# Carnegie Mellon University
## MELLON COLLEGE OF SCIENCE

# THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

**DOCTOR OF PHILOSOPHY IN THE FIELD OF PHYSICS**

TITLE: "Weak lensing science with BOSS"

PRESENTED BY: Sukhdeep Singh

ACCEPTED BY THE DEPARTMENT OF PHYSICS

| | |
|---|---|
| Rachel Mandelbaum | 08/08/17 |
| RACHEL MANDELBAUM, CHAIR PROFESSOR | DATE |

| | |
|---|---|
| Stephen Garoff | 08/08/18 |
| STEPHEN GAROFF, DEPT HEAD | DATE |

APPROVED BY THE COLLEGE COUNCIL

| | |
|---|---|
| Rebecca Doerge | |
| REBECCA DOERGE, DEAN | DATE |

# Weak lensing science with BOSS

Sukhdeep Singh

Department of Physics
Carnegie Mellon University
Pittsburgh, PA 15213

**Thesis Committee:**
Rachel Mandelbaum (Chair)
Rupert Croft
Shirley Ho
Jeffrey A. Newman

August 13, 2017

*Submitted in partial fulfillment of the requirements*
*for the degree of Doctor of Philosophy.*

Copyright © 2017 Sukhdeep Singh

*To my Parents*

*For their love, hard work and dedication*

# Abstract

Observations from the large galaxy surveys, especially the measurements of galaxy clustering and weak gravitational lensing have emerged as a strong probes of cosmology in the last decade and a half. Joint analysis of galaxy clustering and galaxy position-lensing (g-L) cross correlations provides a direct measurement of the clustering of matter in the universe and thus a strong test of models of cosmology, independent of the galaxy bias. Combined with measurements of galaxy velocities from redshift space distortions, these measurements can also test the models of gravity.

In this work, we present new measurements and analysis of the galaxy clustering and g-L cross correlations, using spectroscopic galaxies from the SDSS-III BOSS survey and galaxy shear maps from the SDSS and CMB lensing maps from the Planck survey. Using these measurements, we constrain the amplitude of matter correlations and $E_G$ at $\sim 10\%$ level showing consistency with the Planck $\Lambda$CDM model at $\lesssim 1.5\sigma$ levels. We also study various sources of noise in the measurements using covariance estimations from mocks and theory. We show that using mean subtracted galaxy overdensity field in g-L cross correlations is necessary for the optimal covariance properties and that different methods of covariance estimation in the literature give consistent results when using optimal estimators.

With the increasing precision of cosmological measurements, understanding the impact of systematic errors, both theoretical and observational, is becoming ever more important. Among the theoretical systematics, incomplete modeling of non-linear physics is perhaps the most important. We perform checks on models using mock datasets and apply corrections to control the modeling uncertainties to below 2% levels. For observational systematics, we present the application of a powerful test to constrain relative biases in galaxy and CMB lensing maps by directly comparing their cross correlations with galaxy positions. Given the noise in g-L measurements using CMB lensing, we find the galaxy lensing and CMB lensing to be consistent at $\sim 12\%$ level. We perform additional tests for systematics in galaxy lensing measurements most notably tests for biases due to photometric redshifts, finding no significant source of systematics, though our tests are not exhaustive.

In the second part of this work, we present a detailed analysis of intrinsic alignments (IA) of galaxies using BOSS LOWZ sample. Intrinsic alignments arise due to alignments of galaxy shapes with their local environment. These alignments introduce additional coherent correlations in the galaxy shapes and thus contaminate the weak lensing measurements. IA also probe galaxy physics as they depend sensitively on the processes involved in galaxy formation and evolution. Understanding and accurately modeling the IA is thus one of the major tasks for the weak lensing community.

We perform measurements over a large range of scales and study the properties of IA in both 1-halo and 2-halo regimes. Our measurements are well fit by the tidal alignment models which include non-linear alignment model (NLA) in the 2-halo regime and a halo model with radial alignments of satellite galaxies. These models successfully explains the scaling as well as significant anisotropy in the IA measurements introduced by projected shapes and redshift space distortions. We study the dependence of IA amplitude on several galaxy properties, finding that brighter galaxies or galaxies in overdense regions show stronger alignments. We also study the dependence of IA on shape measurement methods. Shape measurements that give higher weight to outer parts of galaxies show stronger alignments, which suggests that galaxy shapes are twisted, with outer parts being more strongly aligned with tidal fields. These measurements have improved our understanding of IA and will provide sensitive tests of new models of IA being developed using analytical methods as well as cosmological simulations.

# Acknowledgements

This work has been made possible by significant contributions and support from several people. I should start by thanking my Parents who have worked very hard and made numerous sacrifices in their life to provide me with every opportunity to succeed. They are also among the best teachers I have had and have always pushed me to get better. Their efforts helped in building strong foundations during my formative years that prepared me to pursue my passion for research.

I have always been fortunate to have good teachers and mentors throughout my school and college education. While I won't put a long list of names here, I will always be grateful to all of them for educating me. Among them, a very special thanks goes to Rachel Mandelbaum, for being a very supportive advisor and patiently teaching me various aspects of being a good researcher in cosmology. I enjoyed the freedom she gave me to work and try new things. At the same time, her keen eye for details has prevented me from falling into several pitfalls and pushed me to learn new things that I perhaps otherwise would have not. She has always been kind and encouraging during difficult times and I have felt comfortable discussing mistakes and failures with her. I am always inspired by her knowledge, and passion for cosmology and science in general and consider myself lucky to have found an advisor who was a perfect match for me.

I also had lot of support from my friends and colleagues in the physics department and the McWilliams center. I would begin by thanking my dear friend Shadab Alam, who was also my roommate and put up with my snarky personality for four years with a smile. He also taught me cooking and then stomached it for a long time. We also shared several experiences, travelling together many times, teaching driving to one another and even collaborated on research. Finding such a nice friend and roommate was my second stroke of luck at CMU. After Shadab, I got another good friend in Prashant Shrivastava as my roommate and I thank him for his understanding during the hectic final year. I also shared office and became friends with several wonderful people, Sanxi Yao, Zach Mcdargh, Will McGinley and Jun Li in my first year and then with Ananth Tenetti and cosmology postdocs Alex Geringer-Sameth, Marianna Vargas, Sebastien Fromenteau and Anthony Pullen. I learnt a lot from them and especially from discussions with cosmology postdocs also including Ying Zu, Danielle Leonard, Melanie Simet, Layne Price and Francois Lanusse. I would espcially like to thank Alex, who is among the most inquisitive people I have met and Francois, who is probably the epitome of a nice person. They are very good friends and have taught me many things about life and science and have patiently listened to my complaints about random things over the years.

I have had many wonderful collaborators who have taught me many things and pointed out several mistakes. I would especially like to thank Benjamin Joachimi, Uroš Seljak, Anže Slosar, Jonathan Blazek, Elisa Chisari and Shirley Ho for all their contributions and support. I also express my gratitude to the members of faculty in the McWilliams center and the physics department for their support and for facilitating a very friendly and welcoming environment in the department. In particular, I appreciate the perpetual support received from Prof. Manfred Paulini and the members of staff in the physics department as a graduate student.

Last but not least, I thank Hung-Jin for being the sweetest person in my life and for pushing me to be a better and more genial person. I will always cherish the delightful memories of the time we have spent together and I look forward to a future that we build together.

I also thank the taxpayers from India and the United states whose hard earned money has supported/subsidized my education and research throughout.

# Contents

# 1 Introduction

The standard model of cosmology, ΛCDM model, underpins our understanding of the universe and has been successful in explaining a wide array of observations, from the early universe to the geometry and large scale structure in the universe at late times. Within ΛCDM model, the universe began in a homogeneous hot and dense state and has been adiabatically expanding ever since, cooling down and forming structure as we see today.

The expansion of the universe has been well accepted since Edwin Hubble's discovery in 1929 that galaxies farther away from us are receding at higher velocities. In the universe with no special observer, this discovery implies that space part of space-time metric is expanding homogeneously. Discovery of Cosmic Microwave Background radiation (CMB) in 1960s further provided the evidence that the universe was homogeneous (to one part in $10^{-5}$) and was dominated by radiation at early times. Since 1970s (see also Zwicky, 1933), it has also been observed that matter inside galaxies and clusters is moving faster than predicted by models using only observed matter (gas and stars) and our understanding of gravity based on general theory of relativity (GR). Two classes of explanation have been proposed to explain the discrepancy: 1) The cold Dark matter (CDM) explanation: there is dark matter component in the universe than we cannot observe, except for its gravitational effects (see Blumenthal et al., 1984, for an early review) and/or 2) Our understanding of gravity is incomplete and GR needs to be modified (see Clifton et al., 2012, for a review). CDM has gained wider traction for being able to explain a wide set of cosmological observations with a simpler theoretical model though several modified gravity models have not been all ruled out yet.

In the recent decades the field of Cosmology has progressed rapidly with wealth of data provided by several galaxy, supernova and CMB surveys, with bigger and more ambitious surveys being currently carried out and being planned. The most startling discovery in the recent decades has been that the expansion of universe is accelerating. This is contrary to the expectations as in the universe dominated by matter, gravity pulls everything closer, slowing down the expansion. First discovered by Riess et al. (1998); Perlmutter et al. (1999) through studies measuring distances to type Ia supernova, accelerating expansion has been subsequently confirmed by studies of CMB (Jaffe et al., 2001; Spergel et al., 2003; Hinshaw et al., 2013; Planck Collaboration et al., 2015a), Baryon Acoustic Oscillations (BAO) (Cole et al., 2005; Eisenstein et al., 2005; Alam et al., 2016a), galaxy correlations (Efstathiou et al., 2002; Alam et al., 2016a), galaxy velocities (Alam et al., 2016a) and weak gravitational lensing (Hoekstra et al., 2002; Sheldon et al., 2004; Massey et al., 2007; Kilbinger et al., 2013; Mandelbaum et al., 2013; Abbott et al., 2016; Hildebrandt et al., 2017). An excellent and detailed review of probes of cosmic acceleration is provided by Weinberg et al. (2013). If gravity is the dominant force on cosmological scales, the accelerating expansion again suggests that either our understanding of gravity is incomplete or it indicates presence of another mysterious component in universe, the dark energy, with a negative pressure that can cause the expansion. One model for such a component, the cosmological constant, Λ, was introduced by Einstein to counter the gravitational effects of matter and explain the static universe, before Hubble's discovery of expanding universe. In the current ΛCDM model, we use the cosmological constant as a model for dark energy to explain the accelerating expansion of the universe. While ΛCDM model successfully explains a wide range of cosmological observations, understanding the nature of dark matter and dark energy is currently a major challenge in cosmology and indeed in entire field of physics. Within observational cosmology, current focus is to improve the precision of measurements and gain better understanding of effects of dark matter and dark energy on the geometry and large scale structure in the universe.

In this work, we will focus on cosmological observations using weak gravitational lensing (weak lensing hereafter) which has emerged as an important probe of cosmology. In weak regime, gravitational lensing, the bending of path of light rays by gravitational effects of matter, encodes information about the large scale structure as well as geometry of the universe in the tiny but coherent distortions of the shapes and size of galaxies as well as CMB anisotropies. Weak lensing measurements have already been carried out by several current and past surveys, using galaxies: RCS (Hoekstra et al., 2002), SDSS (Sheldon et al., 2004; Mandelbaum et al., 2005), COSMOS (Massey et al., 2007),

CFHTLens (Kilbinger et al., 2013), HSC (Mandelbaum et al., 2017), KiDS (Hildebrandt et al., 2017), DES (Abbott et al., 2016) and using CMB: Planck (Planck Collaboration et al., 2015b), ACT (Das et al., 2011, 2014), SPT (van Engelen et al., 2012). There is now a major push by the community to improve measurements with upcoming surveys, such as LSST (LSST Science Collaboration et al., 2009), EUCLID (Laureijs et al., 2011), WFIRST (Spergel et al., 2013) and CMB-stage IV experiments (Abazajian et al., 2016). Unfortunately weak lensing measurements are also susceptible to a number of systematic errors (systematics hereafter) (Hirata et al., 2004; Mandelbaum et al., 2005; Massey et al., 2013). Though current results are dominated by statistical uncertainties, the systematics will become important in the future with increasing statistical precision of data. In this work we will focus on improving our understanding of several systematic errors in the weak lensing maps using cross-correlations between galaxies and weak lensing maps. Using these measurements, we also perform tests on ΛCDM model and our understanding of gravity based on general theory of relativity. We will also perform measurements of Intrinsic alignments of galaxies, which are a source of systematic error in weak lensing measurements and also a useful probe of galaxy physics.

In the reminder of this chapter, we focus on reviewing the background in cosmology and motivate our study of weak gravitational lensing. This discussion will follow the review by Weinberg et al. (2013) and cosmology texts by Dodelson (2003); Mukhanov (2005); Mo et al. (2010).

## 1.1 Cosmology Background

### 1.1.1 Metric and Distances

The basic assumption underlying our models of cosmology is that the universe is homogeneous and isotropic. Mathematically, we can write the FRW metric for expanding four dimensional space-time that satisfies these assumptions

$$ds^2 = dt^2 - a(t)^2 \left( \frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right) \tag{1.1}$$

$$= dt^2 - a(t)^2 \left( d\chi^2 + f_K^2(\chi)(d\theta^2 + \sin^2\theta d\phi^2) \right) \tag{1.2}$$

Where $t$ is the proper time, $a(t)$ is the time dependent scale factor, and $r$ is the proper distance. $K$ is a constant (note that it has dimensions of $[\text{length}]^{-2}$) representing the curvature of the space, with $K = 0$ being flat Euclidean space, $K = 1$ represents spherical geometry with positive curvature (sphere) and $K = -1$ is the hyperbolic surface with negative curvature. In the second equality, we defined comoving distance $\chi$ as

$$d\chi = \frac{dr}{\sqrt{1 - Kr^2}} = c\frac{dt}{a(t)} = c\frac{da}{H(a)a^2} \tag{1.3}$$

The second equality comes from defining comoving distance as the radial distance between two points along the null geodesics ($ds^2 = 0$). We define Hubble constant as (see section 1.1.3 for more discussion)

$$H(a) = \dot{a}/a \tag{1.4}$$

The transverse comoving distance $f_K(\chi)(= r)$ is the proper distance between two events separated by angular distance $d\theta$ in the observer frame and can be written as

$$f_K(\chi) = \begin{cases} \frac{1}{\sqrt{K}} \sin(\sqrt{K}\chi) & (K = +1) \\ \chi & (K = 0) \\ \frac{1}{\sqrt{|K|}} \sinh(\sqrt{|K|}\chi) & (K = -1) \end{cases} \tag{1.5}$$

In observations, we often measure the angular size of objects, which can then be converted to physical size using the line of sight distance, $D_A$ (called angular diameter distance)

$$D_A(a) = a f_K(\chi) \tag{1.6}$$

In geometric probes such as BAO and position of CMB anisotropy peaks, their size is known from physical models and thus observed angular size can be used to determine the $D_A$. We will discuss the role of distance measures as probe of cosmology in section 1.2.1. A more detailed review of several distance measures used in cosmology is provided by Hogg (1999).

### 1.1.2 Redshift

An important method to estimate the line of sight distance to cosmological objects is through the red shifting of light from these objects. Light signals observed from these objects were emitted at a different time and have travelled over cosmological distances. Since light travels along null geodesics ($ds^2 = 0$) and in comoving coordinates, the separation between two points does not change (or in other words, the wavelength of light does not change in comoving space), we have $d\chi_o = d\chi_e$ (where $o, e$ denote observation and emission frames), which implies

$$\frac{dt_o}{dt_e} = \frac{a(t_o)}{a(t_e)} = \frac{\lambda_o}{\lambda_e} = 1 + z \tag{1.7}$$

Thus, as the photons travel across the universe, their wavelength increases and they lose energy (they are redshifted). Observationally, we measure the wavelengths of photons and by comparing wavelengths of known spectral lines, estimate the redshift of objects. Then using the relations between redshift and distances, we can determine the distances. In section 1.1.1, we defined distance measures as function of scale factor $a$, which can be easily converted to functions of redshift using the fact that $a(z) = \frac{1}{1+z}$, where present day scale factor is set to 1, $a(z = 0) = 1$ .

