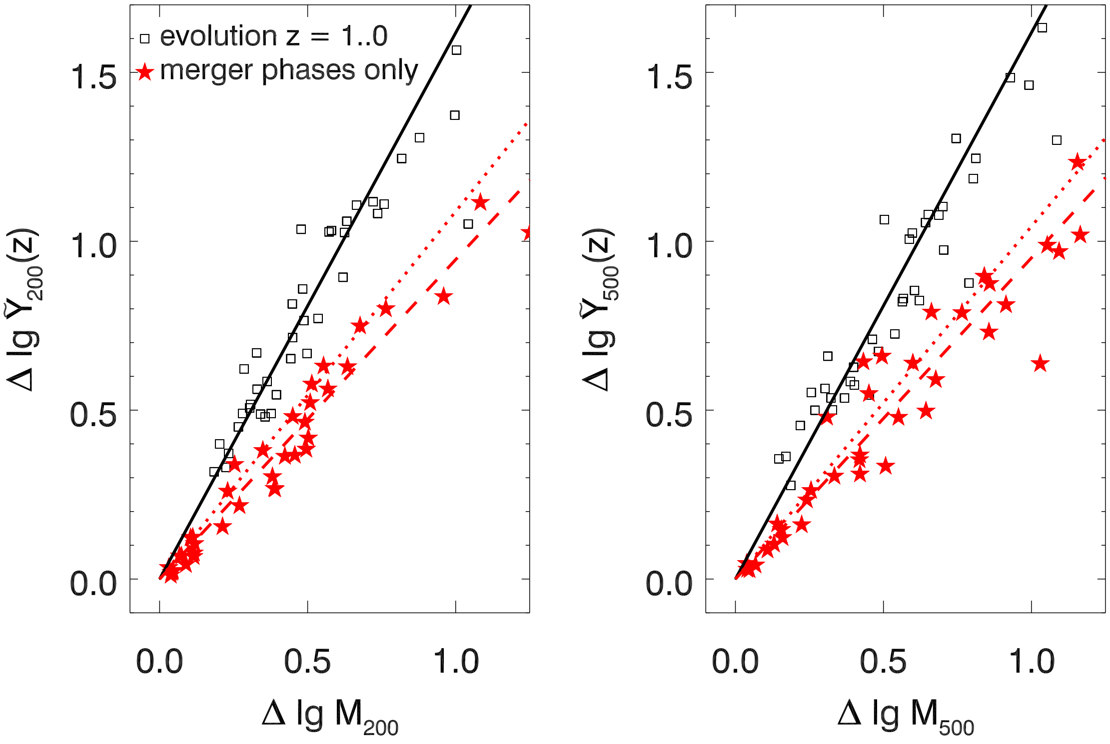


Figure 6.6 Logarithmic mass growth and increase in SZ signal scaled for cosmological evolution for all clusters in sample A. The left panel shows the evolution within $\Delta = 200$, the right panel for $\Delta = 500$. The black open symbols show the overall evolution of individual clusters between $z = 1$ and $z = 0$, the black solid lines are the best linear fit with zero intercept to these points, yielding a slope of 1.62 ± 0.19 at $\Delta = 200$ (1.62 ± 0.29 at $\Delta = 500$), consistent with the slope of the best-fit scaling relation. Filled, red stars show the evolution of each cluster during merger phases, the dashed lines are the best linear fit with zero intercept to the evolution during mergers with slope 0.94 ± 0.15 (0.95 ± 0.22). The dotted lines show the best-fit slope for the evolution during mergers when the merger criterion is relaxed to times when the fractional accretion rate per unit redshift is larger than the mean fractional accretion rate per unit redshift.

Rasia et al. (2011) who analyzed the evolution of X-ray properties of two of these clusters (g8a and g1b) during mergers and find a time delay between mass increase and rise in temperature of order a few hundred megayears. We quantify the difference in evolution during mergers compared to the overall evolution of each cluster in the $M - Y$ plane in figure 6.6. The open symbols show the logarithmic increase in mass,

$$\Delta \lg M = \lg \left(\frac{M(z=0)}{M(z=1)} \right) \tag{6.11}$$

and SZ signal scaled for redshift evolution

$$\Delta \lg \tilde{Y} = \lg \left(\frac{\tilde{Y}(z=0)}{\tilde{Y}(z=1)} \right) \tag{6.12}$$

As expected, the overall evolution from $z = 1$ to $z = 0$ as quantified by the slope of the best-fit linear model with zero intercept is consistent with the slope of the best-fit scaling relation.

The filled star symbols show the evolution of each cluster in the $M - \tilde{Y}$ plane during merger phases only (this corresponds to the sum of the red line segments for each cluster in figure 6.5, treating the different projections separately). The dashed red lines indicate the best-fit slope for the relation between increase in mass and redshift scaled Y during mergers. This shows that the Y signal scaled for redshift evolution increases more slowly during mergers than expected from the overall scaling relation. The dashed lines show the best-fit slope for the relation between increase in mass and redshift scaled Y during mergers when relaxing the merger criterion to include all times at which the fractional accretion rate is above its mean value. This illustrates that the suppression of Y during mergers is robust with respect to the definition of merger event.

We further illustrate the connection between merging events and scatter in the $M_{200}(Y_{200})$ scaling relation in figure 6.7. The top left panel shows how the clusters evolve around the scaling relation, giving the cumulative fraction of clusters evolving into outliers as a function of time, averaged over all clusters and all snapshots. Thick (thin) dashed-dotted or dashed lines show the fraction of clusters which evolve at least 10% (20%) below or above the scaling relation. For example, starting from one simulation snapshot, about 38% of all clusters will move at least 10% below the scaling relation within the next seven snapshots (corresponding to about one gigayear), about 30% deviate at least 10% above the scaling relation during that time period, and about 35% stay within 10% scatter from the scaling relation. The asymmetry between these pairs of lines is due to the non-lognormal distribution of scatter, the thick lines correspond to the 24% and 80% quantile, the thin lines correspond to the 4% and 90% quantile. The top right panel shows the same evolution around