### 1.1.3 Hubble Constant

Hubble constant, defined in eq. 1.4, describes the expansion of the universe. Using the Einstein field equation along with stress energy tensor of an ideal fluid, we can relate the expansion of the universe with the evolution of curvature and energy density of different constituents of the universe and derive the Friedman equation (see chapter 3 of Mo et al., 2010)

$$H^2(a) = \frac{8\pi G}{3} \left(\rho_m(a) + \rho_R(a) + \rho_\Lambda(a) + \rho_K(a)\right) \tag{1.8}$$

where $\rho_K = -\frac{3Kc^2}{a^2 8\pi G}$, $\rho_\Lambda = \frac{\Lambda c^2}{8\pi G}$, $\rho_m$ and $\rho_r$ are the matter and radiation density respectively.

To understand the evolution for various energy density components in an expanding universe, we begin by defining the equation of state, $w$, relating the pressure and density of an ideal fluid

$$\frac{P}{\rho c^2} = w(a) \tag{1.9}$$

Assuming universe is expanding adiabatically and using the conservation of energy we can show

$$\rho \propto \exp\left(\int -3(1 + w(a))d\ln a\right) \propto a^{-3(1+w)} \tag{1.10}$$

where second proportionality is for constant equation of state ($w(a)$ =constant). For pressure less dark matter, $w = 0$, which implies $\rho_m \propto a^{-3}$. This is a statement of conservation of mass that the density of dark matter is inversely proportional to the volume. For Radiation, $w = 1/3$ and hence $\rho_r \propto a^{-4}$, where 3 factors are contributed by change in number density of photons with volume (ignoring absorption and emission at late times) and the additional factor comes from the redshift of photons as they lose energy in the expanding universe. For $\Lambda$, $w = -1$ and $\rho_\lambda$ =constant. Using these relations, we can write Hubble constant from eq.1.8 as

$$H^2 = \frac{8\pi G\rho_{crit}}{3} \left(\Omega_m a^{-3} + \Omega_R a^{-4} + \Omega_K a^{-2} + \Omega_\Lambda\right) \tag{1.11}$$

Where critical energy density $\rho_{crit} = \frac{3H^2}{8\pi G}$ is the energy density required for a flat universe and $\Omega_i = \rho_i/\rho_{crit}$, $\sum \Omega_i = 1$.

From Einstein's equation, we also get a relation to describe the acceleration of the expansion

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\rho(1+3w) \equiv \begin{cases} > 0; & w < -\frac{1}{3} \\ < 0; & w > -\frac{1}{3} \end{cases} \tag{1.12}$$

The expansion decelerates when the universe is dominated by matter or radiation components ($w_{eff} > -1/3$) but the expansion is accelerated when the universe is dominated by $\Lambda$ like component (or Dark energy) with $w_{eff} < -1/3$.

### 1.1.4 Large scale structure

CMB observations have shown that the early universe was remarkably homogeneous, with the deviations from homogeneity (perturbations) of order $\sim 10^{-5}$. These perturbations are believed to have originated during the inflation, when the universe expanded exponentially ($H$ ~constant) driven by a scalar field. Small quantum perturbations in the the energy density of the scalar field lead to the metric perturbations after the inflation ends.

After inflation ends, universe enters the period of radiation domination following reheating and big bang nucleosynthesis. The perturbation modes that enter the horizon during the radiation domination grow slowly as the free streaming of photons decays potential. The evolution of matter perturbations through the radiation dominated area is described in terms of transfer functions, $T(k)$, where $T(k) \propto 1/k^2$ (Hu & Sugiyama, 1996) for modes that enter horizon during the radiation dominated era and $T(k) = 9/10$ for modes that enter horizon during matter dominated era (see chapter 7 of Dodelson (2003) or chapter 4 of Mo et al. (2010) for more details).

After the matter-radiation equality, in the matter dominated epoch, gravity takes over, growing these perturbations which leads to the large scale structure we observe at later times. In Fourier space, the observed perturbation mode at time $t$ can be written in terms of initial mode $\delta_i$ as

$$\delta(\mathbf{k}, t) = \delta_i(\mathbf{k})T(k)D(t) \tag{1.13}$$

$D(t)$ is the growth function after the radiation epoch ends. CMB measurements constrain the power spectrum during the epoch of recombination and the studies of large scale structure at late times, including weak lensing, focus mostly on studying the growth function $D(t)$.

In the matter dominated epoch, gravity acts as the dominant force and gravitational instability leads to the growth of modes as they enter the horizon. Assuming dark matter as pressure-less fluid and using Euler and continuity equations for fluids in expanding universe, the evolution of modes to first (linear) order is described by

$$\frac{\partial^2\delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial\delta}{\partial t} - 4\pi G\bar{\rho}\delta = 0 \tag{1.14}$$

Taking double derivative of Hubble constant from eq. 1.11, it can be shown that the decaying solution of eq.1.14 is simply $\delta_- \propto H$ (decaying because Hubble constant decreases with time). The growing solution, $\delta_+(t) = \delta_i D(t)$ can then be derived be derived in terms of $H$ (see e.g. Lahav & Suto, 2004, for more details)

$$D \propto H(t)\int_0^t \frac{dt'}{a^2(t')H^2(t')} \propto H(z)\int_z^\infty \frac{(1+z')}{E^3(z')}dz' \tag{1.15}$$

where $E(z) = H(z)/H_0$.

Eq. 1.14 and hence growth function in eq. 1.15 is only a good approximation for modes that are linear $\delta \ll 1$. Higher order terms in gravitational evolution couple different (Fourier) modes and once $\delta \sim 1$, evolution becomes non-linear and coupling between modes cannot be ignored. Quasi-linear

evolution (upto few orders in $\delta$) can be studied using higher order perturbation theory (see Bernardeau et al., 2002, for a review), though extending higher order perturbation theory to small scales has proven to be a difficult task. Understanding evolution of modes down to very small scales still requires use of simulations of structure formation. A common strategy is to run few simulations with different cosmological parameters and then construct emulators for the observables by interpolating over these simulations (Smith et al., 2003; Takahashi et al., 2012; Lawrence et al., 2010). For modeling our observations, we use halofit (Smith et al., 2003; Takahashi et al., 2012) to get matter power spectrum beyond linear approximation.

## 1.2 Cosmology Probes

One of the primary goals of cosmological probes is to determine the geometry and energy density constituents of the universe. The cosmology probes can be broadly classified into following categories: 1) Probes of initial conditions, primarily CMB, 2) Local universe measurements of Hubble constant $H_0$, 3) Probes of geometry (distance measurements), and 4) Probes of growth of structure. In this work section we will focus on 3 and 4, and refer the reader to Hu & Dodelson (2002); Planck Collaboration et al. (2015a) for a review and recent measurements of CMB (we will briefly discuss CMB as probe of geometry) and Freedman & Madore (2010); Riess et al. (2016) for review and recent measurements of $H_0$.

Measurements of CMB provide a very sensitive test of the conditions at the time of matter-radiation equality. CMB provides strong constraints on the matter power spectrum, matter and radiation energy density as well as geometry of the universe. However, it is not a very sensitive probe of late time evolution, which is dominated by matter and dark energy. Distance measurements and the late time evolution of the universe depends on the parameter $E(z)$ that determines the evolution of Hubble constant and also the growth of structure

$$E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_R(1+z)^4 + \Omega_K(1+z)^2 + \Omega_{DE}\exp\left(3\int_0^z [1+w(z')]\frac{dz'}{1+z'}\right)} \tag{1.16}$$

In writing this equation, we did not assume that $w = -1$ for dark energy but instead let it be a function of redshift (note that $w(z) \neq -1$ rules out $\Lambda$ as model for dark energy). Measuring $E(z)$ and its evolution with redshift thus provides an important probe of nature and amount of dark energy and dark matter in the universe. Most large scale structure probes aim to measure $E(z)$ by either probing the geometry, distance to a redshift depends on $E(z)$ or by probing the growth of structure which also depends on $E(z)$ as shown in eq. 1.15.

### 1.2.1 Probes of geometry

The geometrical probes of cosmology usually provide an independent measure of distances to a given redshift which can then be compared with the predictions from the model as presented in section 1.1.1.

Observations of acoustic peaks in the CMB power spectrum provide a distance measure to the surface of last scattering ($z \sim 1100$). Before recombination, the photons and baryons are coupled and due to effects of perturbations in potential and the countering radiation pressure, the perturbations in photon-baryon plasma oscillate while the perturbations in dark matter grow in situ though the growth of these perturbations is also affected by radiation. At the time of recombination, photons decouple from baryons and the phase of oscillations get imprinted on the CMB that we observe at later times. The physical size of oscillations depend on the size of sound horizon, $s_*$ (determined by matter density and baryon to photon ratio) at the time of recombination and the observed position of peaks, which follow the harmonic relation $\ell_n = nD_A(z_*)/s_*$, provides a measure of distance to the surface of last scattering ($z_* \sim 1100$). In addition, the relative heights of the peaks provide further information on the dark matter, baryon and radiation densities. While measurements of CMB power spectrum provide strong constraints on matter density, the constraints on dark energy are weaker,

since dark energy was subdominant at time of recombination and CMB alone also can not break the degeneracy between $H_0$ and $E(z)$ which goes into measurements of distances. Combination of CMB with late time probe of $H_0$ and $E(z)$ provides much stronger constraints on dark energy (see for ex. Planck Collaboration et al., 2015a, for CMB alone and CMB+other probes constraints).

Measuring distances to multiple redshifts helps in breaking degeneracy between different parameters as they evolve differently with redshift. One such probe is the Baryon Acoustic Oscillations (BAO), the imprint of sound horizon at the surface of last scattering in the baryon density. After recombination, the acoustic oscillations in baryon distributions are frozen, leading to slightly higher correlations in matter distribution on the scales of sound horizon. Since the BAO scale is large ($\sim 110h^{-1}$Mpc comoving), the BAO peak is relatively less affected by the late time evolution in the structure, though the baryons are pulled back by the dark matter leading to the broadening of the peak. Broadening increases the uncertainty in the BAO measurements as it becomes harder to determine the exact location of the peak. Also the large scale of BAO means means that it requires large volume surveys to measure the peak with good accuracy. Current BAO measurements use large scale galaxy surveys (e.g. BOSS Alam et al., 2016a), though with modest galaxy densities and some form of reconstruction techniques which undo the effects of structure formation and sharpen the peak (see for example Padmanabhan et al., 2009). An example of BAO ring in the recent measurement correlation function of galaxies is shown in figure 1.1

Another probe of distances is to use standard candles, i.e. objects of known absolute luminosity which can then be compared with observed luminosity to compute distances. In an expanding universe, the luminosity distances computed from flux (energy per unit area per unit time) is (Hogg, 1999)

$$D_L = (1+z) f_K(\chi_z) \tag{1.17}$$

It is difficult to find sources which are bright enough to be standard candles over cosmological scales. Fortunately, one such source is type-Ia supernovae (SN), whose peak luminosity and luminosity curve are tightly correlated and thus their light curves can be standardized. Using SN observations, Riess et al. (1998); Perlmutter et al. (1999) first made the confirmed discovery of expanding universe. Though SN are standard candles with low scatter, the diversity of light curves, extinction of light by host galaxy of SN as well as intervening matter between supernova and the observer can complicate the their analysis. Also, supernova observations only measure relative distance to different redshifts (their fundamental calibration is not known independently of other probes) and thus they are not sensitive to $H_0$ values, unlike BAO.

Measurements of weak gravitational lensing also probe the geometry, as gravitational lensing depends on the distances between lens, source and observer (see section 1.3). However, since it depends on ratio of distances, lensing itself is not a very strong probe of geometry and large separations between observer, lens and lensing sources is needed to break degeneracy between different parameters (see discussion in Weinberg et al., 2013). Hu et al. (2007b) proposed an idea to combine CMB lensing measurements with galaxy lensing to maximize the redshift baselines for increased sensitivity. We will present one such measurement of distance ratios by combining CMB lensing with galaxy lensing from SDSS in chapter 2 (see also Miyatake et al. (2016)).

### 1.2.2 Probes of Structure Growth

As discussed in section 1.1.4, the initial perturbations in the matter density grow under the effects of gravity in the epoch of matter dominations. Studying the large scale structure at late times provides a measure of the growth of structure and hence constraints on the cosmological models. We will focus on measuring the two point functions of the density field to study the growth of structure.

#### Correlation function and Power spectrum

CMB observations have shown that the initial density field can be treated as gaussian random field, which is completely described by it's first and second order moments (this can be shown using Wick's

or Isserlis' theorem for gaussian field). The correlation function, $\xi(r)$ and power spectrum, $P(k)$ are defined as

$$\xi(|\mathbf{x}_1 - \mathbf{x}_2|) = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\rangle \tag{1.18}$$

$$P(|k|) = \langle \tilde{\delta}(\mathbf{k})\tilde{\delta}(\mathbf{k}_2)\rangle \delta_D(\mathbf{k} + \mathbf{k}_2) \tag{1.19}$$

Where isotropy ensures that correlation function and power spectrum only depend on absolute values of $\mathbf{k}$ (also $\mathbf{k} + \mathbf{k}_2 = 0$) and separation $|\mathbf{x}_1 - \mathbf{x}_2|$. It is also simple to show that power spectrum and correlation function are Fourier transforms of each other.

From simple slow roll Inflation model with exponential expansion during the inflationary period we get the metric perturbations with nearly scale invariant power spectrum

$$\Delta^2_\Phi(k) \propto k^{n-1} \tag{1.20}$$

where we defined $\Delta^2_\phi = k^3 P_\phi(k)$ and $n = 1$ (Harrison-Zel'dovich spectrum). Using Poisson relation ($k^2\phi_k \propto \delta_k$), we can relate the metric perturbations to perturbations in energy density and show that initial matter power spectrum can be described by a power law

$$P_i(k) \propto k^n \tag{1.21}$$

Under the assumption of linear evolution, using eq. 1.13 we can write late time power spectra as

$$P(k, t) = D(t)^2 T(k)^2 P_i(k) \tag{1.22}$$

Due to differences in $T(k)$ before and after epoch of matter-radiation equality, power spectrum turns around $k_{eq}$, where $k_{eq}$ is horizon size (in fourier space) at time of matter-radiation equality. For modes that enter horizon after matter-radiation equality ($k < k_{eq}$ ), $P(k, t) \propto k$ modulo some evolution due to $D(t)$ as different modes enter horizon at different times. For modes that enter horizon earlier, $k > k_{eq}$, power is suppressed and using $T(k) \propto k^{-2}$ $P(k) \propto k^{-3}$, though this is also modified by late time non-linear evolution. The linear growth assumption works well at large scales, $k \lesssim 0.1h^{-1}\mathrm{Mpc}^{-1}$ at $z \sim 0$, but at smaller scales (larger $k$) the higher order terms become important. Non-linear perturbation theory can describe power spectra to somewhat higher $k$ (Bernardeau et al., 2002), but we need cosmological simulations to model the power spectra down to very small scales (large $k$).

### Galaxy correlations

Galaxy surveys observe and map the positions of galaxies over the sky at different redshifts. Galaxies form inside dark matter halos, which tend to form in the overdense regions (peaks in the matter over-density field). Since overdense regions have preferentially more halos while underdense regions have fewer halos, halo over-density field is generally a biased tracer of the matter field. To describe galaxies, we also need to describe the relation between halos and galaxies. Detailed relation between galaxies and dark matter halo can be described using halo-occupation-distribution (HOD) models (see Cooray & Sheth, 2002, for review). We use samples with low satellite fractions and also restrict most of our analysis to use large enough scales where HOD details are not important. The galaxy bias we measure will be a weighted mean of halo bias (higher mass halos with more galaxies get higher weight). When modeling small scales (1-halo regime), we will assume that correlation function is dominated by central-satellite pairs and hence can be modeled with simple halo profiles such as NFW. For further discussion in this section, we will assume that halo bias and galaxy bias describe the same concept and write everything in terms of galaxies. A detailed review of galaxy biasing is given by Desjacques et al. (2016).