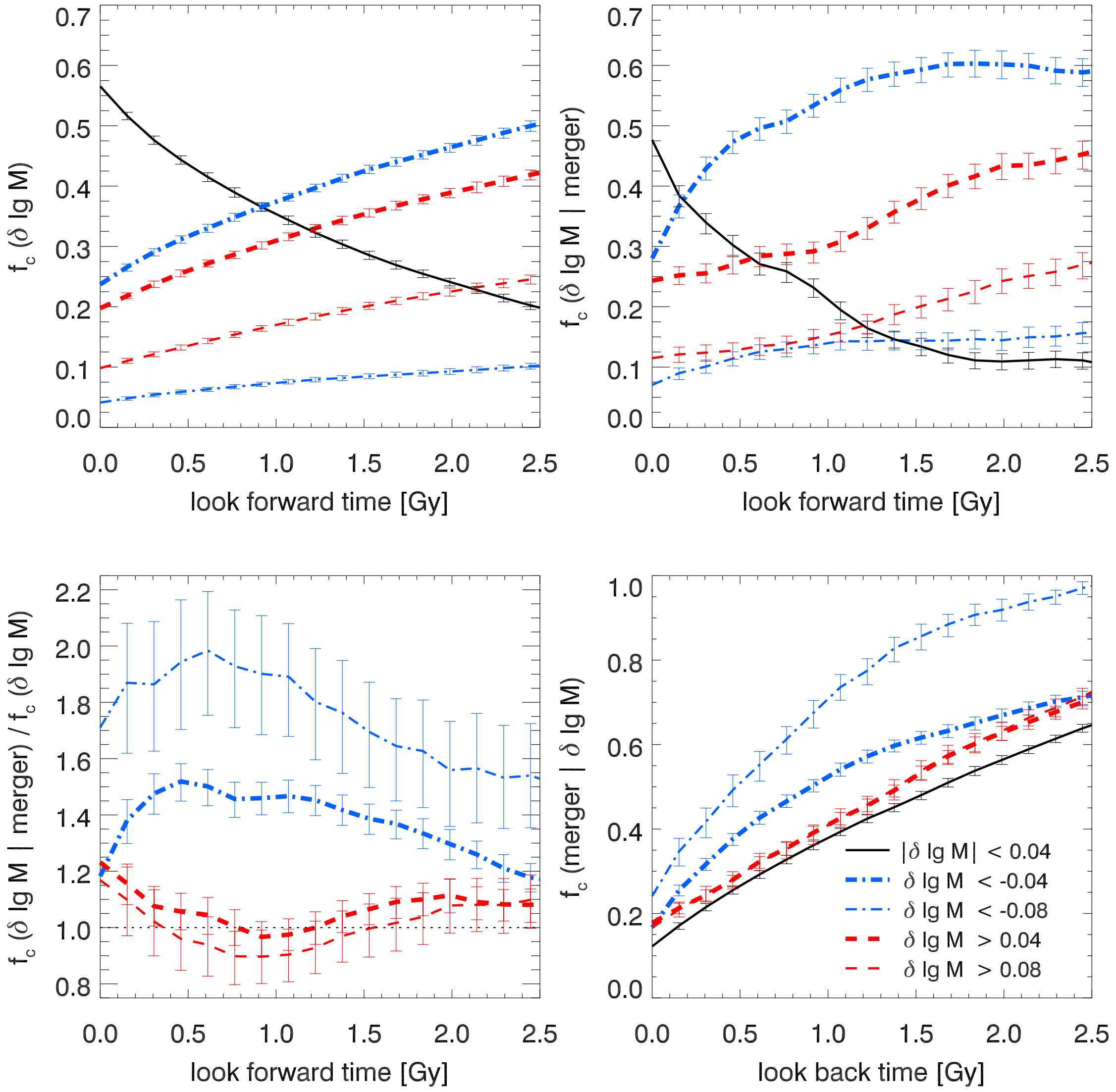


Figure 6.7 Top left: Cumulative probability for a cluster to deviate from the scaling relation by $\delta \lg M$ as a function of time. Thick (thin) dash-dotted blue lines show the fraction of clusters deviating at least 0.04 (0.08) below the scaling relation, corresponding to a bias of 10% (20%) in the inferred mass. Thick (thin) dashed lines show the fraction of clusters deviating at least 0.04 (0.08) above the scaling relation. The black solid line show the fraction of cluster which deviate less that 10% from the scaling relation within a given time. In all panels error bars indicate statistical errors estimated from 100 bootstrap realizations.

Top right: The same for merging clusters. Note that extended merging events are counted as multiple mergers, effectively giving more weight to major mergers.

Bottom left: Ratio of the above panels, highlighting the enhanced probability for mergers to evolve below the scaling relation compared to an average cluster.

Bottom right: Cumulative fraction of clusters which have undergone a merger as a function of look back time and their current deviation from the scaling relation.

the scaling scaling for clusters undergoing a merger at $t = 0$. Within a gigayear after a merger, 55% of all clusters will go through a phase where the inferred mass is biased low by at least 10%, while for only 30% of these cluster the inferred mass will be biased high by more than 10% during this time. The bottom left panel shows the ratio of these two plots, and illustrates the asymmetric evolution of mergers below the scaling relation. The inferred mass of a recently merged cluster is about 50% more likely to be biased low by at least 10% and twice as likely to be biased low by at least 20% compared to an average cluster.

The bottom right panel shows the cumulative fraction of clusters which have undergone a merger as a function of look back time given their current deviation from the scaling relation. This plot shows that 50% (75%) of all clusters with inferred masses biased low by at least 10% (20%) have undergone a merger within the last gigayear.

In summary our analysis shows that the SZ signal changes more slowly than cluster mass during mergers. This indicates that for a cosmological distribution of merger orbits and mass ratios, the delay between mass accretion and heating of the ICM by shocks and partial virialization are more important than merger boosts. Hence the inferred mass of recently merged clusters tends to be biased low, and we find that a large fraction of negative outliers are associated with recent mergers.

Note that throughout this section we have analyzed deviations from a scaling relation determined from a fit to sample A. Since the merger histories of this environment selected sample are not necessarily representative of a volume-limited sample, the calibration of this relation may be biased. However, the results in this section and the correlation between scatter in halo concentration and SZ signal of the volume-limited sample discussed in section 6.3.2 suggest that this bias would increase the normalization B and slope γ at fixed Y. Hence such a calibration bias would downplay the asymmetric scatter induced by mergers that we reported in this section. This suggests that in a volume-limited sample merging clusters may be less frequent, but their inferred masses could be more biased.

6.5 SZ Morphologies

Since we found the dynamical state of clusters to be correlated with scatter in the $M(Y)$ scaling relation we now test if the morphological appearance of SZ maps can be used to identify clusters that deviate from the scaling relation. Quantitative measures of the X-ray surface brightness morphology are commonly used to identify disturbed clusters, observations (e.g., Böhringer et al. 2010, Okabe

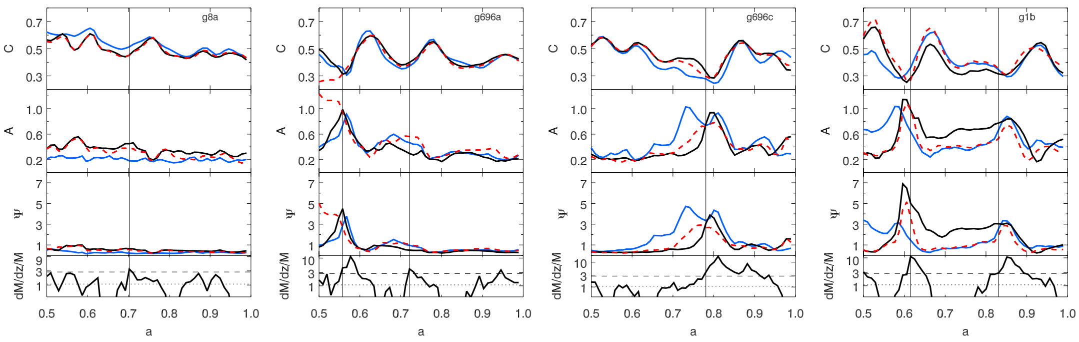


Figure 6.8 Evolution of morphological parameters G, A, Ψ for four massive clusters from sample A, different lines in each panel show the three orthogonal projections. The bottom panel shows the fractional accretion rate on a logarithmic scale, the dotted and dashed line indicate the mean accretion and the accretion rate threshold used to define mergers through out this analysis. Vertical lines mark the onset of mergers, i.e., the time when the fractional accretion rate first crosses the threshold used to define mergers. At the onset of a merger clusters appear less concentrated, more asymmetric and show more substructure.

et al. 2010, Marrone et al. 2011) and simulations (Jeltema et al. 2008, Ventimiglia et al. 2008, Böhringer et al. 2010) find the inferred masses of morphologically disturbed clusters to be biased low. Ventimiglia et al. (2008) analyzed the morphology of clusters from the simulation of Borgani et al. (2004), which is our sample B, and find significant correlations between the centroid shift, axial ratio and power ratios of the X-ray surface brightness distribution of these clusters and scatter in the $T_X(M)$ relation. Böhringer et al. (2010) compared the morphology of these simulated clusters to observed morphologies in the REXCESS sample, and show that the simulated X-ray morphologies show a larger dynamic range and appear more disturbed during mergers. They trace this difference to the fact that cool cores are more pronounced in this simulation.

Here we test the effectiveness of a number of morphological parameters, which are typically used to measure X-ray morphology of clusters or optical morphology of galaxies, at quantifying substructure in projected y maps. Within a circular aperture of radius R_{200} we compute the following quantities:

- *Asymmetry A* measures substructures and differences from circular symmetry, it is defined as the normalized difference between an image I and a copy R of the image rotated by 180 degree, $A = \sum_i |I_i - R_i| / \sum_i I_i$, where sum runs over all pixels in the aperture, and the center of the aperture is chosen to minimize A (Conselice 2003)

- *Centroid shift w* (Mohr et al. 1995) is another measure of the distribution of bright substructures based on the change of the centroid of different isophotal (iso-y) contours. Specifically, we follow the implementation of Ventimiglia et al. (2008) and compute the variance of the centroid for 10 iso-y contours spaced evenly in $\lg y$ between the maximum and minimum of y within the aperture.