In general, the observed galaxy field, $\delta_g$ can be described as a (local) non-linear transform of the matter field, $\delta_m$

$$\delta_g = \mathcal{F}(\delta_m) = b_1\delta_m + b_2\delta_m^2 + \ldots b_n\delta_m^n \tag{1.23}$$

Where in the second equality we taylor expanded the $\delta_g$, assuming it only depends on the local matter density. The correlation function of galaxy field can then be written in terms of matter correlation function $\xi_{mm}$ as

$$\xi_{gg}(r) = \langle \delta_g \delta_g \rangle = \sum_{i,j} b_i b_j [\xi_{mm}(r)]^{\frac{i+j}{2}} \tag{1.24}$$

Note that the terms with $i + j =$ odd number will be zero since we assumed that the over-density field is gaussian with mean zero (and hence their odd moments will be mean zero). For large $r$, $\xi_{mm}$ is small ($\ll 1$) and only the first order term is important in which case we can assume bias is linear and write

$$\xi_{gg}(r) = b_1^2 \xi_{mm}(r) \tag{1.25}$$

$$P_{gg}(r) = b_1^2 P_{mm}(r) \tag{1.26}$$

Linear scale independent bias is typically a good assumption on large scales where the density field is linear, but it breaks down at small scales where density field has gone non-linear, $r \lesssim 10h^{-1}$Mpc and bias is a scale dependent non-linear function on these scales.

We will typically assume linear bias when modeling galaxy correlations which is a good assumption on the scales we model given the uncertainty in the measurements. When necessary, we do apply corrections for the non-linear bias effects by computing them from simulations (see chapter 4).

**Redshift space distortions**

So far we have assumed that the observed redshift of cosmological objects is entirely due to the expansion of the universe. However, most objects are not at rest with their local cosmological rest frame, but instead are moving, primarily under the effects of gravity from large scale structure. The observed line of sight velocity of an object is given by

$$\frac{dr_z}{dt} = \frac{d}{dt} a\chi = \dot{a}\chi + a\dot{\chi} \tag{1.27}$$

Where $\dot{a}\chi$ is the velocity due to cosmological expansion and $a\dot{\chi}$ is the *peculiar* velocity of the object due to its motion with respect to local rest frame. The observed redshift is then

$$1 + z_o = \frac{\lambda_o}{\lambda_e} = \frac{\lambda_o}{\lambda_{e,v}} \frac{\lambda_{e,v}}{\lambda_e} = (1 + z_v)(1 + z) \tag{1.28}$$

where $\lambda_{e,v}$ is the doppler shifted wavelength due to peculiar motion and $1 + z_v$ is the redshift due to this motion. Since we convert the observed redshift into distances using distance-redshift relation, the observed position of the object in redshift space, $\mathbf{s}$, is shifted and is related to its real space position $\mathbf{r}$ by

$$\mathbf{s} = \mathbf{r} + \mathbf{v}_H \cdot \hat{z} \tag{1.29}$$

$\hat{z}$ is along the line of sight and factor of $H(a)$ is included in $\mathbf{v}_H$ ($\Delta s = v/H(a)$,$\mathbf{v}_H = \mathbf{v}/H$). Since the observed positions of objects are shifted in the redshift space, the observed over-density field also changes. Most notably, if the velocities are coherent, the over-density field is no longer isotropic in the redshift space, but has different characteristics along and perpendicular to the line of sight.

Using conservation of mass and the relation between large scale velocity field and growth rate ($\nabla_{\chi_r} \cdot \mathbf{v} = -aHf$), we can show that

$$\delta^s \approx \delta^r + \mu^2 f \delta_r \tag{1.30}$$

where $\mu = k_z/k$ and growth rate $f$

$$f = \frac{\partial \log D(a)}{\partial \log a} \tag{1.31}$$

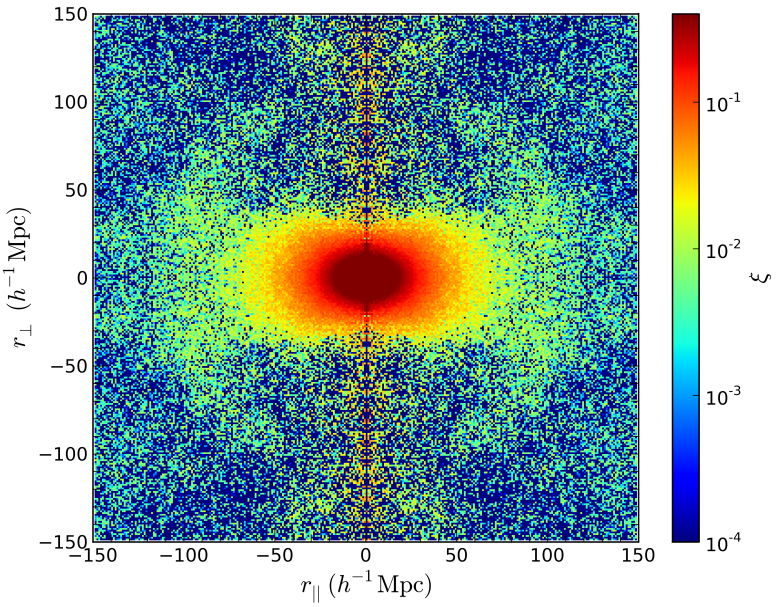



Figure 1.1: Measurement of galaxy correlation function by Samushia et al. (2013). The distortion or anisotropy due to redshift space distortions is clearly visible as elliptical contours (in absence of RSD, contours will be circular). The broadened BAO ring is also visible at scales around $100h^{-1}$Mpc. Plot taken from http://www.sdss3.org/science/boss_publications.php

The additional term in eq. 1.30, $\mu^2 f\delta_r$ always has matter density $\delta_m$ and for biased tracers we get $\delta_b^s \approx \delta_b^r(1+\mu^2\beta)$, where $\beta = f/b$. The redshift space correlation function of galaxies thus includes contributions from matter correlations and galaxy-matter cross correlations and $\mu$ dependence of these terms makes the contributions anisotropic. An effect of RSD on the correlation functions is shown in figure 1.1. In the presence of RSD, the contours of correlation function are elliptical as opposed to being circular for an isotropic universe. By studying the anisotropy of the correlation functions or power spectra, we can measure the growth rate $f$. Most RSD studies achieve this by analyzing the higher order moments of the two point function in the $s, \mu$ space (these measurements are sensitive to $f(z)D(z)$ or $f\sigma_8$) (Alam et al., 2016a). Unfortunately, the non-linear velocity field at small scales, especially the 1-halo term leading to finger-of-god effect, smears the small scale correlations to larger $s$, which makes it challenging to model RSD measurements down to small scales (Carlson et al., 2013; Alam et al., 2015b). In chapter 4, we will perform measurements of RSD using methodology developed by Alam et al. (2015b) and use these measurements in combination with weak lensing measurements to test theories of gravity and ΛCDM model.

## 1.3 Weak lensing

Weak gravitational lensing is a powerful probe of cosmology as as it is sensitive to both geometry and growth of the structure in the universe (see Bartelmann & Schneider, 2001; Weinberg et al., 2013; Kilbinger, 2015, for review). As the light travels across the universe from the source to the observer, the gravitational effects of the intervening structure bends the path of photons. In the weak regime, lensing introduces small but coherent distortions in the shapes (shear) and sizes (convergence) of background sources. These distortions encode the information about the large scale structure between the observer and the source. In case of galaxy observations, we typically measure the correlations in the shapes of galaxies to estimate the lensing shear while in case of CMB, lensing

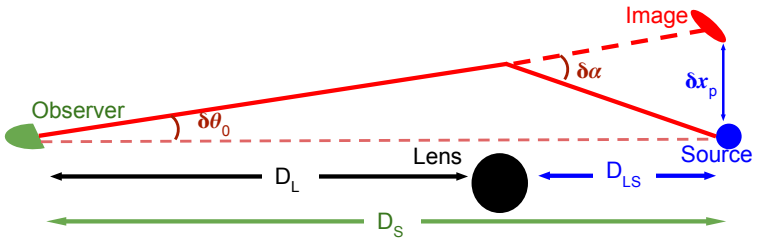



Figure 1.2: An illustration of gravitational lensing by a single lens (the effects are highly exaggerated). As light rays travel from source to observer their paths are distorted, shifting the position of the sources and its shape and size are also distorted. $D_L, D_S, D_{LS}$ are the physical distance between lens, source and the observer (we derived the formalism in corresponding comoving distances $f_k(\chi_l), f_k(\chi_s), f_k(\chi_{l,s})$).

is studied by measuring cross-correlations between different modes to estimate convergence.

Gravitational lensing measurements are sensitive directly to the structure in the matter and can be used to measure matter correlations, unlike other large scale tracers such as galaxies which are biased. Also, since lensing in sensitive to entire structure between source and observer (though this sensitivity varies with distances), a deep lensing survey is capable of studying the growth of large scale structure over a large redshift range. However, lensing measurements are also hard. Typical weak lensing shear is of order 0.01 compared to intrinsic galaxy ellipticities of order 0.2. In a lensing survey we need to accurately measure shapes of a large number of galaxies, while keeping any systematic errors to a level lower than statistical errors. The inferences from lensing are susceptible to a number of systematics originating from shape measurement errors, noisy estimations of redshifts of large number of galaxies, modeling uncertainties and astrophysical sources of coherent shear. Thus a large portion of our work will be focussed on estimating and studying some of the sources of systematics in lensing observables.

### 1.3.1 Gravitational lensing

In this section we develop the basic lensing formalism and derive some useful relations which can be used to model the lensing measurements. To derive the lensing equation, consider the perturbed FRW metric in the newtonian gauge ($c = 1$ units in this section)

$$ds^2 = a(\tau)^2\{-(1+2\psi)d\tau^2 + (1-2\phi)(d\chi^2 + f_\kappa^2(\chi)d\Omega^2)\} \tag{1.32}$$

Where $\psi$ is the newtonian potential and $\phi$ is the curvature potential.

After some algebra, we can derive the acceleration equation for photons traveling along the null geodesic ($ds^2 = 0$)

$$-\nabla\frac{(1+2\psi)}{1-2\phi} = 2d^2\mathbf{x}/d\tau^2 \tag{1.33}$$

Under the assumption of weak fields, $(\psi, \phi << 1)$, we can show that change in velocity vector in the plane of sky (perpendicular to line of sight) is (see figure 1.2 for illustration)

$$\delta\alpha = \frac{\delta x_\perp}{f_\kappa(\chi_s - \chi_l)} = d\hat{U}/d\chi = \nabla_\perp(\psi + \phi) \tag{1.34}$$

As shown in figure 1.2, $\delta\alpha$ is the change in angle with incident and refracted ray at the lens, $\delta x_\perp$ is the physical separation between source and the image in the plane of sky. The shift in the image in frame of observer, $\delta\theta_o$, can be obtained from $\delta x_\perp$

$$\delta\theta_o = \frac{\delta x_\perp}{f_\kappa(\chi_s)} = \int_0^{\chi_s}\frac{f_\kappa(\chi_s - \chi_l)}{f_\kappa(\chi_s)}\nabla_\perp(\psi_l + \phi_l)d\chi_l = \nabla_\theta\Phi_L \tag{1.35}$$

Here we have integrated over the full lensing potential between source and observer. We also defined lensing potential in terms of Weyl Potential, $\Psi = \psi + \phi$

$$\Phi_L = \int d\chi_l \frac{f_\kappa(\chi_s - \chi_l)}{f_\kappa(\chi_s) f_\kappa(\chi_l)} \Psi(f_\kappa(\chi_l)\theta, \chi_l) \tag{1.36}$$

The Jacobian of image to source transformation is

$$A_{ij} = \frac{\partial(\theta_o^i - \delta\theta_o^i)}{\partial\theta_o^j} = \delta_{ij} - \frac{\partial^2 \Phi_l}{\partial\theta_o^i \partial\theta_o^j} \tag{1.37}$$

$$A = \begin{bmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{bmatrix} = (1-\kappa)\begin{bmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{bmatrix}$$

Where convergence $\kappa$ and shear components $\gamma_1, \gamma_2$ are defined as

$$\kappa = 1/2\nabla_\theta^2 \Phi_L \tag{1.38}$$
$$\gamma_1 = 1/2(\Phi_{L,11} - \Phi_{L,22}) \tag{1.39}$$
$$\gamma_2 = \Phi_{L,12} \tag{1.40}$$
$$g_i = \frac{\gamma_i}{1-\kappa} \approx \gamma_i \tag{1.41}$$

Convergence leads to an isotropic change in the area of image while shear leads to a distortion in the shape and contributes to change in total area as well. Typically for galaxy lensing case we do not know the un-lensed size of galaxies and observe the reduced shear $g_i$. In the weak lensing regime, $\kappa \ll 1$, $g_i \approx \gamma_i$

Finally we can also write the convergence and shear expressions in Fourier space and using the poisson relation between potential and density, express them in terms of matter density $\delta$

$$\kappa(\ell) = 1/2\ell^2 \Phi_l(\ell) = \int dz_l \frac{c}{H(z_l)} \frac{\overline{\rho}_m}{\Sigma_c(z_l, z_s)} \delta\left(\frac{\ell}{\chi(z_l)}\right) \tag{1.42}$$
$$\gamma_1(\ell) = 1/2(\ell_1^2 - \ell_2^2)\Phi_l(\ell) = \int dz_l \frac{c}{H(z_l)} \frac{\overline{\rho}_m}{\Sigma_c(z_l, z_s)} \frac{(\ell_1^2 - \ell_2^2)}{\ell^2} \delta\left(\frac{\ell}{\chi(z_l)}\right) \tag{1.43}$$
$$\gamma_2(\ell) = \ell_1 \ell_2 \Phi_\ell = \int dz_l \frac{c}{H(z_l)} \frac{\overline{\rho}_m}{\Sigma_c(z_l, z_s)} \frac{\ell_1 \ell_2}{\ell^2} \delta\left(\frac{\ell}{\chi(z_l)}\right) \tag{1.44}$$
$$\tag{1.45}$$

Writing $\ell_1, \ell_2$ into cylindrical coordinates, $\ell_\perp, \phi_\ell$ we can also write shear as

$$\gamma_1 + i\gamma_2 = |\gamma| e^{i2\phi_\ell} = \int dz_l \frac{c}{H(z_l)} \frac{\overline{\rho}_m}{\Sigma_c(z_l, z_s)} \delta\left(\frac{\ell}{\chi(z_l)}\right) e^{i2\phi_\ell} \tag{1.46}$$

We also defined the geometric factor $\Sigma_c$ as

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{f_k(\chi_s)}{(1+z_l) f_k(\chi_l) f_k(\chi_s - \chi_l)} \tag{1.47}$$

### 1.3.2 Estimating shear from galaxy shapes

In galaxy lensing, we infer the lensing shear from the ensemble average of galaxy shapes. The observed galaxies have some intrinsic shapes which are sheared by the lensing and also by local tidal fields (intrinsic alignments). In this section we present a simplified discussion of how the galaxy shapes respond to shear and can be used to estimate shear.

For discussion in this section, we will assume that shear has only one component, $g_1$ with $g_2 = 0$. This can be achieved by applying a rotation to the coordinate axis, such that $\phi = 0$ in eq.1.46. Following (Bernstein & Jarvis, 2002), we represent the shear tensor as

$$S_\eta = \begin{bmatrix} e^{\eta/2} & 0 \\ 0 & e^{-\eta/2} \end{bmatrix} \equiv \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

where we parameterized the shearing of the shape with $\eta$. The axis ratio $q$ is defined as

$$q = \frac{b}{a} = e^{-\eta} \tag{1.48}$$

The ellipticity, $\epsilon$ (and analogously the reduced shear $g$), is defined as

$$\epsilon = \frac{1-q}{1+q} \approx \frac{\eta}{2} = g \tag{1.49}$$

In weak lensing measurements, distortion $e$ is also used to measure shear. Starting with the second moments of the image

$$Q_{ij} = \frac{\int d^2x I(\mathbf{x}) W(\mathbf{x}) x_i x_j}{\int d^2x I(\mathbf{x}) W(\mathbf{x})} \tag{1.50}$$

Where $W(\mathbf{x})$ is a weight function. $W(\mathbf{x})$ is usually chosen to down weight and/or truncate the integral beyond a certain range ($W(\mathbf{x}) = 1$ will lead to divergent noise since the integral is over whole sky). Alternative way is to fit a model for galaxies to image and then measure moments from the model. Distortion is then defined as

$$e = e_1 + ie_2 = \frac{Q_{11} - Q_{22} + i2Q_{12}}{Q_{11} + Q_{22}} = |e| \exp(i\phi) \tag{1.51}$$

$$|e| = \frac{1-q^2}{1+q^2} = \tanh \eta \tag{1.52}$$

In the absence of lensing (or any other source of coherent shear), galaxies are randomly oriented (random $\phi$) and each shape component can take positive or negative values. Thus the ensemble average of intrinsic ellipticity of galaxies is expected to be zero.

$$\langle e_i \rangle = 0 \tag{1.53}$$

The response of the ensemble average to a small coherent shear is given by the responsivity $\mathcal{R}$

$$\mathcal{R} = \frac{\partial \langle e_i \rangle}{\partial \eta} = \langle 1 - e_i^2 \rangle \tag{1.54}$$

Such that the shear is

$$g = \frac{\langle e \rangle}{2\mathcal{R}} \tag{1.55}$$

The zero mean from ensemble average of ellipticities greatly simplifies the problem of estimating shear. If ensemble average is not zero, then one needs to estimate the ensemble average accurately and subtract it out of the measured shear.