- *Concentration C* We quantify the apparent concentration of the y distribution by the fraction of integrated Y contained within $0.3 \times R_{200}$, $C = Y_{0.3R_{200}} / Y_{200}$
- *Ellipticity* $\epsilon = 1 - B/A$ is defined as the ratio of semimajor (A) and semiminor axis (B) and is calculated directly from the second-order moments of the y distribution (Hashimoto et al. 2007)
- *Gini coefficient G* measures the uniformness of pixel values regardless of their spatial distribution (Lotz, Primack and Madau 2004). It is based on the Lorentz curve, the rank–ordered cumulative distribution of pixel values. It is defined as

  $$G = \frac{1}{2\bar{y}n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j| \,, \tag{6.13}$$

  where n is the number of pixels inside the aperture, y_i the value of the ith pixel, and $\bar{y}$ is the mean pixel value. The Gini coefficient of a uniform distribution is zero, and it is one if one pixel contains all the signal. It increases with the fraction of y in compact components.
- *Second-order brightness moment M_{20}* (Lotz, Primack and Madau 2004): The total second-order moment M is the signal in each pixel y_i weighted by the squared distance to the center of the galaxy cluster $(x_{1,c}, x_{2,c})$, summed over all pixel inside the aperture:

  $$M = \sum_{i}^{n} M_i = \sum_{i}^{n} y_i \left((x_{1,i} - x_{1,c})^2 + (x_{2,i} - x_{2,c})^2 \right) \,. \tag{6.14}$$

  Again, the center is determined by finding $(x_{1,c}, x_{2,c})$ that minimizes M. The second-order moment of the brightest regions measures the spatial distribution of bright subclumps. M_{20} is defined as the normalized second-order moment of the brightest 20% of the cluster's flux. M_{20} is computed from the pixels rank ordered by y,

  $$M_{20} = \log\left(\frac{\sum_i M_i}{M}\right) \quad \text{while} \quad \sum_i y_i < 0.2 Y_{200} \,. \tag{6.15}$$

  M_{20} is similar to C, but it is more sensitive to the spatial distribution of luminous regions and is not based on any symmetry assumptions.

- *Multiplicity* Ψ (Law et al. 2007) is another measure of the amount (multiplicity) of bright substructures. Using the observed y distribution as a tracer of the cluster's projected mass, one can calculate a "potential energy" of the y distribution,

  $$\Psi_{\mathrm{actual}} = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{y_i y_j}{r_{ij}}\,, \tag{6.16}$$

  where r_{ij} is the distance between pixels i and j. This value is normalized by the most compact possible re–arrangement of the pixel values, i.e., a circular configuration with pixel values decreasing with radius. The "potential energy" of this most compact light distribution is

  $$\Psi_{\mathrm{compact}} = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{y_i y_j}{r'_{ij}}\,, \tag{6.17}$$

  where r'_{ij} is the distance between pixels i and j in the most compact configuration. The multiplicity coefficient is defined as

  $$\Psi = 100 \times \log\left(\frac{\Psi_{\mathrm{compact}}}{\Psi_{\mathrm{actual}}}\right)\,. \tag{6.18}$$

  It is similar to A and M_{20}, but is has a larger dynamical range than M_{20} and requires no center or symmetry assumption.

- *Power ratio P_n* (Buote and Tsai 1995) correspond to a multipole expansion of the y map inside an aperture centered on the y centroid. We measure the power ratio P_2/P_0 which is related to the projected cluster ellipticity.

We measure morphology at a fixed physical resolution of 17.6 kpc/pixel and do not include any noise or observational effects.

Figure 6.8 shows the morphology as measured by C, A, and Ψ of four massive clusters from simulation A during their evolution since $a = 0.5$. The evolution of these clusters around the $M(Y)$ scaling relation is shown in figure 6.5. Vertical lines indicate the onset of mergers. Clusters g696a, g696c, and g1b illustrate the expected course of a merger: As a merging object enters the aperture within which morphologies are computed, the clusters appear less symmetric (higher A), less concentrated (lower C) and shows more substructure (higher Ψ). As the infalling clump sinks toward the cluster center and dissolves, the cluster appears less disturbed again. However, linking accretion history to morphology is complicated by extended merger phases (g696c, g1b at $a > 0.8$) with multiple infalling clumps. It is also apparent from these examples that fluctuation in morphology are not always linked to major accretion events (e.g., g8a, late time evolution of g696a).

For a more representative distribution of dynamical states and morphologies, we show the distribution of scatter in the $M(Y)$ relation and morphological parameters for all clusters in sample B in figure 6.9. Shaded region contain the 25% most disturbed/most elongated/least concentrated clusters. Overall, the inferred mass $M(Y)$ has larger scatter for clusters with disturbed morphologies, but it is nearly unbiased. Splitting the cluster sample by mass shows that morphologically disturbed clusters with low mass ($M_{200} < 10^{14} M_\odot/h$, open star symbols) tend to be biased toward larger inferred masses, while massive clusters ($M_{200} > 2 \times 10^{14} M_\odot/h$, filled red triangles) with disturbed morphologies are preferentially biased low in inferred mass. We quantify this trend using the Spearman rank order correlation coefficient for different mass samples and show the correlation coefficients in figure 6.9. If the significance level s of a correlation between a morphology parameter and mass bias is low ($s > 0.01$), we do not list a correlation coefficient. We find a significant correlation between morphology and mass bias in all three mass bins ($M > 2 \times 10^{14} M_\odot/h, M > 10^{14} M_\odot/h, M < 10^{14} M_\odot/h$) for the multiplicity, concentration, M_{20} and asymmetry parameter. These different morphology parameters consistently show that the correlation between disturbed morphology and negative mass bias increases with mass threshold, and the correlation coefficient changes sign for the low mass clusters. For centroid shifts and the Gini coefficient, we only find significant correlations with scatter in the $M(Y)$ relation in two mass bins, which follow the same pattern as just described. Power ratio P_2/P_0 and ellipticity are correlated with mass bias only for the most massive clusters, such that less circular clusters tend to be biased low in mass.

This segregation in mass, which is consistent among all morphological parameters, suggests that a large fraction of morphologically disturbed clusters which are biased high in inferred mass is caused by projection effects. The more massive clusters, which are less affected by projection effects, show correlations with disturbed morphology corresponding to a negative bias in inferred mass as expected from X-ray results. We expect cool cores to have a smaller influence on the SZ morphology than is found in X-ray, as the SZ signal is linear in density and less sensitive to physics in the cluster core. Projection effects due to uncorrelated large-scale structure along the line of sight are on average more diffuse than the projection effects from nearby structure that is included in our analysis. Hence we do not expect the morphology of massive clusters to become dominated by projection effects for line of sight projections which include all intervening structure.

As a first step toward including resolution effects, we convolve all projected y maps with a circular Gaussian beam with full width at half maximum (FWHM) of 150 kpc, and sample the maps at a resolution of four pixels per FWHM. For a telescope with a 1 arc minute beam, this physical resolution is reached for a source at $z \sim 0.15$; for an experiment with beam width of about 20 arc seconds, this corresponds to $z \sim 0.8$. Figure 6.10 shows the correlation between mass bias and cluster morphology as measured from these blurred maps for all massive clusters with $M > 2 \times 10^{14} M_\odot/h$ from sample B. For this choice of beam and pixel scale, cluster morphology and bias in inferred mass are well correlated and resolution effects are small. However, since this analysis is based on noise- and background-free y maps and a simplistic map making procedure, more realistic simulations are required to assess whether SZ based morphology can in practice be used as a proxy for the dynamical state of a cluster.

6.6 Summary and Discussion

Using projected Compton y maps of galaxy clusters extracted from cosmological hydrodynamical simulations, we analyze the clusters' thermal SZ signal and its scaling relation with cluster mass. We study the detailed time evolution of a sample of 39 clusters around the scaling relation using simulations with outputs closely spaced in time. Compared to previous studies, which focused either on the evolution of isolated, idealized mergers or on large samples of clusters at widely spaced redshifts, this sample enables us to isolate the effect of merging events for a cosmologically representative distribution of merger orbits, mass ratios, and impact parameters. Our main results can be summarized as follows:

1. The best-fit scaling relations to the integrated Y_{200} signal of these clusters are close to self-similar predictions and agree well with other simulations that include comparable gas physics.

2. The scatter around these scaling relations is small (of order 10% scatter in mass at fixed Y_{200}) and it is overall well correlated with the scatter in halo concentration, such that more concentrated halos have larger Y signal at fixed mass.

3. The scatter in the scaling relation deviates from a log normal distribution and is skewed toward clusters with Y signals larger than expected from their mass. We find projection effects due

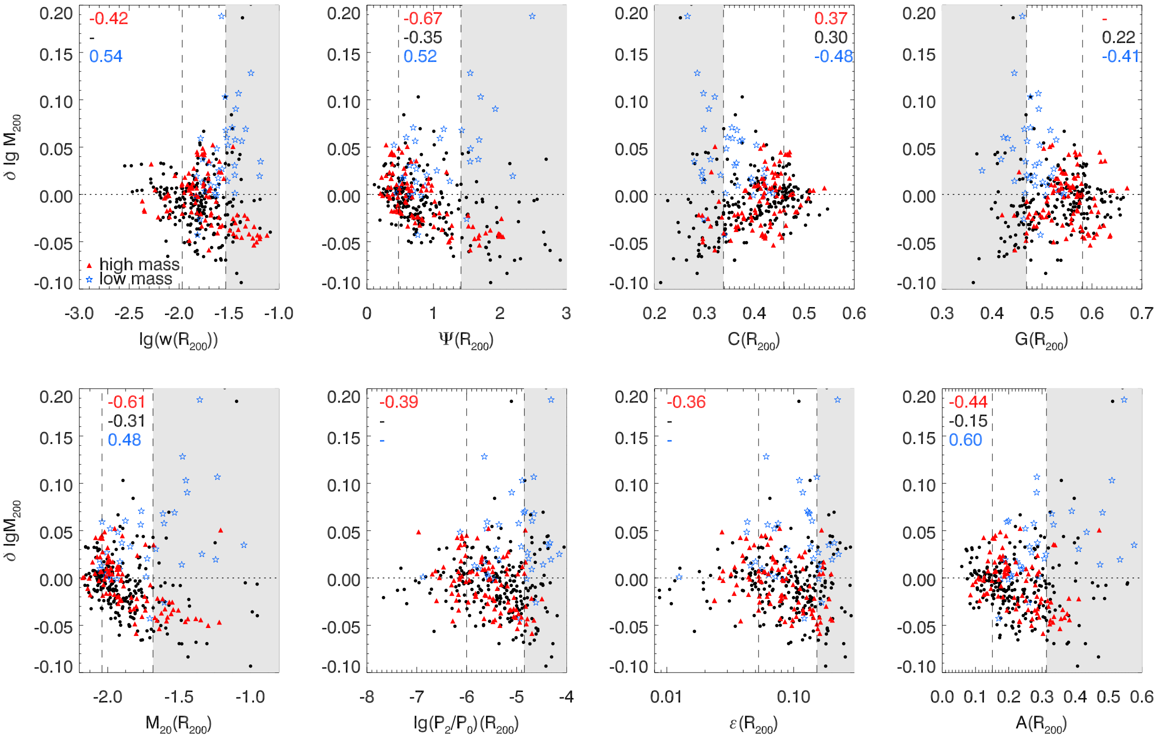


Figure 6.9 Relation between scatter in the $M_{200}(Y_{200})$ relation $\delta \lg M_{200}$ and morphological parameters for all clusters from sample B measured within an aperture of size R_{200}. Open star symbols show clusters with $M < 10^{14} M_{\odot}/h$, filled circles how clusters with $10^{14} M_{\odot}/h < M < 2\times10^{14} M_{\odot}/h$, and filled triangles show massive clusters with $M > 2 \times 10^{14} M_{\odot}/h$. Dashed vertical lines indicated the 25% and 75% quantiles of the morphology distribution. Shaded regions contain the 25% of the data points which are classified as most disturbed by that morphological parameter. Numbers in the upper left or right corner give the Spearman rank correlation coefficient between the morphological parameter and scatter in the $M(Y)$ relation. From top to bottom these numbers are for mass samples $M > 2 \times 10^{14} M_{\odot}/h, M > 10^{14} M_{\odot}/h, M < 10^{14} M_{\odot}/h$. If a correlation is not significant (significance level > 0.01), we do not list the correlation coefficient.

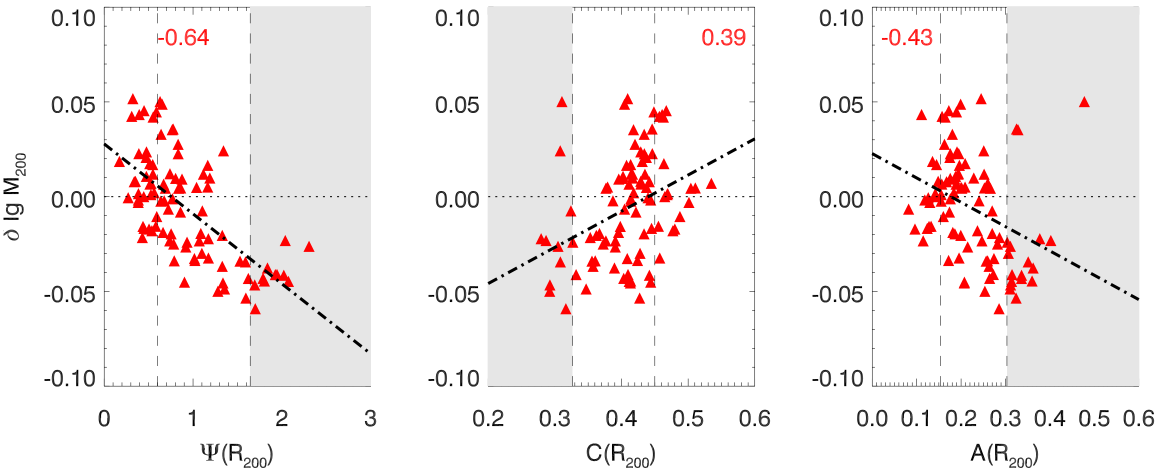


Figure 6.10 Relation between scatter in the $M_{200}(Y_{200})$ relation $\delta \lg M_{200}$ and morphological parameters for clusters with $M > 2\times10^{14} M_\odot/h$ from sample B, measured from SZ maps smoothed with a Gaussian beam with a FWHM of 150 kpc and sampled at a pixel scale of 37.5 kpc. Dashed vertical lines indicated the 25% and 75% quantiles of the morphology distribution. Shaded regions contain the 25% of the data points which are classified as most disturbed by that morphological parameter. Numbers in the upper left or right corner give the Spearman rank correlation coefficient between the morphological parameter and scatter in the $M(Y)$ relation. Dashed-dotted lines show the best-fit linear relation.

to nearby structures to be an important source of this upward scatter. However, due to the small magnitude of the scatter in the mass scaling, projection effects are not expected to be a significant contamination for cosmological constraints from SZ cluster surveys.

4. Merging clusters fall below the scaling relation, such that their inferred masses are biased low. More quantitatively, we find that within a Gyr following a merger, clusters are twice as likely as the average cluster to undergo a phase during which their inferred mass is biased low by more than 10%.

5. We identify merging events to be a major source of downward scatter in the scaling relation: a large fraction of clusters whose inferred masses are biased low recently underwent a merger (c.f. figure 6.7).

6. For massive clusters, we find the morphology of SZ maps to be well correlated with deviations from the scaling relation. While the robustness of this result with respect to noise and imaging artifacts requires further analysis, it suggests that SZ morphology may be useful to reduce the scatter of mass estimates, and to infer merger rates of massive halos and hence test theories

of halo formation.

Our analysis of the time evolution of merging events is in agreement with the conclusions drawn from earlier studies comparing morphologically disturbed and undisturbed clusters in cosmological simulations at fixed redshifts (e.g., Mathiesen and Evrard 2001, Kravtsov, Vikhlinin and Nagai 2006, Nagai 2006, Jeltema et al. 2008, Ventimiglia et al. 2008). Specifically, it supports the hypothesis that for a cosmological distribution of merger parameters partial virialization and non-thermal pressure support due to mergers are more important than merger boosts found in simulations of direct collisions between mergers. For simulated clusters the intrinsic scatter in the scaling relation and the mass segregation between morphologically relaxed and disturbed clusters are significantly smaller than recent observational results based on SZ measurements, X-ray morphology and weak lensing inferred masses (Marrone et al. 2011). However, as these authors note, the observed scatter is in agreement with the scatter expected in weak lensing mass measurements (Becker and Kravtsov 2010). Similarly, the mass segregation is enhanced by the sensitivity of weak lensing mass estimates to cluster triaxiality, and these observational constraints on the intrinsic scatter and bias in SZ mass estimates are limited by the accuracy of weak lensing mass reconstruction.

Further complications arise when inferring cluster masses from SZ observations as most Y measurements are derived from fitting parametric profiles (e.g., Nagai, Kravtsov and Vikhlinin 2007, Arnaud et al. 2010) to the data which assume radial symmetry (but see Plagge et al. 2010, Marrone et al. 2011, Sayers et al. 2011, for alternate methods and discussions).The distorted geometry of merging clusters may introduce additional scatter to mass estimates derived from profile fits, but an experiment specific analysis of such effects is beyond the scope of this work.