### 1.3.3 Two point functions

For cosmological analysis from lensing, we can construct the two point functions out of shear or convergence measurements. We can measure either the lensing-lensing correlations or cross correlations of lensing with other tracers of large scale structure, e.g. galaxy positions.

### Lensing correlations

Auto correlations of the lensing estimators, convergence or shear, are sensitive to the integrated matter power spectrum between the source and observer. While these correlations contain the integrated information along the line of sight, we can obtain more information by cross correlating sources at different redshift, in which case the information comes from the observer to the source at lower redshift.

In Fourier space, the power spectra for different combinations of shear and convergence are (note that throughout we will work under Limber approximation and assume flat sky wherever necessary)

$$C_\ell^{\kappa\kappa} = \int dz_{s1} p(z_{s1}) \int dz_{s2} p(z_{s2}) \int_0^{z_s} dz_l \frac{H(z_l)}{c} \frac{\bar{\rho}_m^2}{\Sigma_c(z_l, z_{s1})\Sigma_c(z_l, z_{s2})} P_{\delta\delta}\left(\frac{\ell}{\chi(z_l)}\right) \tag{1.56}$$

where $z_s = \min(z_{s1}, z_{s2})$. The other correlations can be written in terms of $C_\ell^{\kappa\kappa}$ as (this is strictly valid only in flat-sky approximation (Kilbinger et al., 2017))

$$C_\ell^{\gamma\gamma *} = C_\ell^{\kappa\kappa} \tag{1.57}$$

$$C_\ell^{\gamma\gamma} = C_\ell^{\kappa\kappa} e^{i4\phi_k} \tag{1.58}$$

$$C_\ell^{\kappa\gamma} = C_\ell^{\kappa\kappa} e^{i2\phi_k} \tag{1.59}$$

Correlation functions under flat sky approximation can be written using Hankel transforms

$$\langle\kappa\kappa\rangle(\theta) = \langle\gamma\gamma *\rangle(\theta) = \int \frac{d\ell\ell}{2\pi} C_\ell^{\kappa\kappa} J_0(\ell\theta) \tag{1.60}$$

$$\langle\gamma\gamma\rangle(\theta) = \int \frac{d\ell\ell}{2\pi} C_\ell^{\kappa\kappa} J_4(\ell\theta) \tag{1.61}$$

$$\langle\kappa\gamma\rangle(\theta) = \int \frac{d\ell\ell}{2\pi} C_\ell^{\kappa\kappa} J_2(\ell\theta) \tag{1.62}$$

Lensing auto correlations measurements have been performed now by several studies. Convergence auto correlations from CMB lensing have been measured by Das et al. (2011, 2014); van Engelen et al. (2012); Planck Collaboration et al. (2015b), with current state of the art measurements by Planck Collaboration et al. (2015b) measuring convergence auto correlations at $\sim 40\sigma$ levels and constraining the amplitude of matter fluctuations at $\sim 3\%$ level (lensing alone).

Shear auto correlations, cosmic shear, have also been detected by a number of surveys, CFHTlens (Heymans et al., 2013a; Joudaki et al., 2017), KiDS (Hildebrandt et al., 2016), DES (Abbott et al., 2016). Some of these measurements are in mild tension with Planck$\Lambda$CDM model, with weak lensing preferring lower amplitude (Heymans et al., 2013a; Hildebrandt et al., 2016). Recent re-analysis of CFHTlens by (Joudaki et al., 2017) found that some of the tension can be elevated by marginalizing over conservative models for systematics (see also Efstathiou & Lemos, 2017). With the increasing precision of upcoming results from DES and HSC surveys and even higher precision from LSST, WFIRST, Euclid in future, consistency tests between lensing and other cosmological probes will become more stringent and modeling of systematics will also need to be better to distinguish between effects of systematics and incomplete cosmological models.

A way of testing for systematics is to measure cross correlations between convergence from CMB lensing and galaxy shear, which are independent of several additive systematics in both CMB lensing and galaxy shear. These cross correlations have been detected by Hand et al. (2015); Kirk et al. (2016); Liu & Hill (2015); Harnois-Déraps et al. (2016). Galaxy shear and CMB lensing have very different lensing kernels with CMB lensing being most sensitive to $z \in [1, 5]$ while galaxy shear from current surveys being sensitive to $z < 1$. The cross correlations between the two probe intermediate regions where auto correlations of either probe are not very sensitive. However, since one or both lensing kernels are rather low in this regime, these cross correlation measurements are also noisy. Most

of the current measurements are consistent with Planck$\Lambda$CDM model predictions though statistical uncertainties are rather high. We will present our measurement of cross correlations between SDSS shear and Planck CMB lensing maps in chapter 2. SDSS lensing kernel is sensitive to only $z \lesssim 0.4$ where CMB lensing kernel is quite low. This will increase the statistical noise in our measurements compared to other similar studies, though SDSS and Planck overlap area is also larger than used in other works which compensates for the lost precision from redshift kernel.

### Galaxy lensing cross correlations

Cross correlations between galaxy positions and lensing maps provide a powerful probe to map the dark matter distribution around galaxies and recover the information about matter correlations on cosmological scales. These measurements are also independent of additive systematic biases in lensing maps. Measurements of galaxy-lensing cross correlations and their applications are the main focus of our work in chapters 2–4.

We measure the cross-correlations in real-comoving space as

$$\langle \Sigma_c g\kappa \rangle(r_p) = \overline{\rho}_m b_g \int d\chi \int d^3k e^{ik\cdot r} P(k) \approx \overline{\rho}_m b_g \int dk_\perp J_0(k_\perp r_p) P(k_\perp) = \overline{\rho}_m w_{gm} = \Sigma(r_p) \tag{1.63}$$

$$\langle \Sigma_c g\gamma \rangle(r_p) = \overline{\rho}_m b_g \int d\chi \int d^3k e^{ik\cdot r} P(k) e^{i2\phi_k} \approx \overline{\rho}_m b_g \int dk_\perp J_2(k_\perp r_p) P(k_\perp) = \overline{\rho}_m w_{gm+} = \Delta\Sigma(r_p) \tag{1.64}$$

Where we converted shear to projected matter density using $\Sigma_c$ factor, $r_p$ is the physical separation between lens and source at lens redshift and to get the approximate expression, we used the fact that lensing line of sight window is broad and we can use limber approximation to write $P(k) \approx P(k_\perp)$.

An important application of galaxy-lensing cross correlations is to combine them with galaxy clustering and obtain matter power spectrum independent of the galaxy bias. The matter correlation function can be recovered using the relation (Seljak et al., 2005; Baldauf et al., 2010)

$$\xi_{mm} = \frac{\xi_{gm}^2}{\xi_{gg} r_{cc}} \tag{1.65}$$

Where $r_{cc}$ is the galaxy-matter cross-correlation coefficient and is $\sim 1$ at large scales. At smaller scales it can deviate significantly from 1 due to stochasticity in galaxy-matter correlations. Exact effect of $r_{cc}$ also depends on the estimator used, with $\Delta\Sigma$ being more affected by $r_{cc}$ due to information from small scales affecting larger scales (see Baldauf et al., 2010, for detailed discussion). Cosmological analysis using galaxy-lensing cross correlations have been performed by Mandelbaum et al. (2013) using estimator suggested by Baldauf et al. (2010) to remove small scale information. Other studies, (e.g. More et al., 2015) have also used HOD modeling to use the clustering and lensing information down to small scales.

Galaxy-lensing cross correlations can also be used to construct powerful tests for theories of gravity. As shown in eq. 1.36, the lensing potential depends on the newtonian and curvature potentials. The two potentials are equal in the General theory of relativity in the absence of anisotropic stresses, which is true within $\Lambda$CDM model as long as effects of relativistic species , e.g. neutrinos, are subdominant. The growth of structure depends only on the newtonian potential and thus the combination lensing with the other probes of growth can be used to test the equality of two potentials. Zhang et al. (2007) suggested one such test, the $E_G$ parameter, measuring ratio of galaxy-lensing cross correlations and galaxy-velocity cross correlations (see also Leonard et al., 2015). $E_G$ measurements have been performed by Reyes et al. (2010); Blake et al. (2016); Pullen et al. (2016); Alam et al. (2017), with results being consistent with $\Lambda$CDM predictions, except for Pullen et al. (2016), who found a $2.6\sigma$ discrepancy with the Planck$\Lambda$CDM model.

Since galaxy-lensing cross correlations map the matter distribution around galaxies, they are also an important probe of relations between galaxies and their surrounding dark matter halos. Several galaxy-galaxy lensing measurements have probed the connection between galaxies and their host dark

matter halos (e.g. Mandelbaum et al., 2006; Sifón et al., 2015a; van Uitert et al., 2016; Mandelbaum et al., 2016). Similarly, cross-correlations between weak lensing and clusters are also used to study galaxy clusters, measure their masses and also calibrate other methods to obtain masses (e.g. Hoekstra et al., 2015; Simet et al., 2017). Cluster masses can then be used to put constraints on cosmological models as number of observed clusters in a given mass range depend on the growth of structure (see Allen et al., 2011, for review).

Galaxy-lensing measurements also provide tests for several systematics in the lensing measurements, since lens galaxy positions are not expected to correlated with many of the lensing systematics. Most notably, galaxy-lensing cross correlations are independent of additive systematics in the estimation of shear from galaxy shapes. Comparison of galaxy-lensing measurements using same galaxy positions and different lensing maps, e.g. galaxy lensing and CMB lensing, can also help to estimate the relative biases, including hard to estimate multiplicative biases between the two maps (Schaan et al., 2016).

### 1.3.4 Systematics

As stated in section 1.3.3, some weak lensing measurements are in mild tension with results from Planck CMB measurements, which can potentially be caused by some systematics in the lensing measurements. Here we highlight few of these systematics, resulting from biases in shear estimation from galaxy shapes and photometric redshifts. We will discuss intrinsic alignments of galaxies, which introduces a non-lensing coherent shear, in section 1.3.5. Our discussion of systematics is not complete as there are other systematics, such as non-linear physics and impact of baryons which impact inferences from small scale correlations. We refer the reader to Joudaki et al. (2016) who tested the impact of several of these systematics for more details.

#### Shape measurements

Estimating shear from galaxy shapes is a non-trivial problem as, 1) Galaxies have a wide range of complicated morphologies (shape and structure), 2) the observed shapes from a telescope are convolved with point spread function (PSF) due atmosphere (ground based observations only) and instrumental effects, 3) Measurements are noisy.

The rich diversity in galaxy morphologies makes it harder to construct general models for galaxy images. This can be challenging for methods that use simple morphological models to fit for the galaxy shapes which may not capture all features in a given galaxy shape leading to wrong shear estimates. Even for general moments based methods not using morphological models, accurate estimate of galaxy morphologies is required to derive the shear responsivity factors (eq. 1.54) for the population.

The effects of PSF are two fold. It blurs the galaxies, making them rounder and in case of anisotropic PSF, can also shear the galaxies in a preferential direction. While PSF can be estimated from stars in the observational field, removing effects of PSF is a hard problem. If there are only few stars in the observational field, the interpolation of PSF may not be able to capture spatial variation of PSF across the whole field. Further, PSF can depend on the wavelength of light and since galaxies and stars have different spectra, PSF for galaxies can be biased. Since PSF blurs the objects, getting morphological information for galaxies smaller than the PSF size is hard, which introduces a limit on the size of galaxies that can be used. Finally, since measurements of galaxies and PSF are noisy, deconvolution is hard in the limit of low S/N galaxies and if noise is not accounted for correctly, noise biases can be severe.

Current strategies (atleast for ground based surveys) to overcome these problems are to use deep space based observations to learn galaxy morphologies and then construct larger simulations containing realistic galaxies (see GREAT3 challenge handbook Mandelbaum et al. (2014)). The biases in shear estimate are usually quantified as multiplicative bias, $m$ (making galaxies appear rounder or more elliptical) and additive bias, $c$, which shears galaxies in some preferred direction. The observed shear

component are then related to true components as

$$\widehat{\gamma}_i = (1+m)\gamma_i + c_i \tag{1.66}$$

For our measurements, we use shape measurements that already exist from Reyes et al. (2012). While these measurements are already well characterized and calibrated, we do perform several checks for systematics which directly relate to shape measurements. These shape measurements are known to have some additive systematics caused by PSF anisotropies. We estimate the impact of these additive biases by measuring shear around random points on the sky (or using random realizations of galaxy sample) and then remove them by subtracting the measurement around randoms from galaxy-lensing cross correlation signal. Multiplicative biases in shear are harder to estimate as they only change the amplitude of measurements and are de-generate with the cosmological parameters (especially amplitude of correlations, $\sigma_8$). In chapter 2, we use CMB lensing to search for relative multiplicative biases between galaxy lensing and CMB lensing. In chapter 6, in context of intrinsic alignments, we also perform several null tests on measured shapes, including two additional methods, de Vaucouleurs and isophotal shapes, which have been used to measure intrinsic alignments in other works.

### Photometric redshifts

Weak lensing measurements require shape measurements from a large number of source galaxies. Interpretation of shear measurements also require an estimate of line of sight distance to these galaxies. Obtaining spectroscopy to get redshift estimates for all galaxies in a photometric lensing survey is prohibitively expensive with the current technology. Instead we attempt to obtain noisy photometric redshift estimates from the broadband measurements. Essentially, the photometric redshifts are estimates from *extremely* low resolution spectra. In ground based surveys photometric observations are done in $\sim 5$ bands and differences of flux (or magnitudes) in these bands (colors hereafter) are mapped to redshift. This mapping works best for galaxies with spectral features that lead to large differences in flux across different bands (one such feature is $4000\mathring{A}$ feature in red galaxies). There are several algorithm based on machine learning or template fitting to estimate photometric redshifts, but they require a representative calibration sample with spectroscopic redshifts for training (Feldmann et al., 2006; Csabai et al., 2007; Nakajima et al., 2012; Hildebrandt et al., 2010), which can also be hard to obtain for deep samples. These methods also suffer from degeneracies as different spectral features at different redshifts can look the same to the algorithm.

Another method that attempts to overcome these limitations is to cross correlate the photometric galaxies with a spectroscopic sample (Newman, 2008). Since the clustering of galaxies is a local phenomenon, photometric galaxies will show clustering only with nearby galaxies and thus we can get an estimate of redshift distribution of these galaxies. This method still requires a spectroscopic sample that spans the entire redshift range of the photometric sample, though a sparse sample will work. A major uncertainty in this method is the estimation of galaxy bias for photometric sample which affects the clustering amplitude. In principle galaxy bias can be constrained by auto correlations of photometric sample, but the complicated selection function of photometric galaxies across the sky and along the line of sight can introduce modes in the auto correlations and severely bias the clustering measurements. Another uncertainty in clustering measurements is that lensing effects can also correlate the distribution of background photometric galaxies with the foreground spectroscopic sample.

For many weak lensing applications, we do not require accurate per object photometric redshifts but an accurate estimate of the redshift distribution of galaxies. For the case of tomographic measurements and galaxy lensing measurements, it is useful to have good estimates of per object redshifts to select galaxies in some redshift bins or galaxies behind a lens. While photo-$z$ bias and scatter can both affect the lensing measurements, it is possible to calibrate for them if bias and scatter can be estimated. Calibrating photometric redshifts with any of the aforementioned methods though requires hard to obtain deep spectroscopic samples.

We will use the photometric redshifts from Nakajima et al. (2012) obtained using template fitting method (Feldmann et al., 2006) and calibrated using representative but smaller area spectroscopic

samples (see Nakajima et al., 2012, for more details). In chapter 4, we test the calibrations obtained by Nakajima et al. (2012) using the clustering photo-$z$ method by cross correlating photometric sample with spectroscopic samples (LOWZ+CMASS) from SDSS-BOSS survey.

### 1.3.5 Intrinsic Alignments

Intrinsic alignments (IA) of galaxy shapes are an important systematic in the weak lensing measurements (Hirata & Seljak, 2004; Krause et al., 2016; Chisari et al., 2015a). These alignments lead to violation of the assumption that mean ellipticity of galaxies is zero, $\langle e_i \rangle = 0$, in the absence of lensing shear. In the presence of IA, we can write the observed shear as sum of lensing shear $\gamma_G$ and an intrinsic shear $\gamma_I$ and the observed two-point functions get modified as

$$\langle \gamma\gamma \rangle = \langle \gamma^G\gamma^G + \gamma^G\gamma^I + \gamma^I\gamma^I \rangle = \xi_{GG} + \xi_{GI} + \xi_{II} \tag{1.67}$$

$$\langle g\gamma \rangle = \langle g\gamma^G + g\gamma^I \rangle = \xi_{gG} + \xi_{gI} \tag{1.68}$$

Where $G$ refers to lensing shear and $g$ refers to galaxy positions. $GG$ and $gG$ are the terms of interest for lensing measurements while the other terms are the contaminants. $II$ term is the auto correlation of intrinsic shear, affecting cosmic shear measurements when two physically nearby source galaxies are correlated. This can be removed by removing source galaxy pairs that are close to each other, though it is hard to do with photometric redshifts. $GI$ term is caused by two source galaxies at different redshifts that are correlated. In this case the foreground galaxy is intrinsically aligned with the large scale structure that also lenses the background galaxies. $gI$ is similarly caused in galaxy-galaxy lensing when lens and source galaxies are nearby and source galaxy is intrinsically aligned with the lens instead of being lensed. Most current lensing surveys account for IA contamination by including the IA model at the time of inference, though there are other strategies such as nulling (Joachimi & Schneider, 2008) which attempt to exploit different redshift kernel of IA and lensing and remove IA at the time of measurement.

For the current generation of surveys, IA contamination has not been found to be a significant source of systematic and attempts to detect IA using the full galaxy samples have resulted in null detections (Blazek et al., 2012; Chisari & Dvorkin, 2013; Heymans et al., 2013a; Joudaki et al., 2016). However, for next generation surveys such as LSST, Krause et al. (2016) have shown that IA can significantly bias the inferences from lensing measurements if not accounted for.

The IA of galaxies depends intricately on the processes involved in the galaxy formation and evolution, such as tidal forces from large scale structure, accretion of matter along filaments, merger of galaxies. IA are also known to depend on the galaxy types, with several robust detections for alignments of elliptical galaxies (Mandelbaum et al., 2006; Hirata et al., 2007; Okumura et al., 2009; Joachimi et al., 2011; Blazek et al., 2012) while IA have not been detected yet for the case of spiral galaxies (see Mandelbaum et al., 2011, for a case of null detection).

For the elliptical galaxies, the galaxy shapes are assumed to be preferentially aligned with the dark matter halo shapes with some scatter (misalignment angle) and halo shapes are aligned with the local large scale structure. The halo alignments are caused by the effects of tidal fields from large scale structure, in which case we can write expression for intrinsic shear as (Catelan et al., 2001)

$$\gamma^I = (\gamma^I_+, \gamma^I_\times) = -\frac{C_1}{4\pi G}(\partial_x^2 - \partial_y^2, \partial_x\partial_y)\phi_p \tag{1.69}$$

Note that expression for intrinsic shear in eq. 1.69 is very similar to the tangential lensing shear though they have opposite signs (IA is radial while lensing is tangential).Hirata & Seljak (2004) further developed the linear alignment model for IA and showed that the galaxy-intrinsic shear cross correlation can be written in analogous expression to galaxy-galaxy lensing

$$w_{g+} = A_I b C_1 \rho_{\rm crit}\Omega_m \int \frac{{\rm d}k_\perp k_\perp}{2\pi} P(k_\perp) J_2(k_\perp r_p) \tag{1.70}$$

Where we introduced $A_I$ as amplitude of IA, $b$ is the galaxy bias and $C_1$ is a constant conventionally set such that $C_1\rho_{\rm crit} = 0.0134$ (Joachimi et al., 2011). Similar expression can be written for IA-IA and lensing-IA correlations. An important point to note is that IA is a local process and hence is sensitive to the local potential around galaxies while lensing is sensitive to entire potential between source and observer. Bridle & King (2007) showed that linear alignment model can be extended to quasi-linear scales by using the non-linear matter power spectrum. Schneider & Bridle (2010) also developed the halo model of IA assuming that central galaxies are aligned with the halo shapes while satellite galaxies are radially aligned. Radial alignment of satellites can also be viewed in the context of tidal alignments, as tidal fields within a halo are predominantly radial. Blazek et al. (2015) showed that for sample with low satellite fraction, IA observations can be modeled by simple NFW profiles. A full halo model for IA has not been developed at the time of writing.

There are several observations of galaxy alignments in the the literature using different statistics and shape measurement methods (for example Mandelbaum et al., 2006; Hirata et al., 2007; Okumura et al., 2009; Joachimi et al., 2011; Blazek et al., 2011; Hao et al., 2011; Hung & Ebeling, 2012; Li et al., 2013; Chisari et al., 2014; Sifón et al., 2015b; Singh et al., 2015). Measurements of two point functions of IA for elliptical galaxies show scaling consistent with the predictions from tidal alignment models (Hirata et al., 2007; Joachimi et al., 2011; Blazek et al., 2012), while the amplitude evolves with galaxy magnitude, with brighter galaxies showing stronger alignments (Hirata et al., 2007; Joachimi et al., 2011). These observations are also consistent with other studies using different statistics, such as measuring alignment angles, where brighter and redder galaxies have stronger alignments (see e.g. Zhang et al., 2013). Since brighter and redder galaxies are also known to form in relatively overdense regions (they have higher clustering bias also), the evolution of amplitude with luminosity is consistent with the expectations from tidal alignment model since tidal fields are stronger in overdense regions. Small scale processes such as galaxy mergers, accretion along preferred directions (e.g. from filaments), also affect the IA. Thus IA observations can help in studying the importance of different processes, making them an interesting probe of galaxy physics as well.

The IA observations for elliptical galaxies are qualitatively consistent with the results from simulations (see e.g. Croft & Metzler, 2000; Heymans et al., 2006; Faltenbacher et al., 2008; Kang et al., 2007; Kuhlen et al., 2007; Joachimi et al., 2013a). In dark matter only simulations, general consensus has been that the halo shapes are aligned with the large scale structure, with more massive halos showing stronger alignments. Alignments and ellipticities are also observed to evolve with the radii, with outer regions of halos being strongly aligned Kuhlen et al. (2007). This again can be explained by the tidal alignment model as the outer shapes of halos are more susceptible to the tidal fields and hence are more strongly aligned (Kormendy, 1982). Subhalos are usually aligned radially, though near pericenter alignments can be tangential where tidal torque is not able to align the rapidly accelerating halo (Pereira et al., 2008).

To relate the galaxy alignments with halo alignments, we can use semi-analytic models to populate the halos with galaxies and assume that the galaxy shapes are aligned with the halo shapes with some random misalignment angles. Such approach has been used by (Faltenbacher et al., 2007; Okumura et al., 2009; Joachimi et al., 2013a) to study alignments, finding that one needs misalignment angles of order 35°-40° to explain the observations.

Hydrodynamic simulations provide another way to study the alignments of galaxies and also test the influence on alignments of various processes involved in galaxy physics (Tenneti et al., 2014b,a, 2015, 2016, 2017; Chisari et al., 2015b, 2016, 2017; Velliscig et al., 2015b). While Hydro simulations are also qualitatively consistent with the picture of IA presented above, there are notable inconsistencies in results from different simulations. Galaxy shapes are aligned with the shapes of host halos but the distribution of misalignment angles has been found to depend on the details of physical mechanisms put into the simulations (Tenneti et al., 2016, 2017; Velliscig et al., 2015b). The scaling of two point functions is consistent with the predictions from tidal alignment models, though there appears to be some discrepancy with the observations in terms of the amplitude. Tenneti et al. (2014a) found alignments in simulations to be stronger than compared to observations while Chisari et al. (2016) finding alignments to be weaker in simulations. There is also no consensus yet on redshift

evolution of the alignments, with Tenneti et al. (2014a) finding no significant evolution of redshift evolution while Chisari et al. (2016) observed redshift evolution though the results depend on shape measurement method used. Another notable discrepancy between different hydro simulations has been the differences in alignments of disk galaxies, with different simulations giving different amplitude and different signs of alignments (Tenneti et al., 2016).

## 1.4 Thesis Structure

In chapter 2 and appendix 4.B we will present new measurements of galaxy-lensing cross-correlations and galaxy clustering and their applications to test the cosmological models and sources of systematics errors. These measurement use BOSS spectroscopic galaxy samples (Reid et al., 2016) and galaxy shear measurements from SDSS Reyes et al. (2012) and CMB lensing maps from Planck Planck Collaboration et al. (2015b). Using these measurements we will perform cosmological analysis similar to Mandelbaum et al. (2013) (work in progress) and measure the halo mass of galaxies. The halo mass measurements will also be used in chapter 5 to study the dependence of intrinsic alignments of galaxy shapes with halo mass. Comparing the lensing correlations using galaxy shear and CMB lensing, we measure distance ratios and test for relative calibration biases between CMB and galaxy lensing.

In chapter 3, we test different methods of estimating covariance for galaxy-lensing cross correlations and perform a detailed study on sources of noise in galaxy-lensing cross correlations. We also derive expressions for covariance for the popular estimators used for galaxy-lensing cross correlations.

In chapter 4, we present constraints on the $E_G$ parameter using the measurements from chapter 2. We will also perform tests for several systematics in the measurements, including the modeling uncertainties in $r_{cc}$ parameter.

Chapters 5 and 6 contain a detailed study of intrinsic alignments of galaxy shapes. In chapter 5, we present our measurements of IA using BOSS LOWZ galaxies. The LOWZ sample allows us to extend the IA measurements to lower luminosities compared to previous work. Using these measurements we will test the non-linear alignment model and the evolution of IA amplitude with various galaxy properties. We also extend to IA measurements to smaller scales and probe IA within 1-halo regime using the halo model of Schneider & Bridle (2010). In chapter 6, we further extend this work, testing the validity of linear alignment model in two dimensional space (separation perpendicular and parallel to line of sight, $r_p$ and $\Pi$) when using projected galaxy shapes in redshift space. We also perform IA measurements with different shape measurement methods and different estimators used in literature.

All the chapters from chapter 2–6 are published papers in MNRAS journal, with the exception of chapter 4, which at the time of writing is ready for submission after a review by collaborators. Chapter 2 is published as Singh et al. (2016a), chapter 3 published as Singh et al. (2016b), chapter 5 published as Singh et al. (2015) and chapter 6 published as Singh & Mandelbaum (2016). Since these works are published, I will be including them in this thesis with minimal changes. Hence each chapter will be a self contained work with its own introduction and conclusions which describe the problems being addressed in that work in more detail than has been covered in this chapter. A brief summary of our conclusions will be presented in chapter 7. I will also ignore the general convention of putting appendices at the end of the thesis. Each chapter will have its own appendices, consistent with the stated aim of keeping every chapter as a self contained work.

# 2 Cross-correlating Planck CMB lensing with SDSS: Lensing-lensing and galaxy-lensing cross-correlations

Sukhdeep Singh[1], Rachel Mandelbaum[1], Joel R. Brownstein[2]

[1]McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[2]Department of Physics and Astronomy, University of Utah, 115 S 1400 E, Salt Lake City, UT 84112, USA

**Abstract**

We present results from cross-correlating Planck CMB lensing maps with the Sloan Digital Sky Survey (SDSS) galaxy lensing shape catalog and BOSS galaxy catalogs. For galaxy position vs. CMB lensing cross-correlations, we measure the convergence signal around the galaxies in configuration space, using the BOSS LOWZ ($z \sim 0.30$) and CMASS ($z \sim 0.57$) samples. With fixed Planck 2015 cosmology, doing a joint fit with the galaxy clustering measurement, for the LOWZ (CMASS) sample we find a galaxy bias $b_g = 1.75 \pm 0.04$ ($1.95 \pm 0.02$) and galaxy-matter cross-correlation coefficient $r_{\rm cc} = 1.0 \pm 0.2$ ($0.8 \pm 0.1$) using $20 < r_p < 70 h^{-1}$Mpc, consistent with results from galaxy-galaxy lensing. Using the same scales and including the galaxy-galaxy lensing measurements, we constrain $\Omega_m = 0.284 \pm 0.024$ and relative calibration bias between the CMB lensing and galaxy lensing to be $b_\gamma = 0.82^{+0.15}_{-0.14}$. The combination of galaxy lensing and CMB lensing also allows us to measure the cosmological distance ratios (with $z_l \sim 0.3$, $z_s \sim 0.5$) $\mathcal{R} = \frac{D_s D_{l,*}}{D_* D_{l,s}} = 2.68 \pm 0.29$, consistent with predictions from the Planck 2015 cosmology ($\mathcal{R} = 2.35$). We detect the galaxy position-CMB convergence cross-correlation at small scales, $r_p < 1 h^{-1}$Mpc, and find consistency with lensing by NFW halos of mass $M_h \sim 10^{13} h^{-1} M_\odot$. Finally, we measure the CMB lensing-galaxy shear cross-correlation, finding an amplitude of $A = 0.76 \pm 0.23$ ($z_{\rm eff} = 0.35$, $\theta < 2^\circ$) with respect to Planck 2015 $\Lambda$CDM predictions ($1\sigma$-level consistency). We do not find evidence for relative systematics between the CMB and SDSS galaxy lensing.

## 2.1 Introduction

As photons travel from the source towards the observer, their paths are distorted by the gravitational potential of the intervening matter. This phenomenon, known as gravitational lensing, has become an important tool in cosmology to study the growth of structure in the dark matter distribution as well as cosmic acceleration (Weinberg et al., 2013). In the weak regime, the gravitational lensing introduces small but coherent distortions in the shapes of background galaxies, which can be measured through correlations of galaxy images. Weak gravitational lensing also remaps the CMB anisotropies, blurring the acoustic peaks and correlating different modes, which can then be exploited to generate mass maps from the CMB observations with good resolution and signal-to-noise ratio (Zaldarriaga & Seljak, 1999; Hu & Okamoto, 2002; Lewis & Challinor, 2006).

CMB lensing is most sensitive to structure at high redshifts, $z = 1 - 5$, and thus provides a unique probe to study the structure in the dark matter distribution at these high redshifts. The lensing of CMB has been robustly detected by several CMB experiments including the Atacama Cosmology Telescope (ACT; Das et al., 2011, 2014), the South Pole Telescope (SPT; van Engelen et al., 2012), the Planck telescope (Planck Collaboration et al., 2014b, 2015b), The POLARBEAR experiment (Ade et al., 2014) and the BICEP2/Keck array (Keck Array et al., 2016), with measurements being consistent with the $\Lambda$CDM predictions. Cross-correlating CMB lensing maps with galaxy surveys provides opportunities to probe the large-scale structure (LSS), calibrate different probes of large-

scale structure (particularly galaxy lensing) and carry out consistency tests. Several studies cross-correlating CMB lensing with galaxy position catalogs as well as galaxy lensing catalogs have already been performed.