An additional limitation of our analysis is the range of non-gravitational physics included in the simulations. While recent studies show the impact of AGN-feedback on overall cluster profiles and scaling relations (Sijacki et al. 2007, Puchwein, Sijacki and Springel 2008, Battaglia et al. 2010, Fabjan et al. 2011), this mainly affects the cluster center. Consequently, we do not expect AGN feedback to significantly alter the slow virialization of newly accreted material at larger radii, which we found to be the main source of scatter during merging events. In the cluster outskirts, electrons and ions are not in thermal equilibrium. Rudd and Nagai (2009) and Wong and Sarazin (2009) show that detailed treatment of the multi-temperature structure of the intracluster medium leads to a significant suppression of electron temperature and SZ signal. Based on a sample of three simulated cluster, Rudd and Nagai (2009) find this effect to be especially pronounced in clusters undergoing major mergers. Under specific conditions, this effect may cause a bias of up to 5% in integrated Y, corresponding to an additional negative bias of about 3% in the inferred mass of merging clusters.

Overall, we find that merger events cause a temporary negative bias in inferred cluster mass of order 10%-15%. Due to the increased fraction of recently merged objects at higher redshift, we conclude that this merger bias should be accounted for when modeling the redshift evolution in the scatter of scaling relations.

Acknowledgements

We thank Silvia Ameglio for providing SZ maps and halo concentration measurements for cluster sample B, and Nick Battaglia, Chris Hirata, and James Taylor for useful discussions. The JobRunner web application was constructed by Laurent Bourges and Gerard Lemson as part of the activities of the German Astrophysical Virtual Observatory.

Appendix A

Halo Model Trispectrum

The trispectrum $T(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)$ of the dark matter density contrast $\tilde{\delta}(\mathbf{k})$ is defined as

$$\left\langle \tilde{\delta}(\mathbf{k}_1)\tilde{\delta}(\mathbf{k}_2)\tilde{\delta}(\mathbf{k}_3)\tilde{\delta}(\mathbf{k}_4) \right\rangle_{\mathrm{c}} = (2\pi)^3 \delta_{\mathrm{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4) T(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \,. \tag{A.1}$$

We model the dark matter trispectrum using the halo approach (Seljak 2000, Cooray and Sheth 2002), which assumes that all matter is bound in virialized structures, which are assumed to be biased tracers of the density field. Then the statistics of the density field can be described by the dark matter distribution within halos on small scales, and is dominated by the clustering properties of halos and their abundance on large scales. In this model, the trispectrum splits into four terms, which describe the 4-point correlation within one halo (the *one-halo* term $T^{1\mathrm{h}}$), and between 2 to 4 halos (*two-, three-, four-halo* term)

$$T = T^{1\mathrm{h}} + \left(T^{2\mathrm{h}}_{22} + T^{2\mathrm{h}}_{13} \right) + T^{3\mathrm{h}} + T^{4\mathrm{h}} \,. \tag{A.2}$$

The *two-halo* term is split into two parts, representing correlations between two or three points in the first halo and two or one point in the second halo.

As halos are the building blocks of the density field in the halo approach, we need to choose models for their internal structure, abundance and clustering in order to build a model for the trispectrum. In the following we summarize the main ingredients of our implementation of the halo model convergence trispectrum following (Cooray and Hu 2001).

We assume the halo profiles to follow the NFW profile (Navarro, Frenk and White 1997),

$$\rho(r, c) = \frac{\Delta_{\mathrm{vir}} \bar{\rho} c^2}{3(\ln(1 + c) - c/(1 + c))} \frac{1}{r/r_{\mathrm{vir}}\,(1 + cr/r_{\mathrm{vir}})^2} \,, \tag{A.3}$$

where $\Delta_{\rm vir}$ and $\bar{\rho}$ are the density contrast and mean density of the universe at virilization, and $c(M, z)$ is the halo concentration, which we model using the Bullock et al. (2001) fitting formula. We model the halo abundance using the Sheth and Tormen (1999) mass function,

$$\frac{{\rm d}n}{{\rm d}M}{\rm d}M = \frac{\bar{\rho}}{M}f(\nu){\rm d}\nu = \frac{\bar{\rho}}{M}A\left[1 + (a\nu)^{-p}\right]\sqrt{a\nu}\exp\left(-\frac{a\nu}{2}\right)\frac{{\rm d}\nu}{\nu}, \tag{A.4}$$

where A and p are fit parameters, and ν is the peak height $\nu = \delta_c/(D(z)\sigma(M))$. $\sigma(M)$ is the rms fluctuation of the present day matter density smoothed over a scale $R = (3M/4\pi\bar{\rho})^{1/3}$, and $D(z)$ is the growth factor. To describe the biased relation between the dark matter halo distribution and the density field, we assume a scale independent bias and use the fitting formula of Sheth and Tormen (1999),

$$b(\nu) = 1 + \frac{a\nu - 1}{\delta_{\rm c}} + \frac{2p}{\delta_{\rm c}(1 + (a\nu)^p)}, \tag{A.5}$$

and neglect higher-order bias functions (b_2, etc.). Following the notation of Cooray and Hu (2001) we introduce

$$I^{\beta}_{\mu}(k_1, \cdots, k_{\mu}; z) = \int {\rm d}M\frac{{\rm d}n}{{\rm d}M}\left(\frac{M}{\bar{\rho}}\right)^{\mu} b_{\beta}(M)\tilde{\rho}(k_1, c(M, z))\cdots\tilde{\rho}(k_{\mu}, c(M, z)) , \tag{A.6}$$

which describes the correlation of μ points within the same halo, and where $b_0 = 1$ and b_1 is given by (A.5). Then

$$T^{\rm 1h}(k_1, k_2, k_3, k_4; z) = I^0_4(k_1, k_2, k_3, k_4; z) \tag{A.7}$$

$$T^{\rm 2h}_{31}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4; z) = P_{\rm lin}(k_1)D(z)I^1_3(k_2, k_3, k_4; z)I^1_1(k_1; z) + 3 \text{ perm.} \tag{A.8}$$

$$T^{\rm 2h}_{22}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4; z) = P_{\rm lin}(k_{12})D(z)I^1_2(k_1, k_2; z)I^1_2(k_3, k_4; z) + 2 \text{ perm.} \tag{A.9}$$

$$T^{\rm 4h}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4; z) = T^{\rm pt}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4; z)I^1_1(k_1; z)\cdots I^1_1(k_4; z) , \tag{A.10}$$

where $\mathbf{k}_{ab} \equiv \mathbf{k}_a + \mathbf{k}_b$. We neglect the *3-halo* term, as it has negligible effect on our calculation, and simplify the *4-halo* term using just the trispectrum given by perturbation theory $T^{\rm pt}$ (Fry 1984).

Finally the tomographic convergence trispectrum can be written as

$$\begin{aligned}T_\kappa(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, -\mathbf{l}_{123}; z_\alpha, z_\beta, z_\gamma, z_\delta) &= l_1^2 l_2^2 l_3^2 l_{123}^2 \int \mathrm{d}\chi \frac{W(\chi, \chi_\alpha) W(\chi, \chi_\beta) W(\chi, \chi_\gamma) W(\chi, \chi_\delta)}{\chi^6} \\ &\quad \times T_\Phi(\mathbf{l}_1/\chi, \mathbf{l}_2/\chi, \mathbf{l}_3/\chi, -\mathbf{l}_{123}/\chi; z(\chi)) \\ &= \left(\frac{3}{2}\Omega_m H_0^2\right)^4 \int \mathrm{d}\chi \frac{W(\chi, \chi_\alpha) W(\chi, \chi_\beta) W(\chi, \chi_\gamma) W(\chi, \chi_\delta)}{\chi^6} (1 + z(\chi))^4 \\ &\quad \times T(\mathbf{l}_1/\chi, \mathbf{l}_2/\chi, \mathbf{l}_3/\chi, -\mathbf{l}_{123}/\chi; z(\chi)) \,, \qquad \text{(A.11)}\end{aligned}$$

where we have used the Poisson equation to relate the potential trispectrum to the matter density trispectrum.