Given that CMB lensing and galaxy lensing have very different lensing kernels, cross-correlating the two can help constrain the amplitude of matter fluctuations at low redshift (whereas CMB lensing auto-correlations are not very sensitive to this). The past year has seen a number of detections of this effect in several surveys (Hand et al., 2015; Kirk et al., 2016; Liu & Hill, 2015; Harnois-Déraps et al., 2016), typically in Fourier space but in the last work, in configuration space as well. Assuming a fixed cosmological model, these cross-correlations also provide a test for relative calibration biases between the two lensing maps (Vallinotto, 2012; Das et al., 2013). The results of recent work has been largely consistent with ΛCDM predictions, with at most slight tension ($\sim 2\sigma$) that has at times (e.g., Liu et al., 2016) been interpreted as residual systematics in the galaxy lensing.

Cross-correlating CMB lensing with galaxy positions also provides a probe of structure growth and of the matter distribution around galaxies. In addition, when compared with galaxy-galaxy lensing, such correlations can also provide a handle on systematics such as biases in photometric redshift distributions (de Putter et al., 2014). Many cross-correlations of galaxy positions with CMB lensing exist in the literature (Smith et al., 2007; Hirata et al., 2008; Bleem et al., 2012; Sherwin et al., 2012; Planck Collaboration et al., 2014b; Giannantonio & Percival, 2014; Bianchini et al., 2015; Pullen et al., 2016; Giannantonio et al., 2016). As for the lensing-lensing cross-correlation, the results of measurement of the galaxy position vs. CMB lensing cross-correlation are typically consistent with ΛCDM predictions, at times with low-level ($2\sigma$) tension. For example, Pullen et al. (2015a) used the cross correlations between BOSS CMASS galaxies and CMB lensing in combination with galaxy clustering and velocity measurements to constrain the theory of gravity on large scales. They found $\sim 2\sigma$ deviations from ΛCDM predictions, with the discrepancy being primarily driven by the lower amplitude of galaxy position vs. CMB lensing cross-correlations at $r_p \gtrsim 80h^{-1}$Mpc. These cross-correlation measurements provide a valuable consistency check when compared with results using galaxy lensing.

Miyatake et al. (2016) measured the ratio of CMB lensing and galaxy lensing signal around the BOSS CMASS galaxies (Alam et al., 2015a), using Planck 2015 CMB lensing maps (Planck Collaboration et al., 2015b) and galaxy lensing sources from CFHTLens (Erben et al., 2013). Such a ratio depends on the geometric factors involving the distance between the observer, the lens galaxies, the CMB last scattering surface and the source galaxies, providing a measurement of cosmic distance ratios (Hu et al., 2007a) and hence a strong consistency check on the cosmological model.

Most of the CMB lensing vs. galaxy position cross-correlation measurements have been performed at large scales. Current and next-generation CMB surveys will have sufficiently high resolution and low noise levels to measure the lensing signals even on the scales of dark matter halos, and provide mass constraints (Hu et al., 2007a). This has the potential to be particularly powerful at the high redshifts that are beyond the reach of galaxy lensing surveys. The first such measurement has already been performed by Madhavacheril et al. (2015) by cross-correlating CMB lensing maps from ACTPol with CMASS sample galaxies from the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS; Alam et al., 2015a). Similar measurements using more massive SZ selected clusters have also been performed by Baxter et al. (2015) and Planck Collaboration et al. (2015c).

In this work, we perform the galaxy-CMB lensing cross-correlations using SDSS-III BOSS galaxies (Alam et al., 2015a). We compute the projected matter density $\Sigma$ in real comoving space by stacking the convergence obtained from Planck 2015 lensing maps around the positions of BOSS galaxies. Using the low redshift sample, LOWZ, we also present a direct comparison between the results from galaxy lensing and CMB lensing, and test for relative calibration biases between the two lensing signals. Using the combination of CMB lensing and galaxy lensing, we also present a measurement of the cosmic distance ratio at an effective lens redshift of 0.26. In addition, we also cross-correlate the two lensing measurements in configuration space. The SDSS lensing source sample is at relatively low redshift compared to other galaxy lensing surveys, where the rapid decrease of the CMB lensing kernel reduces the amplitude of the cross-correlation signal. However, the large sky area of the SDSS

($\gtrsim$ 8000 square degrees) compared to other existing lensing surveys compensates for the relatively lower expected signal, particularly in the case where noise in the CMB lensing maps dominates the statistical error budget.

This paper is organized as follows: In Sec. 2.2, we discuss the theoretical background and the estimators used in our measurements. In Sec. 2.3, we describe the datasets used in this work. Our results are in Sec. 2.4, and we conclude in Sec. 2.5. Throughout, we use the Planck 2015 cosmology (Planck Collaboration et al., 2015a), with $\Omega_m = 0.309$, $n_s = 0.967$, $A_s = 2.142 \times 10^{-9}$, $\sigma_8 = 0.82$. To compute theoretical predictions in this paper, we use the linear+halofit (Smith et al., 2003; Takahashi et al., 2012) matter power spectrum generated using the CAMB software (Lewis & Bridle, 2002b).

## 2.2 Formalism and Methodology

In this section, we present the theoretical models and estimators used to carry out and interpret the measurements.

### 2.2.1 Weak Lensing Introduction

Here we provide a very brief review of weak lensing, and refer the reader to Bartelmann & Schneider (2001) for details. Gravitational lensing measurements are sensitive to the lensing potential, defined as

$$\Phi_L = \int \mathrm{d}\chi_l \frac{f_k(\chi_s - \chi_l)}{f_k(\chi_s) f_k(\chi_l)} \Psi(f_k(\chi_l)\theta, \chi_l) \tag{2.1}$$

where the Weyl Potential $\Psi = \psi + \phi$, $\phi$ and $\psi$ are the Newtonian and curvature potentials, $\theta$ is the angular coordinate on the sky, $\chi_s, \chi_l$ are comoving radial distances to source and lens respectively and $f_k$ are the generalized (not assuming flatness) transverse comoving distances (in the case of a flat universe, $f_k(\chi) = \chi$). Within $\Lambda$CDM, $\phi = \psi$ and $\nabla^2\phi = 4\pi G\rho_m$. When source size is smaller than the angular scales over which lens properties change, the relation between source and image positions can be linearized and the Jacobian of the image to source transformation is (Bartelmann & Schneider, 2001)

$$A_{ij} = \frac{\partial(\theta_o^i - \delta\theta_o^i)}{\partial\theta_o^j} \tag{2.2}$$

$$A_{ij} = \delta_{ij} - \frac{\partial^2\Phi_l}{\partial\theta_o^i \partial\theta_o^j} \tag{2.3}$$

where the subscript $o$ indicates that the derivative is carried out in observer (or image) coordinates. The matrix $A$ is

$$A = \begin{bmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{bmatrix}.$$

In the weak gravitational lensing regime, the primary observables are the convergence $\kappa$ for CMB lensing and the shear $\gamma = \gamma_1 + \mathrm{i}\gamma_2 = |\gamma| e^{2i\theta}$ for galaxy lensing. The measured $\gamma$ can be rotated into the lens-source frame, $\gamma = \gamma_t + i\gamma_\times$, where $\gamma_t$ is the shear along the line joining the lens and source galaxy while $\gamma_\times$ is the shear with respect to the 45° lines. These observables relate to the underlying gravitational potential as

$$\kappa = \frac{1}{2}\nabla_\perp^2 \Phi_L \tag{2.4}$$

$$\gamma_t = \frac{1}{2}(\nabla_{x,x}^2 - \nabla_{y,y}^2)\Phi_L \tag{2.5}$$

$$\gamma_\times = \frac{1}{2}(\nabla_{x,y}^2)\Phi_L \tag{2.6}$$

where the derivatives are with respect to the plane-of-sky coordinates ($x$ is along the line joining the lens and source positions, and $y$ is orthogonal to $x$). Note that the lensing observables are sensitive only to the matter density contrast ($\bar{\rho}_m\delta = \rho_m - \bar{\rho}_m$) and not the mean matter density. The equivalent expressions for convergence and shear in the Fourier space are

$$\tilde{\kappa} = \frac{1}{2}k_\perp^2\tilde{\Phi}_L \tag{2.7}$$

$$\tilde{\gamma}_t =\equiv \frac{1}{2}(k_{\perp,x}^2 - k_{\perp,y}^2)\tilde{\Phi}_L \tag{2.8}$$

$$\tilde{\gamma}_\times = \frac{1}{2}(k_{\perp,x}k_{\perp,y})\tilde{\Phi}_L, \tag{2.9}$$

### 2.2.2 Lensing-lensing cross-correlation

When cross-correlating CMB lensing with galaxy shear, we will be measuring the tangential shear $\gamma_t$ in the galaxies around each pixel of the CMB map, and weight that shear with the CMB convergence value $\kappa$ within that pixel. Under the Limber approximation, using expressions for $\kappa$ and $\gamma_t$ from Eqs. (2.5) and (2.6), we can write the cross-correlation function for shear and convergence as

$$\begin{aligned}\langle\kappa\gamma\rangle(\theta) = &\int \mathrm{d}z_\kappa p(z_\kappa)\int \mathrm{d}z_\gamma p(z_\gamma)\\ &\int \mathrm{d}z_l \frac{H(z_l)}{c}W_L(\chi_\kappa,\chi_l)W_L(\chi_\gamma,\chi_l)\\ &\int \frac{\mathrm{d}k}{2\pi}kP_{\delta\delta}(k)J_2\left[kf_k(\chi_l)\theta\right],\end{aligned} \tag{2.10}$$

where $p(z_\kappa)$ and $p(z_\gamma)$ are the redshift distribution of the source samples used to measure the convergence and shear respectively. In the case that $\kappa$ is measured from CMB lensing, $p(z_\kappa) = \delta_D(z_\kappa - 1100)$. $\gamma_t$ is the galaxy tangential shear defined with respect to the line joining the galaxy with the pixel center of $\kappa$ map. The lensing weight $W_L$ is defined as

$$W_L(\chi_s,\chi_l) = \frac{3}{2}\frac{H_0^2}{c^2}\Omega_{m,0}(1+z_l)\frac{c}{H(z_l)}\frac{f_k(\chi_l)f_k(\chi_s-\chi_l)}{f_k(\chi_s)} \tag{2.11}$$

when $\chi_s > \chi_l$ and $W_l = 0$ for $\chi_s < \chi_l$.

The CMB lensing weight (or kernel) decreases sharply at low redshift, while the galaxy lensing weights depends on the redshift distribution of source galaxies. In Fig. 2.1 we show the redshift-dependent weights of both CMB lensing and galaxy lensing (calculated using the SDSS source sample redshift distribution as described in Sec. 2.3.3) as a function of redshift, as well as the combined weight given by

$$W_L(z)^{\text{Combined}} = \frac{H(z)}{c}D(z)^2W_L(z)^{\text{CMB}}W_L(z)^{\text{galaxy shear}}. \tag{2.12}$$

$W_L(z)$ for CMB and galaxy shear are defined in Eq. (2.11); $D(z)$, the growth function normalized to 1 at $z = 0$, accounts for the growth of matter perturbations with redshifts. In this work, we will use source galaxies with photometric redshift values satisfying $z_p > 0.1$ (see section 2.3.3). The galaxy lensing weight peaks around $z \sim 0.2$ and the combined weight peaks around $z \sim 0.3$, though it is skewed towards higher redshift with $\langle z\rangle = 0.35$.

### 2.2.3 Lensing signal around galaxies

When studying the lensing signal around galaxies, we are interested in measuring the projected average surface mass density $\Sigma$ around the lens galaxy sample. We start by writing the convergence and shear

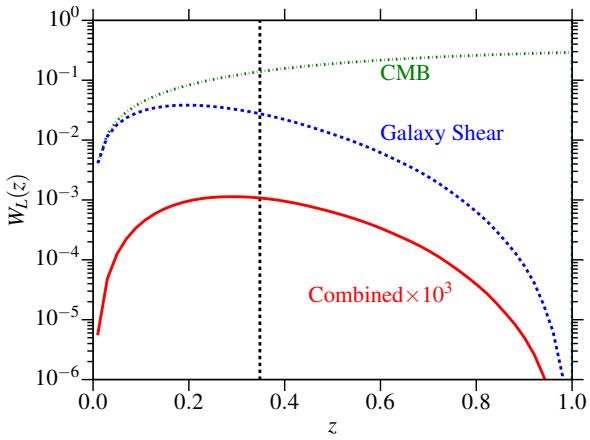



Figure 2.1: We show the weight functions that enter the lensing-lensing correlations as defined in Eqs. (2.11) and (2.12). The vertical black line marks the effective redshift for the lensing-lensing cross-correlation signal, measured using combined weights. For galaxy lensing we used the SDSS source sample redshift distribution as defined in Sec. 2.3.3.

as

$$\kappa(r_p) = \frac{\Sigma(r_p)}{\Sigma_c} \tag{2.13}$$

$$\gamma_t(r_p) = \frac{\bar{\Sigma}(< r_p) - \Sigma(r_p)}{\Sigma_c} \tag{2.14}$$

$\bar{\Sigma}(< r_p)$ is the mean projected surface mass density within the radius $r_p$ (transverse comoving), and the critical surface density is defined in comoving coordinates as

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{f_k(\chi_s)}{(1+z_l) f_k(\chi_l) f_k(\chi_s - \chi_l)}. \tag{2.15}$$

The $(1+z_l)$ factor is required to convert the $c^2/G$ factor to comoving space since it has dimensions of $\left[\frac{\text{Mass}}{\text{length}}\right]$. We can write $\Sigma$ in terms of the projected surface mass density as

$$\Sigma(r_p) = \bar{\rho}_m \int \mathrm{d}\Pi\, \xi_{gm}(r_p, \Pi) = \bar{\rho}_m w_{gm}(r_p), \tag{2.16}$$

where $\Pi$ is the line of sight distance from the galaxy in redshift space, and $\xi_{gm}$ is the three dimensional galaxy-matter cross-correlation function. Here we use $\xi_{gm}$ rather than $1+\xi_{gm}$ since the shear is only sensitive to the matter density contrast, i.e., $\Sigma^{\text{lensing}} = \Sigma(r_p) - \bar{\Sigma}$, where $\bar{\Sigma}$ is the background projected surface mass density. Throughout this paper when using $\Sigma$, we mean $\Sigma^{\text{lensing}}$. The projected matter-galaxy correlation function can be derived from matter power spectrum as

$$w_{gm}(r_p) = b_g r_{\text{cc}} \int \mathrm{d}z W(z) \int \frac{\mathrm{d}^2 k_\perp}{(2\pi)^2} P_{\delta\delta}(\mathbf{k}, z) e^{i(\mathbf{r}_p \cdot \mathbf{k}_\perp)} \tag{2.17}$$

where $b_g$ is the linear galaxy bias and $r_{\text{cc}}$ is the galaxy-matter cross-correlation coefficient, assumed here to be independent of redshift. The linear bias assumption is only valid at linear scales $r_p \gtrsim 10h^{-1}$Mpc, and at smaller scales there are substantial contributions from non-linear bias (Baldauf et al., 2010; Mandelbaum et al., 2013). In this work we primarily use scales where the linear bias model is valid, and analysis with a scale-dependent bias will be presented in future work. To first order, lensing measurements are not affected by redshift space distortions and hence we do not include

any corrections for that. The weight function $W(z)$ depends on the redshift distribution of source galaxies and on any redshift-dependent weights used when estimating the signal (see Sec. 2.2.5). We explicitly calculate these weights and include them in the theory calculations, which integrates over the whole redshift range while using the weights (but does assume redshift-independent $b$ and $r_{\rm cc}$). For a nearly volume-limited sample like LOWZ, this is a quite good assumption, but in general if the bias and $r_{\rm cc}$ do evolve with redshift, we will measure an effective value averaged over redshift.

### 1-halo term: Galaxy-galaxy lensing

To model the small-scale signal ($r_p \lesssim 1h^{-1}$Mpc), we use the NFW profile (Navarro et al., 1996), for which the 3D density is

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2} \tag{2.18}$$

The NFW profile can be integrated over the line-of-sight to get the projected mass density, $\Sigma$ and then $\Delta\Sigma$

$$\Sigma(r_p) = 2\int_0^{r_{\rm vir}} \rho(r = \sqrt{r_p^2+\chi^2})\mathrm{d}\chi. \tag{2.19}$$

$$\Delta\Sigma(r_p) = \bar{\Sigma}(<r_p) - \Sigma(r_p) \tag{2.20}$$

For NFW profiles, we define the concentration, $c_{200b} = r_{200b}/r_s$, and mass, $M_{200b}$, using a spherical overdensity of 200 times the mean density:

$$M_{200b} = \frac{4\pi}{3} r_{200b}^3 (200\bar{\rho}_m) \tag{2.21}$$

We use the COLOSSUS[1] software (Diemer & Kravtsov, 2015) to compute the NFW profiles with the mass-concentration relation from Bhattacharya et al. (2013). To avoid the contamination from the host halo (for those galaxies that are satellites within some larger host halo) and halo-halo terms, we fit the NFW profile only in the range $0.05 < r_p < 0.3h^{-1}$Mpc in case of galaxy-galaxy lensing.

### 1-halo term: Galaxy-CMB lensing

In the case of CMB lensing, as described in Sec. 2.3.4, the Planck CMB lensing maps only go up to $l_{\rm max} = 2048$, which corresponds to a hard edge in $\ell$ that effectively smoothes out the configuration-space maps at $\sim 6'$ scales. To model the signal, we then have to smooth out the NFW profile by convolving it with a sinc kernel ($\propto \frac{J_1(l_{\rm max}\theta)}{l_{\rm max}\theta}$) and with a tophat kernel with size set by the resolution of the healpix map being used. However, $6'$ is $\gtrsim$ the virial radius at all redshifts and expected halo masses of our samples, and thus the measured profiles on all scales will have contributions from satellite and halo-halo terms. Since the simple NFW profile does not contain these contributions, mass estimates using only the NFW profiles will be biased high. To get more correct mass estimates, one should use a halo model. However, since the signal-to-noise ratio of the CMB lensing at these scales is fairly low, and there are additional uncertainties in the measurement from the Planck beam and possible leakage of foregrounds and astrophysical systematics such the thermal SZ effect (van Engelen et al., 2014), we do not attempt more complicated modeling in this work. Instead, for a qualitative comparison, we will present convolved profiles by defining $\Sigma_{gg}$ and $\Sigma_{gm}^{\rm approx}$ as

$$\Sigma_{gg}(r_p) = \frac{\bar{\rho}_m}{b_g} w_{gg}(r_p) \tag{2.22}$$

$$\Sigma_{gm}^{\rm approx}(r_p) = \Sigma_{\rm NFW}(r_p) + e^{-(0.5/r_p)^2}\Sigma_{gg}(r_p) \tag{2.23}$$

where $w_{gg}$ is measured directly from the data as described in Sec. 2.2.4 and 2.2.5. At very small scales, the $w_{gg}$ measurement is likely to be biased because of imperfect corrections for fiber collisions

[1] http://www.benediktdiemer.com/code/

(see Sec. 2.3.2). Hence we down-weight $w_{gg}$ at these scales and use the NFW profile, which should provide an adequate estimates for the small-scale profile. We fit this profile to the data with the NFW halo mass as a free parameter. However, these mass constraints could be biased because of several approximations being made in this model, so they too should be taken as a rough guide.

Finally, we will also present a comparison with $\Sigma_{gm}$ measured from the 'Med-Res' $N$-body simulations that were first presented in Reid et al. (2014), using the $z = 0.25$ and $z = 0.6$ snapshots for LOWZ and CMASS galaxies, respectively. The sample of halos we use is generated using an HOD model from Zheng et al. (2005) fit to the clustering of galaxies assuming fixed abundance of parent halos of the galaxies, with priors on abundance being derived from range of redshift-dependent abundance of galaxies (see Reid et al., 2014, for more details). To get $\Sigma_{gm}$, we cross-correlate the halos with matter particles to obtain $w_{gm}$, which is then multiplied with $\bar{\rho}_m$ to get $\Sigma_{gm}$.

We note that the smoothing kernel is only applied to the theoretical predictions when fitting a model to the small-scale signals below the resolution of the Planck convergence maps. When fitting large scales ($r_p \gtrsim 5h^{-1}$Mpc), we have confirmed that the effects of smoothing are negligible, and hence we do not apply the smoothing kernel to the theoretical predictions in those cases.

### 2.2.4 Galaxy Clustering

We also measure the two-point correlation function of galaxies to constrain their large-scale (linear) bias. In the linear bias regime, the two-point correlation function of galaxies is given by

$$\xi_{gg}(r_p, \Pi) = b_g^2 \int \mathrm{d}z\, W(z) \int \frac{\mathrm{d}^2 k_\perp \mathrm{d}k_z}{(2\pi)^3} \tag{2.24}$$

$$P_{\delta\delta}(\mathbf{k}, z)(1 + \beta\mu_k^2)^2 e^{i(\mathbf{r}_p.\mathbf{k}_\perp + \Pi k_z)}. \tag{2.25}$$

The Kaiser factor, $(1 + \beta\mu_k^2)$, accounts for the redshift space distortions in the linear regime (Kaiser, 1987), where $\beta = f(z)/b_g$, $f(z)$ is the linear growth rate factor at redshift $z$ and $\mu_k = k_z/k$. The weight function $W(z)$ is given by Mandelbaum et al. (2011) as

$$W(z) = \frac{p(z)^2}{\chi^2(z)\mathrm{d}\chi/\mathrm{d}z} \left[ \int \frac{p(z)^2}{\chi^2(z)\mathrm{d}\chi/\mathrm{d}z} \mathrm{d}z \right]^{-1}. \tag{2.26}$$

Here $p(z)$ is the redshift probability distribution for the galaxy sample.

Finally we integrate $\xi_{gg}$ over the line-of-sight separation to get the projected correlation function

$$w_{gg}(r_p) = \int_{-\Pi_{\max}}^{\Pi_{\max}} \mathrm{d}\Pi\, \xi_{gg}(r_p, \Pi) \tag{2.27}$$

We use $\Pi_{\max} = 100h^{-1}$Mpc to reduce the effects of redshift space distortions (van den Bosch et al., 2013).

### 2.2.5 Estimators

In this section we present the estimators used for measuring various signals. For all measurements, we use 100 approximately equal-area ($\sim 9^\circ$ on a side) jackknife regions to obtain the jackknife mean and errors for each bin, as described in Singh et al. (2015) in more detail. When fitting measurements to theoretical predictions, we use a weighted least squares method to fit each jackknife region using just the diagonal elements of jackknife covariance matrix, and then quote the jackknife mean and errors on the best-fitting parameters. When doing MCMC fits (Sec. 2.4.3), we use the jackknife covariance matrix.

### Galaxy-Galaxy Lensing

For galaxy-galaxy lensing we measure $\Delta\Sigma$ as

$$\widehat{\Delta\Sigma}(r_p) = \frac{\sum_{ls} w_{ls}\gamma_t^{(ls)}\Sigma_c^{(ls)}}{\sum_{rs} w_{rs}} - \frac{\sum_{rs} w_{rs}\gamma_t^{(rs)}\Sigma_c^{(rs)}}{\sum_{rs} w_{rs}} \tag{2.28}$$

The summation is over all lens-sources (ls) pairs, where the weight $w_{ls}$ for each lens-source pair is defined as (see e.g. Singh et al., 2015, for more detail)

$$w_{ls} = \frac{\Sigma_c^{-2}}{\sigma_\gamma^2 + \sigma_{SN}^2}. \tag{2.29}$$

$\sigma_{SN}$ is the shape noise and $\sigma_\gamma$ is the measurement noise for the source galaxy. The $\Sigma_c^{-2}$ enters the weight because we defined the $\Delta\Sigma$ in Eq. (2.28) as the maximum-likelihood estimator (Sheldon et al., 2004). Note that the denominator in the first term in Eq. (2.28) has a weight $w_{rs}$, measured by using random lenses rather than real lenses. This accounts for the dilution of the shear by unsheared "source" galaxies that are actually associated with the lens but are put behind the lens due to photometric redshift scatter. The correction factor for this effect, $\sum w_{ls}/\sum w_{rs}$, is usually called the boost factor (Sheldon et al., 2004; Mandelbaum et al., 2005). Finally, the second term in Eq. (2.28) is the subtraction of the $\Delta\Sigma$ measured around the random lenses, to remove the effect of spurious shear at large scales (Mandelbaum et al., 2005).

To measure only the tangential shear around galaxies, we use the estimator

$$\widehat{g\gamma_t}(r_p) = \frac{\sum_{ls} w_{ls}^{\gamma}\gamma_t^{(ls)}}{\sum_{rs} w_{rs}^{\gamma}} - \frac{\sum_{rs} w_{rs}^{\gamma}\gamma_t^{(rs)}}{\sum_{rs} w_{rs}^{\gamma}} \tag{2.30}$$

where the weight $w_{ls}^{\gamma}$ for galaxy-galaxy lensing (but not cosmography; see Sec. 2.2.7) is defined as

$$w_{ls}^{\gamma} = \frac{1}{\sigma_\gamma^2 + \sigma_{SN}^2}. \tag{2.31}$$

### Cross-correlation between galaxy positions and CMB convergence

CMB lensing measurements provide convergence $\kappa$ measurements on the sky, with the sky being divided into equal area pixels using healpix[2] (Górski et al., 2005). Using these measurements, we can obtain the projected surface mass density around lens galaxies as

$$\widehat{\Sigma}(r_p) = \frac{\sum_{lp} w_{lp}\kappa_p\Sigma_{c,*}}{\sum_{lp} w_{lp}} - \frac{\sum_{Rp} w_{Rp}\kappa_p\Sigma_{c,*}}{\sum_{Rp} w_{Rp}} \tag{2.32}$$

where the summation is carried over all lens galaxy-CMB pixel pairs ($lp$) separated by comoving projected distance $r_p \in [r_{p,\min}, r_{p,\max}]$ at the lens redshift, where $r_{p,\min}$ and $r_{p,\max}$ define the bin edges. $\Sigma_{c,*}$ is the geometric factor defined in Eq. (2.15) with CMB as the source. The weight factor is defined as

$$w_{lp} = \Sigma_{c,*}^{-2}. \tag{2.33}$$

We do not include pixel noise in the weights since each pixel has the same statistical noise. Finally we subtract out the signal measured around random galaxies ($Rp$ pairs), to remove the spurious signal from noise (more discussion in section 2.4.1).

To measure the convergence signal around galaxies, we use the estimator

$$\widehat{g\kappa}(r_p) = \frac{\sum_{lp} w_{lp}^{\kappa}\kappa_p}{\sum_{lp} w_{lp}^{\kappa}} - \frac{\sum_{Rp} w_{Rp}^{\kappa}\kappa_p}{\sum_{Rp} w_{Rp}^{\kappa}} \tag{2.34}$$

[2] http://healpix.sf.net/
https://github.com/healpy/healpy

The rationale for and effect of subtracting the mean convergence around random points is discussed in Sec. 2.4.1. Under the assumption that each pixel has the same statistical noise, we adopt uniform weights: $w_{lp}^{\kappa} = 1$.

#### Cross-correlation between galaxy shear and CMB convergence

For lensing-lensing cross-correlations, we measure the tangential shear around the pixels of the CMB map and multiply it with the CMB convergence measured in that pixel.

$$\widehat{w}_{\kappa\gamma_t}(\theta) = \frac{\sum_{sp}^{\theta} w_{sp}\gamma_t^{(sp)}\kappa_p}{\sum_{sp}^{\theta} w_{sp}} \tag{2.35}$$

The summation is over all pixel ($p$) and source galaxy ($s$) pairs with separation $\theta \in [\theta_{\rm min}, \theta_{\rm max}]$, where $\theta_{\rm min}$ and $\theta_{\rm max}$ define the bin edges. The weights are inverse variance weights for source galaxies, accounting for shape noise and measurement noise.

$$w_{sp} = \frac{1}{\sigma_\gamma^2 + \sigma_{SN}^2} \tag{2.36}$$

#### Galaxy Clustering

We use the Landy-Szalay (Landy & Szalay, 1993) estimator to compute the two-point correlation function:

$$\widehat{\xi}_{gg}(r_p, \Pi) = \frac{DD - 2DR + RR}{RR} \tag{2.37}$$

$DD$ is the count of galaxy-galaxy pairs, $DR$ is count of galaxy-random pairs and $RR$ are the random-random pairs. The projected correlation function is obtained by integrating over the line-of-sight separation ($\Pi$) bins

$$\widehat{w}_{gg}(r_p) = \sum_{-\Pi_{\rm max}}^{\Pi_{\rm max}} \Delta\Pi\, \xi_{gg}(r_p, \Pi) \tag{2.38}$$

We use $\Pi_{\rm max} = 100h^{-1}$Mpc, with 20 line-of-sight bins of size $\Delta\Pi = 10h^{-1}$Mpc. The choice of bin size does not significantly impact our results since the redshift extent of our sample is $\gg 10h^{-1}$Mpc. The choice of $\Pi_{\rm max}$ also does not affect our measurements; measurements with $\Pi_{\rm max} = 50h^{-1}$Mpc are not signifcantly different from $\Pi_{\rm max} = 100h^{-1}$Mpc. We do use the correct $\Pi_{\rm max}$ in our theory predictions to account for redshift-space distortion effects.

### 2.2.6 Removing small-scale information

When measuring galaxy-galaxy lensing for which the observable is $\Delta\Sigma$, information from the matter distribution on small scales affects the signal measured at large scales as well. If we have a valid model for correlation functions on all scales, this is not a problem. However, perturbation theory-based methods cannot model signals within the virial radius of virialized halos, and even at somewhat larger scales ($r_p \lesssim 10h^{-1}$Mpc), lowest order perturbation theory is not valid. To some extent this problem can be alleviated by using non-linear prescriptions that describe quasi-linear scales with some success. Baldauf et al. (2010) suggested an approach for removing the small-scale information by defining a new estimator

$$\Upsilon_{gm}(r_p, r_0) = \Delta\Sigma(r_p) - \left(\frac{r_0}{r_p}\right)^2 \Delta\Sigma(r_0) \tag{2.39}$$

$\Sigma(r_p)$ and $w_{gg}$ ($\Sigma_{gg} = w_{gg}$) can also be converted to $\Upsilon$ using the relations between $\Sigma$ and $\Delta\Sigma$ both in the data and theory. Baldauf et al. (2010) showed that using $\Upsilon$ not only removes small-scale

information that is difficult to robustly model, it also reduces the impact of cosmic variance and redshift-space distortions in the projected correlation functions. The trade-off made when using $\Upsilon$ is that we are removing signal when computing $\Upsilon$, so the signal-to-noise ratio ($S/N$) in the measurements decreases, especially at scales near $r_0$. Thus we want to choose lower $r_0$ to use more signal and have higher $S/N$, and higher $r_0$ to be able to remove non-linear galaxy bias more effectively. Baldauf et al. (2010) suggests using $r_0 \gtrsim 2r_{\rm vir}$, where $r_{\rm vir}$ is the virial radius of haloes in the sample. For BOSS LOWZ galaxies, $r_{\rm vir} \lesssim 1h^{-1}$Mpc and hence we will use $r_0 = 2h^{-1}$Mpc or greater in our analysis. As stated, using $\Upsilon$ helps in reducing the impact of cosmic variance, which improves the $S/N$, particularly on large scales and at least partially compensates for the lost $S/N$ at small scales.

We will use $\Upsilon$ to derive most of the constraints on galaxy bias $b_g$, galaxy lensing amplitude, and the galaxy-matter cross-correlation coefficient $r_{\rm cc}$. In principle, for the combination of CMB lensing and galaxy clustering, using $\Upsilon$ is unnecessary since the observables at a given $r_p$ are not contaminated by information from smaller $r_p$. However, use of a consistent estimator for all probes is valuable. Moreover, while use of $\Upsilon$ reduces the $S/N$ to some extent (this is dominated by reduced $S/N$ at small scales, as we will show explicitly in Fig. 2.9), the $S/N$ is moderately improved on the large scales that dominate our constraints.