Appendix B

Tree-level galaxy trispectrum

To calculate the tree-level matter trispectrum we need to consider the density contrast to third-order as the tree-level Trispectrum splits into two types of connected terms, $\left\langle\tilde{\delta}^{(1)}\tilde{\delta}^{(1)}\tilde{\delta}^{(2)}\tilde{\delta}^{(2)}\right\rangle_{\mathrm{c}}$ and $\left\langle\tilde{\delta}^{(1)}\tilde{\delta}^{(1)}\tilde{\delta}^{(1)}\tilde{\delta}^{(3)}\right\rangle_{\mathrm{c}}$. The third-order density contrast is given by (Fry 1984)

$$\tilde{\delta}^{(3)}(\mathbf{k}) = \int \frac{\mathrm{d}^3\mathbf{k}_1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{k}_2}{(2\pi)^3} F_3\left(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}-\mathbf{k}_1-\mathbf{k}_2\right) \tilde{\delta}^{(1)}(\mathbf{k}_1)\tilde{\delta}^{(1)}(\mathbf{k}_2)\tilde{\delta}^{(1)}(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2)\,, \tag{B.1}$$

with the third-order coupling function F_3. One finds for the matter trispectrum

$$\begin{aligned}(2\pi)^3\delta_{\mathrm{D}}(\mathbf{k}_{1234})T_{\mathrm{pt}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) \approx\; & \left\langle\tilde{\delta}^{(1)}(\mathbf{k}_1)\tilde{\delta}^{(1)}(\mathbf{k}_2)\tilde{\delta}^{(1)}(\mathbf{k}_3)\tilde{\delta}^{(3)}(\mathbf{k}_4)\right\rangle + 3\text{ perm.}\\ & + \left\langle\tilde{\delta}^{(1)}(\mathbf{k}_1)\tilde{\delta}^{(1)}(\mathbf{k}_2)\tilde{\delta}^{(2)}(\mathbf{k}_3)\tilde{\delta}^{(2)}(\mathbf{k}_4)\right\rangle + 5\text{ perm.}\end{aligned} \tag{B.2}$$

After some algebra one obtains

$$\begin{aligned}T_{\mathrm{pt}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) =\; & 6F_3^{\mathrm{s}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)P(k_1)P(k_2)P(k_3) + 3\text{ perm.}\\ & + 4\left[P(k_{13})F_2(\mathbf{k}_1,-\mathbf{k}_{13})F_2(\mathbf{k}_2,\mathbf{k}_{13}) + P(k_{23})F_2(\mathbf{k}_1,\mathbf{k}_{23})F_2(\mathbf{k}_2,-\mathbf{k}_{23})\right]\\ & \times P(k_1)P(k_2) + 5\text{ perm.}\end{aligned} \tag{B.3}$$

If one assume the third-order galaxy bias (b_3) to be zero, two types of additional terms containing the quadratic galaxy bias contribute to the galaxy trispectrum, $\left\langle b_1\tilde{\delta}^{(1)}\; b_1\tilde{\delta}^{(1)}\; b_1\tilde{\delta}^{(2)}\; b_2\tilde{\delta}^{(1)}\otimes\tilde{\delta}^{(1)}\right\rangle_{\mathrm{c}}$ and $\left\langle b_1\tilde{\delta}^{(1)}\; b_1\tilde{\delta}^{(1)}\; b_2\tilde{\delta}^{(1)}\otimes\tilde{\delta}^{(1)}\; b_2\tilde{\delta}^{(1)}\otimes\tilde{\delta}^{(1)}\right\rangle_{\mathrm{c}}$. Hence our model for the galaxy trispectrum is given by

$$\begin{aligned} T_{\rm gal}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) &\approx b_1^4 T_{\rm pt}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) \\ &+ 2b_1^3 b_2 P(k_1)P(k_2)\left[P(k_{13})F_2(\mathbf{k}_1,-\mathbf{k}_{13}) + P(k_{24})F_2(\mathbf{k}_2,-\mathbf{k}_{23})\right] + \text{ 5 perm.} \\ &+ 4b_1^2 b_2^2 P(k_1)P(k_2)\left[P(k_{13}) + P(k_{23})\right] + \text{ 5 perm.} \end{aligned} \quad \text{(B.4)}$$

Bibliography

Abazajian K. N. et al., 2009, *ApJS*, 182, 543

Albrecht A. et al., 2009, ArXiv 0901.0721

Allen S. W., Evrard A. E., Mantz A. B., 2011, ArXiv e-prints

Amara A., Réfrégier A., 2008, *MNRAS*, 391, 228

Ameglio S., Borgani S., Pierpaoli E., Dolag K., 2007, *MNRAS*, 382, 397

Ameglio S., Borgani S., Pierpaoli E., Dolag K., Ettori S., Morandi A., 2009, *MNRAS*, 394, 479

Andersson K. et al., 2011, *ApJ*, 738, 48

Angrick C., Bartelmann M., 2011, ArXiv e-prints

Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, *A&A*, 517, A92+

Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5

Balogh M. L., Navarro J. F., Morris S. L., 2000, *ApJ*, 540, 113

Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291

Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., Sijacki D., 2010, *ApJ*, 725, 91

Becker M. R., Kravtsov A. V., 2010, ArXiv e-prints

Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379

Benjamin J. et al., 2007, *MNRAS*, 381, 702

Benson A. J., 2005, *MNRAS*, 358, 551

Berlind A. A., Weinberg D. H., 2002, *ApJ*, 575, 587

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Phys. Rep.*, 367, 1

Bernardeau F., van Waerbeke L., Mellier Y., 1997, *A&A*, 322, 1

Bernstein G. M., Jarvis M., 2002, *AJ*, 123, 583

Bhattacharya S., Habib S., Heitmann K., 2011, ArXiv e-prints

Binggeli B., 1982, *A&A*, 107, 338

Blanton M. R. et al., 2003a, *AJ*, 125, 2348

Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003b, *AJ*, 125, 2276

Blanton M. R., Roweis S., 2007, *AJ*, 133, 734

Blanton M. R. et al., 2005, *AJ*, 129, 2562

Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, *ApJ*, 301, 27

Böhringer H. et al., 2010, *A&A*, 514, A32+

Borgani S. et al., 2004, *MNRAS*, 348, 1078

Bridle S. et al., 2009, ArXiv 0908.0945

Broadhurst T. J., Taylor A. N., Peacock J. A., 1995, *ApJ*, 438, 49

Buchalter A., Kamionkowski M., Jaffe A. H., 2000, *ApJ*, 530, 36

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, *MNRAS*, 321, 559

Bunn E. F., Zaldarriaga M., Tegmark M., de Oliveira-Costa A., 2003, *Phys. Rev. D*, 67, 023501

Buote D. A., Tsai J. C., 1995, *ApJ*, 452, 522

Catelan P., Kamionkowski M., Blandford R. D., 2001, *MNRAS*, 320, L7

Comerford J. M., Moustakas L. A., Natarajan P., 2010, *ApJ*, 715, 162

Conselice C. J., 2003, *ApJS*, 147, 1

Cooper M. C. et al., 2007, *MNRAS*, 376, 1445

Cooper M. C. et al., 2006, *MNRAS*, 370, 198

Cooray A., 2004, *MNRAS*, 348, 250

Cooray A., Hu W., 2001, *ApJ*, 554, 56

Cooray A., Hu W., 2002, *ApJ*, 574, 19

Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1

Coupon J. et al., 2011, ArXiv e-prints

Crittenden R. G., Natarajan P., Pen U., Theuns T., 2002, *ApJ*, 568, 20

Crittenden R. G., Natarajan P., Pen U.-L., Theuns T., 2001, *ApJ*, 559, 552

Davis M., Huchra J., Latham D. W., Tonry J., 1982, *ApJ*, 253, 423

Dodelson S., Shapiro C., White M., 2006, *Phys. Rev. D*, 73, 023009

Dodelson S., Zhang P., 2005, *Phys. Rev. D*, 72, 083001

Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497

Dolag K., Hansen F. K., Roncarelli M., Moscardini L., 2005, *MNRAS*, 363, 29

Dolag K., Meneghetti M., Moscardini L., Rasia E., Bonaldi A., 2006, *MNRAS*, 370, 656

Dolney D., Jain B., Takada M., 2006, *MNRAS*, 366, 884

Doroshkevich A. G., 1970, Astrophysics, 6, 320

Efstathiou G., Bond J. R., White S. D. M., 1992, *MNRAS*, 258, 1P

Eifler T., Schneider P., Krause E., 2010, *A&A*, 510, A7+

Erickson B. M. S., Cunha C. E., Evrard A. E., 2011, ArXiv e-prints

Faber S. M. et al., 2007, *ApJ*, 665, 265

Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, *MNRAS*, 416, 801

Fakhouri O., Ma C., 2008, *MNRAS*, 386, 577

Fakhouri O., Ma C., 2009, *MNRAS*, 394, 1825

Faltenbacher A., Li C., Mao S., van den Bosch F. C., Yang X., Jing Y. P., Pasquali A., Mo H. J., 2007, *ApJL*, 662, L71

Faltenbacher A., Li C., White S. D. M., Jing Y., Shu-DeMao, Wang J., 2009, Research in Astronomy and Astrophysics, 9, 41

Feldman H. A., Frieman J. A., Fry J. N., Scoccimarro R., 2001, Physical Review Letters, 86, 1434

Fry J. N., 1984, *ApJ*, 279, 499

Fry J. N., 1994, Physical Review Letters, 73, 215

Fry J. N., Gaztanaga E., 1993, *ApJ*, 413, 447

Fry J. N., Thomas D., 1999, *ApJ*, 524, 591

Fu L. et al., 2008, *A&A*, 479, 9

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748

Gao L., Springel V., White S. D. M., 2005, *MNRAS*, 363, L66

Gao L., White S. D. M., 2007, *MNRAS*, 377, L5

Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, *ApJ*, 616, 16

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759

Guo H., Jing Y. P., 2009a, *ApJ*, 698, 479

Guo H., Jing Y. P., 2009b, *ApJ*, 702, 425

Hallman E. J., O'Shea B. W., Burns J. O., Norman M. L., Harkness R., Wagner R., 2007, *ApJ*, 671, 27

Hamilton A. J. S., Tegmark M., 2004, *MNRAS*, 349, 115

Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399

Hashimoto Y., Böhringer H., Henry J. P., Hasinger G., Szokoly G., 2007, *A&A*, 467, 485

Hatton S., Cole S., 1998, *MNRAS*, 296, 10

Heavens A. F., Matarrese S., Verde L., 1998, *MNRAS*, 301, 797

Heinis S. et al., 2009, *ApJ*, 698, 1838

Heinis S. et al., 2007, *ApJS*, 173, 503

Heitmann K., Higdon D., White M., Habib S., Williams B. J., Wagner C., 2009, ArXiv 0902.0429

Heitmann K., White M., Wagner C., Habib S., Higdon D., 2008, ArXiv 0812.1052

Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, *A&A*, 499, 31

Hirata C. M., 2009, *MNRAS*, 399, 1074

Hirata C. M., Seljak U., 2003, *Phys. Rev. D*, 68, 083002

Hirata C. M., Seljak U., 2004, *Phys. Rev. D*, 70, 063526

Hoyle F., 1949, *MNRAS*, 109, 365

Huff E. M., Eifler T., Hirata C. M., Mandelbaum R., Schlegel D., Seljak U., 2011, ArXiv e-prints

Hui L., Gaztañaga E., Loverde M., 2007, *Phys. Rev. D*, 76, 103502

Hui L., Zhang J., 2008, *ApJ*, 688, 742

Huterer D., Takada M., Bernstein G., Jain B., 2006, *MNRAS*, 366, 101

Ivezic Z., Tyson J. A., Acosta E., Allsman R., Anderson S. F., et al., 2008, ArXiv e-prints

Jackson J. C., 1972, *MNRAS*, 156, 1P

Jain B., 2002, *ApJL*, 580, L3

Jarvis M., Bernstein G., Jain B., 2004, *MNRAS*, 352, 338

Jeltema T. E., Hallman E. J., Burns J. O., Motl P. M., 2008, *ApJ*, 681, 167

Jing Y. P., Börner G., 2004, *ApJ*, 607, 140

Jing Y. P., Zhang P., Lin W. P., Gao L., Springel V., 2006, *ApJL*, 640, L119

Joachimi B., Schneider P., 2008, *A&A*, 488, 829

Joachimi B., Schneider P., 2009, ArXiv 0905.0393

Joachimi B., Shi X., Schneider P., 2009, *A&A*, 508, 1193

Kaiser N., 1984, *ApJL*, 284, L9

Kaiser N., 1987, *MNRAS*, 227, 1

Kaiser N., 1992, *ApJ*, 388, 272

Kaiser N., Squires G., Broadhurst T., 1995, *ApJ*, 449, 460

Kauffmann G. et al., 2003, *MNRAS*, 341, 33

Kayo I. et al., 2004, *PASJ*, 56, 415

Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, 556, 121

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, *MNRAS*, 372, 961

Kilbinger M., Schneider P., Eifler T., 2006, *A&A*, 457, 15

King L. J., 2005, *A&A*, 441, 47

King L. J., Schneider P., 2003, *A&A*, 398, 23

Komatsu E. et al., 2009, *ApJS*, 180, 330

Komatsu E. et al., 2010, ArXiv e-prints

Kravtsov A. V., Vikhlinin A., Nagai D., 2006, *ApJ*, 650, 128

Krumpe M., Miyaji T., Coil A. L., 2010, *ApJ*, 713, 558

Kulkarni G. V., Nichol R. C., Sheth R. K., Seo H., Eisenstein D. J., Gray A., 2007, *MNRAS*, 378, 1196

Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64

Lau E. T., Kravtsov A. V., Nagai D., 2009, *ApJ*, 705, 1129

Laureijs R., Amiaux J., Arduini S., Auguères J. ., Brinchmann J., et al., 2011, ArXiv e-prints

Law D. R., Steidel C. C., Erb D. K., Pettini M., Reddy N. A., Shapley A. E., Adelberger K. L., Simenc D. J., 2007, *ApJ*, 656, 1

Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R., 2011, ArXiv e-prints

Lee J., Pen U.-L., 2000, *ApJ*, 532, L5

Lee J., Pen U.-L., 2007, *ApJL*, 670, L1

Levine R., Gnedin N. Y., 2006, *ApJL*, 649, L57

Li C., Kauffmann G., Wang L., White S. D. M., Heckman T. M., Jing Y. P., 2006, *MNRAS*, 373, 457

Lima M., Hu W., 2005, *Phys. Rev. D*, 72, 043006

Loh Y. et al., 2010, *MNRAS*, 407, 55

Lotz J. M., Primack J., Madau P., 2004, *AJ*, 128, 163

Ma Z., Hu W., Huterer D., 2006, *ApJ*, 636, 21

Majumdar S., Mohr J. J., 2004, *ApJ*, 613, 41

Mandelbaum R. et al., 2011, *MNRAS*, 410, 844

Mandelbaum R., Seljak U., Cool R. J., Blanton M., Hirata C. M., Brinkmann J., 2006, *MNRAS*, 372, 758

Manera M., Gaztanaga E., 2009, ArXiv e-prints

Marín F. A., Wechsler R. H., Frieman J. A., Nichol R. C., 2008, *ApJ*, 672, 849

Marriage T. A. et al., 2011, *ApJ*, 737, 61

Marrone D. P. et al., 2011, ArXiv e-prints

Martin D. C. et al., 2007, *ApJS*, 173, 342

Massey R. et al., 2007a, *MNRAS*, 376, 13

Massey R. et al., 2007b, *ApJS*, 172, 239

Mathiesen B. F., Evrard A. E., 2001, *ApJ*, 546, 100

McBride J., Fakhouri O., Ma C., 2009, *MNRAS*, 398, 1858

McDonald P., 2006, *Phys. Rev. D*, 74, 103512

McDonald P., Roy A., 2009, *JCAP*, 8, 20

Meneghetti M., Fedeli C., Pace F., Gottlöber S., Yepes G., 2010, *A&A*, 519, A90+

Meneghetti M., Fedeli C., Zitrin A., Bartelmann M., Broadhurst T., Gottlöber S., Moscardini L., Yepes G., 2011, *A&A*, 530, A17+

Miyaji T., Krumpe M., Coil A. L., Aceves H., 2011, *ApJ*, 726, 83

Mohr J. J., Evrard A. E., Fabricant D. G., Geller M. J., 1995, *ApJ*, 447, 8

Moran S. M., Ellis R. S., Treu T., Smith G. P., Rich R. M., Smail I., 2007, *ApJ*, 671, 1503

Nagai D., 2006, *ApJ*, 650, 538

Nagai D., Kravtsov A. V., Vikhlinin A., 2007, *ApJ*, 668, 1

Nagai D., Vikhlinin A., Kravtsov A. V., 2007, *ApJ*, 655, 98

Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493

Neto A. F. et al., 2007, *MNRAS*, 381, 1450

Nichol R. C. et al., 2006, *MNRAS*, 368, 1507

Niederste-Ostholt M., Strauss M. A., Dong F., Koester B. P., McKay T. A., 2010, ArXiv e-prints

Nock K., Percival W. J., Ross A. J., 2010, ArXiv e-prints

Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, *MNRAS*, 396, 19

Okabe N., Zhang Y.-Y., Finoguenov A., Takada M., Smith G. P., Umetsu K., Futamase T., 2010, *ApJ*, 721, 875

Padmanabhan N. et al., 2007, *MNRAS*, 378, 852

Peacock J. A. et al., 2001, *Nature*, 410, 169

Pedrosa S., Tissera P. B., Scannapieco C., 2009, *MNRAS*, 395, L57

Peebles P. J. E., 1969, *ApJ*, 155, 393

Peebles P. J. E., Hauser M. G., 1974, *ApJS*, 28, 19

Plagge T. et al., 2010, *ApJ*, 716, 1118

Planck Collaboration et al., 2011a, ArXiv e-prints

Planck Collaboration et al., 2011b, ArXiv e-prints

Planck Collaboration et al., 2011c, ArXiv e-prints

Poole G. B., Babul A., McCarthy I. G., Fardal M. A., Bildfell C. J., Quinn T., Mahdavi A., 2007, *MNRAS*, 380, 437

Poole G. B., Fardal M. A., Babul A., McCarthy I. G., Quinn T., Wadsley J., 2006, *MNRAS*, 373, 881

Postman M. et al., 2011, ArXiv e-prints

Puchwein E., Sijacki D., Springel V., 2008, *ApJL*, 687, L53

Randall S. W., Sarazin C. L., Ricker P. M., 2002, *ApJ*, 577, 579

Rasia E. et al., 2006, *MNRAS*, 369, 2013

Rasia E., Mazzotta P., Evrard A., Markevitch M., Dolag K., Meneghetti M., 2011, *ApJ*, 729, 45

Reid B. A., Spergel D. N., Bode P., 2009, *ApJ*, 702, 249

Rudd D. H., Nagai D., 2009, *ApJL*, 701, L16

Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, *ApJ*, 672, 19

Salim S. et al., 2007, *ApJS*, 173, 267

Sayers J., Golwala S. R., Ameglio S., Pierpaoli E., 2011, *ApJ*, 728, 39

Schäfer B. M., 2009, International Journal of Modern Physics D, 18, 173

Schmidt F., Rozo E., Dodelson S., Hui L., Sheldon E., 2009a, *ApJ*, 702, 593

Schmidt F., Rozo E., Dodelson S., Hui L., Sheldon E., 2009b, Physical Review Letters, 103, 051301

Schmidt F., Vallinotto A., Sefusatti E., Dodelson S., 2008, *Phys. Rev. D*, 78, 043513

Schneider P., 1996, *MNRAS*, 283, 837

Schneider P., Eifler T., Krause E., 2010, *A&A*, 520, A116+

Schneider P., Kilbinger M., 2007, *A&A*, 462, 841

Schneider P., Kilbinger M., Lombardi M., 2005, *A&A*, 431, 9

Schneider P., Lombardi M., 2003, *A&A*, 397, 809

Schneider P., Seitz C., 1995, *A&A*, 294, 411

Schneider P., van Waerbeke L., Jain B., Kruse G., 1998, *MNRAS*, 296, 873

Schneider P., van Waerbeke L., Mellier Y., 2002, *A&A*, 389, 729

Schrabback T. et al., 2007, *A&A*, 468, 823

Sciama D. W., 1955, *MNRAS*, 115, 3

Scoccimarro R., Couchman H. M. P., 2001, *MNRAS*, 325, 1312

Scoccimarro R., Couchman H. M. P., Frieman J. A., 1999, *ApJ*, 517, 531

Scoccimarro R., Feldman H. A., Fry J. N., Frieman J. A., 2001, *ApJ*, 546, 652

Sefusatti E., Crocce M., Pueblas S., Scoccimarro R., 2006, *Phys. Rev. D*, 74, 023522

Sehgal N., Bode P., Das S., Hernandez-Monteagudo C., Huffenberger K., Lin Y.-T., Ostriker J. P., Trac H., 2010, *ApJ*, 709, 920

Seibert M., et al, in preparation

Sehgal N. et al., 2011, *ApJ*, 732, 44

Seitz C., Schneider P., 1997, *A&A*, 318, 687

Seljak U., 2000, *MNRAS*, 318, 203

Seljak U., Zaldarriaga M., 1996, *ApJ*, 469, 437

Shapiro C., 2009, *ApJ*, 696, 775

Shapiro C., Cooray A., 2006, Journal of Cosmology and Astro-Particle Physics, 3, 7

Shaw L. D., Holder G. P., Bode P., 2008, *ApJ*, 686, 206

Shaw L. D., Holder G. P., Dudley J., 2010, *ApJ*, 716, 281

Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119

Shi X., Schneider P., Joachimi B., 2011, *A&A*, 533, A48

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, *MNRAS*, 380, 877

Smith G. P., Edge A. C., Eke V. R., Nichol R. C., Smail I., Kneib J.-P., 2003a, *ApJL*, 590, L79

Smith J. A. et al., 2002, *AJ*, 123, 2121

Smith R. E. et al., 2003b, *MNRAS*, 341, 1311

Smith R. E., Sheth R. K., Scoccimarro R., 2008, *Phys. Rev. D*, 78, 023523

Springel V., Di Matteo T., Hernquist L., 2005, *ApJL*, 620, L79

Springel V., Hernquist L., 2003, *MNRAS*, 339, 289

Springel V. et al., 2005, *Nature*, 435, 629

Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, *ApJ*, 715, 1508

Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, *MNRAS*, 387, 1391

Takada M., Jain B., 2003, *ApJL*, 583, L49

Takada M., Jain B., 2004, *MNRAS*, 348, 897

Tegmark M. et al., 2004, *ApJ*, 606, 702

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709

Tinker J., Wetzel A., Conroy C., 2011, ArXiv e-prints

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878

Tinker J. L., Weinberg D. H., Zheng Z., Zehavi I., 2005, *ApJ*, 631, 41

Tormen G., Bouchet F. R., White S. D. M., 1997, *MNRAS*, 286, 865

Treu T., Ellis R. S., Kneib J., Dressler A., Smail I., Czoske O., Oemler A., Natarajan P., 2003, *ApJ*, 591, 53

Valageas P., Sato M., Nishimichi T., 2011, ArXiv e-prints

van den Bosch F. C., Abel T., Croft R. A. C., Hernquist L., White S. D. M., 2002, *ApJ*, 576, 21

Vanderlinde K. et al., 2010, *ApJ*, 722, 1180

Ventimiglia D. A., Voit G. M., Donahue M., Ameglio S., 2008, *ApJ*, 685, 118

Verde L., Heavens A. F., Matarrese S., 2000, *MNRAS*, 318, 584

Verde L., Heavens A. F., Matarrese S., Moscardini L., 1998, *MNRAS*, 300, 747

Verde L. et al., 2002, *MNRAS*, 335, 432

Vishniac E. T., 1983, *MNRAS*, 203, 345

von der Linden A., Wild V., Kauffmann G., White S. D. M., Weinmann S., 2010, *MNRAS*, 404, 1231

Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, *ApJ*, 568, 52

Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, *ApJ*, 652, 71

Wetzel A. R., Tinker J. L., Conroy C., 2011, ArXiv e-prints

White M., 2002, *ApJS*, 143, 241

White S. D. M., 1984, *ApJ*, 286, 38

Wik D. R., Sarazin C. L., Ricker P. M., Randall S. W., 2008, *ApJ*, 680, 17

Williamson R. et al., 2011, *ApJ*, 738, 139

Wong K.-W., Sarazin C. L., 2009, *ApJ*, 707, 1141

Wyder T. K. et al., 2007, *ApJS*, 173, 293

Yang H., Bhattacharya S., Ricker P. M., 2010, *ApJ*, 725, 1124

Yang H., Ricker P. M., Sutter P. M., 2009, *ApJ*, 699, 315

York D. G. et al., 2000, *AJ*, 120, 1579

Yoshida N. et al., 2001, *MNRAS*, 325, 803

Zaldarriaga M., Scoccimarro R., 2003, *ApJ*, 584, 559

Zehavi I. et al., 2005, *ApJ*, 630, 1

Zehavi I. et al., 2010, ArXiv e-prints

Zentner A. R., Rudd D. H., Hu W., 2008, *Phys. Rev. D*, 77, 043507

Zhang P., 2008, ArXiv 0811.0613

Zheng Z. et al., 2005, *ApJ*, 633, 791