### 2.2.7 Cosmography

Hu et al. (2007b) proposed the idea of using the ratio of CMB convergence and galaxy lensing convergence as a way to measure the distance ratio (distance to surface of last scattering relative to the distance to the source galaxy sample used to estimate the galaxy lensing) and hence constrain the geometry, $\Omega_k$ and the equation of state of dark energy. The ratio is defined as

$$\mathcal{R}(z_l) = \frac{\kappa(z_l, z_*)}{\kappa(z_l, z_s)} = \frac{\Sigma_c(z_l, z_s)}{\Sigma_c(z_l, z_*)} \tag{2.40}$$

Similar distance ratio tests have also been proposed using galaxy or galaxy cluster lensing alone, in both strong lensing (eg. Link & Pierce, 1998; Golse et al., 2002) and weak lensing regimes (eg. Jain & Taylor, 2003; Bernstein & Jain, 2004). Several studies have already measured the distance ratios (e.g. Taylor et al., 2012; Diego et al., 2015; Kitching et al., 2015; Caminha et al., 2016, and references therein), though they are afflicted by several systematics such uncertainties in modeling cluster profiles and cosmic variance in case of multiple strong lens systems, and photometric redshift uncertainties as well as imaging systematics that cause a redshift-dependent shear calibration in the case of weak lensing. The small redshift baseline also limits the cosmological applications of these measurements using optical weak lensing alone (see discussion in Hu et al., 2007b; Weinberg et al., 2013). Using CMB lensing in cosmographic measurements is advantageous in several ways. First, the source redshift for the CMB (redshift of surface of last scattering) is well known, so one of the two redshift slices being compared has no redshift uncertainty. The long redshift baseline between CMB and galaxy lensing sources also improves the sensitivity of $\mathcal{R}$ to cosmological parameters (Hu et al., 2007b). However, using CMB lensing with galaxy lensing makes $\mathcal{R}$ become more sensitive to some of the systematics in galaxy lensing (for example, multiplicative bias) and $\mathcal{R}$ can also be used as test for presence of such systematics.

To measure $\mathcal{R}$, we work with galaxy lensing shear measured using estimator defined in eq. (2.30), not convergence. Instead we convert the convergence measurement from CMB (measured using estimator defined in eq. (2.34)) to the shear. Motivated by the estimator in Sec. 2.2.6, we define the estimator $\upsilon_t$ as

$$\upsilon_t(r_p, r_0) = \gamma_t(r_p) - \left(\frac{r_0}{r_p}\right)^2 \gamma_t(r_0) \tag{2.41}$$

Just as $\gamma_t = \Delta\Sigma/\Sigma_c$, we can write $\upsilon_t = \Upsilon_{gm}/\Sigma_c$. In the limit that $r_0 = 0$, $\upsilon_t$ is simply $\gamma_t$. The CMB lensing convergence $\kappa$ averaged around lens galaxy positions can be converted to $\upsilon_t$ using $\gamma_t(r_p) = \bar{\kappa}(< r_p) - \kappa(r_p)$ and then converting $\gamma_t$ to $\upsilon_t$.

One of the primary motivations for defining $\Upsilon_{gm}$ was to remove information for small scales which are more difficult to model. When measuring $\mathcal{R}$, we do not need to model those small scales (any nonlinear bias, etc. will cancel in the ratio) and hence using $\upsilon_t$ is not strictly necessary. However, when computing the convergence signal from the CMB, the smallest scales are smoothed (see Sec. 2.3.4) and it is desirable to completely remove information from those scales. Thus we will use $\upsilon_t$ to compute $\mathcal{R}$ and our final definition of the estimator $\widehat{\mathcal{R}}$ is

$$\widehat{\mathcal{R}}(z_l) = \frac{\upsilon_t(z_l, z_*)}{\upsilon_t(z_l, z_s)} = \frac{\Sigma_c(z_l, z_s)}{\Sigma_c(z_l, z_*)} \tag{2.42}$$

Note that our estimator is different from the one used by Miyatake et al. (2016), who use $\gamma_t$ to compute $\mathcal{R}$ and exclude the scales which are affected by smoothing. We will show our measurement using small value of $r_0$ ($\ll$ smoothing scale), in which case our estimator is equivalent to one using $\gamma_t$ and excluding scales smaller than smoothing scale. Also in the estimator of Miyatake et al. (2016), the galaxy-galaxy lensing measurement is in the numerator, so their estimator is effectively $1/\mathcal{R}$. We keep the CMB lensing measurement in the numerator since it is noisier. Later in this section we describe how we account for the bias that comes from taking the expectation value of the ratio of noisy quantities.

To model the measurement, we begin by computing the galaxy position-convergence and galaxy position-shear cross-correlations

$$\langle g\kappa\rangle(r_p) = \int \mathrm{d}z_l\, p(z_l) \frac{\Sigma(z_l)(r_p)}{\Sigma_c(z_l, z_*)} \tag{2.43}$$

$$\langle g\gamma_t\rangle(r_p) = \int \mathrm{d}z_l\, p(z_l) \int_{z_l}^{\infty} \mathrm{d}z_s p(z_s|z_{\rm ph} > z_l) \frac{\Delta\Sigma(z_l)(r_p)}{\Sigma_c(z_l, z_s)} \frac{1}{\sigma_\gamma^2 + \sigma_{SN}^2} \tag{2.44}$$

where source weights $\frac{1}{\sigma_\gamma^2+\sigma_{SN}^2}$ are defined in Sec. 2.2.5 The theory computations can be converted to $\upsilon_t$ using similar method as in the data. Note that due to variations in the number density of source galaxies, galaxy-CMB and galaxy-source cross correlations measurement are at different effective lens redshifts. In principle this can be modeled, but a desirable feature in $\mathcal{R}$ is to define it as a simple ratio without requiring complicated modeling for small scale signals. To get both signals at the same effective redshift, we explicitly compute the weights from galaxy-galaxy lensing as a function of lens redshift, and use them as weights when computing the galaxy-CMB lensing cross-correlations. These weights decrease strongly with redshift due to the decrease in the number of source galaxies behind a lens for increasing lens redshift. Since CMB lensing kernel increases with redshift, these weights are suboptimal for galaxy-CMB lensing cross correlations and increase the noise in $g\kappa$ measurements. Since the $g\kappa$ measurement is noisier and dominates the noise in $\mathcal{R}$, it is desirable to modify the weights to give higher weight to higher redshift galaxies. Thus we add an additional factor of $\Sigma_c(z_l, z_*)^{-2}$ to weights. This increases the noise in the $g\gamma_t$ measurement from galaxy-galaxy lensing, but this increase is more than balanced by the reduced noise in $g\kappa$ from galaxy-CMB lensing cross correlations. The final lens weights are defined as

$$\mathcal{W}_{\mathcal{R}}(z_l) = \int_{z_l}^{\infty} \mathrm{d}z_s p(z_s|z_{\rm ph} > z_l) \frac{1}{\sigma_\gamma^2 + \sigma_{SN}^2} \Sigma_c(z_l, z_*)^{-2} \tag{2.45}$$

Finally $\mathcal{R}$ is given as

$$\mathcal{R} = \frac{\int \mathrm{d}z_l D(z_l)^2 \mathcal{W}_{\mathcal{R}}(z_l) \Sigma_c^{-1}(z_l, z_*)}{\int \mathrm{d}z_l D(z_l)^2 \Sigma_c(z_l, z_*)^{-2} \int_{z_l}^{\infty} \mathrm{d}z_s \frac{p(z_s|z_{\rm ph}>z_l)}{\sigma_\gamma^2+\sigma_{SN}^2} \Sigma_c^{-1}(z_l, z_s)} \tag{2.46}$$

$D(z_l)^2$ is the matter growth function and enters because $\Sigma$ scales as $D(z_l)^2$ in linear theory. In principle, due to non-linear effects in $\Sigma$, the weights in $\mathcal{R}$ can vary with the scale. Since non-linear

effects also evolve with redshift, the measurement of $\mathcal{R}$ at different scales can be at somewhat different effective lens redshifts, $z_l$, and hence $\mathcal{R}$ can in principle be scale-dependent. However, this effect is likely to be small given the narrow redshift range of LOWZ sample, and should be subdominant to the noise in our measurements. When comparing with data, we show this effect by replacing $D(z_l)$ with scale dependent growth function ($D(z_l, r_p)$) estimated from correlation function using linear theory with halofit nonlinear evolution. We do note that halofit does not capture the full non-linear evolution at small scales and thus our estimate will only be approximate.

To estimate the bias that comes from taking the ratio of noisy quantities before taking the expectation value, we make mock realizations of the numerator and denominator in Eq. (2.46), using the signal-to-noise ratio from CMB measurements for numerator and galaxy lensing measurements for the denominator.

$$\delta X(r_p) = \left| X \frac{\delta \upsilon_t(r_p)}{\upsilon_t(r_p)} \right| \tag{2.47}$$

where $X$ is either the numerator or denominator and $\upsilon_t$ is measured from CMB (galaxy) lensing for numerator (denominator). Using $\delta X(r_p)$, we generate random realizations, $\tilde{X}(r_p)$ assuming gaussian distribution with mean $X$ and standard deviation $\delta X$. We then recompute the $\tilde{R}$ taking the ratio of $\tilde{X}$ and then compute the mean, $\langle \tilde{R} \rangle$, using the same scales as $\widehat{R}$. We include the bias estimated from this exercise in the theory prediction before comparing with the data.

## 2.3 Data

### 2.3.1 SDSS

The SDSS (York et al., 2000) imaged roughly $\pi$ steradians of the sky, and the SDSS-I and II surveys followed up approximately one million of the detected objects spectroscopically (Eisenstein et al., 2001; Richards et al., 2002; Strauss et al., 2002). The imaging was carried out by drift-scanning the sky in photometric conditions (Hogg et al., 2001; Ivezić et al., 2004), in five bands ($ugriz$) (Fukugita et al., 1996; Smith et al., 2002) using a specially-designed wide-field camera (Gunn et al., 1998) on the SDSS Telescope (Gunn et al., 2006). These imaging data were used to create the catalogues of shear estimates that we use in this paper. All of the data were processed by completely automated pipelines that detect and measure photometric properties of objects, and astrometrically calibrate the data (Lupton et al., 2001a; Pier et al., 2003; Tucker et al., 2006). The SDSS-I/II imaging surveys were completed with a seventh data release (Abazajian et al., 2009), though this work will rely as well on an improved data reduction pipeline that was part of the eighth data release, from SDSS-III (Aihara et al., 2011); and an improved photometric calibration ('ubercalibration', Padmanabhan et al., 2008).

### 2.3.2 SDSS-III BOSS

Based on the photometric catalog, galaxies were selected for spectroscopic observation (Dawson et al., 2013), and the BOSS spectroscopic survey was performed (Ahn et al., 2012) using the BOSS spectrographs (Smee et al., 2013). Targets are assigned to tiles of diameter 3° using an adaptive tiling algorithm (Blanton et al., 2003), and the data were processed by an automated spectral classification, redshift determination, and parameter measurement pipeline (Bolton et al., 2012). In this paper we use the BOSS data release 12 galaxies (Alam et al., 2015a).

The number densities of the BOSS samples and of various subsamples used in this work are shown in Fig. 2.2.

#### LOWZ

The LOWZ sample consists of Luminous Red Galaxies (LRGs) at $z < 0.4$, selected from the SDSS DR8 imaging data and observed spectroscopically in the BOSS survey (Reid et al., 2016). The sample is approximately volume-limited in the redshift range $0.16 < z < 0.36$, with a number density of

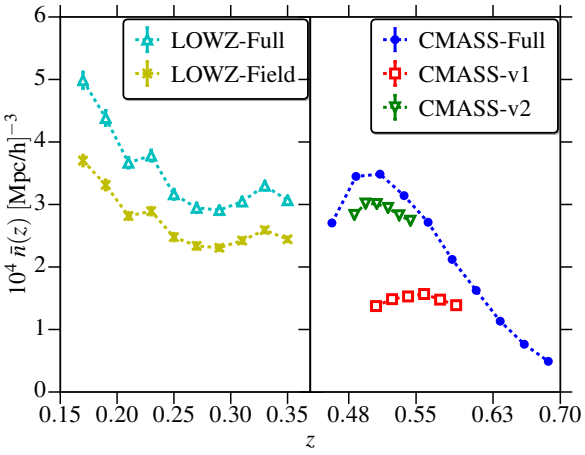



Figure 2.2: Number density as a function of redshift for different samples. LOWZ and field sample are defined in $0.16 < z < 0.36$, while CMASS, CMASS-v1 and CMASS-v2 are defined in $0.45 < z < 0.7$. No weights were used in computing the number densities presented here.

$\bar{n} \sim 3 \times 10^{-4}\ h^3\mathrm{Mpc}^{-3}$ (Manera et al., 2015; Reid et al., 2016). BOSS DR12 has 249 938 LOWZ galaxies within the redshift range used in this work, $0.16 < z < 0.36$. After combining with the Planck lensing map mask and the SDSS shape catalog mask, which masks out certain regions that have higher Galactic extinction or poor imaging quality (Reyes et al., 2012), we are left with 225 181 LOWZ galaxies, with redshift-dependent number density shown in Fig. 2.2.

We also define a sample of field galaxies using the Counts in Cylinders methods (Reid & Spergel, 2009), which was used in Singh et al. (2015) to define a sample of groups from LOWZ galaxies. Here we apply the same technique to DR12, and select field galaxies by requiring that they are in groups of one galaxy and that their fiber collision and redshift failure weights are equal to 1 (see Sec. 2.3.2). These cuts ensure that there is no neighboring target LOWZ galaxy within $r_p < 0.8h^{-1}$Mpc and $|\Pi| < 20h^{-1}$Mpc for the field galaxies. Field galaxies by construction have no redshift failure weights (see Sec. 2.3.2) and thus will provide a test on any possible systematics from these weights. They also tend to reside in lower mass halos, and are less likely to be satellites (which have lensing profiles that are more complex to interpret on small scales due to contamination from the host halo lensing profile).

### CMASS

The BOSS CMASS sample consists of higher redshift galaxies ($0.4 < z < 0.7$) targeted using color and magnitude cuts intended to select a uniform sample of massive galaxies (Reid et al., 2016). The DR12 CMASS sample has 849 637 galaxies, of which we use 682 298 after applying the masks and a redshift cut $z \in [0.45, 0.7]$. The number density of the CMASS sample varies across the redshift range (see Fig. 2.2), which can bias the inferences of its properties from stacked galaxy position-lensing cross correlations. To overcome this problem, we also define two volume-limited samples: CMASS-v1, with $z \in [0.5, 0.6]$ and $M_r \in [-23, -22]$, and CMASS-v2, with $z \in [0.48, 0.55]$ and $M_r \in [-21.5, -22.8]$, where $M_r$ is the absolute magnitude $k + e$ corrected to $z = 0$ using method described in Wake et al. (2006). CMASS-v1 has 188 586 galaxies with $\bar{n} = (1.55 \pm 0.10) \times 10^{-4}\ h^3\mathrm{Mpc}^{-3}$, where $\pm 0.10$ denotes the maximum variation across the redshift range. CMASS-v2 has 236 676 galaxies with $\bar{n} = (3 \pm 0.2) \times 10^{-4} h^3\mathrm{Mpc}^{-3}$. Since both volume-limited samples have a narrow redshift range, we use $\Pi_{\max} = 60h^{-1}$Mpc when calculating the clustering signal for these samples.

**Weights**

In their large scale structure (LSS) samples the BOSS collaboration provides several weights for each galaxy to correct for systematics when estimating the galaxy clustering(Reid et al., 2016). The most important of these are the fiber collision and incompleteness (redshift failure) weights. Due to the finite size of the spectroscopic fibers, it is impossible to simultaneously take spectra of BOSS galaxies that are separated by less than $62''$. Many of these cases are resolved by revisiting the field multiple times. However, some target galaxies lack spectroscopic redshifts either due to fiber collisions and redshift failures. This introduces a bias in the clustering measurements, since the fiber-collided galaxies are preferentially located in overdense regions. An approach that has been shown to work on large scales ($\theta \gtrsim 2'$) (Reid et al., 2014) is to upweight the nearest neighbor of the fiber-collided galaxies, based on the assumption that they are likely to be in the same group due to their proximity on the sky.

In addition, we also use systematics weights, which correct for the effects of varying target density as a function of stellar density for the CMASS sample (Ross et al., 2012) The final weights used for CMASS are

$$w = w_{\rm sys}(w_{\rm no\text{-}z} + w_{\rm cp} - 1) \tag{2.48}$$

where $w_{cp}$ corrects for fiber collisions, and $w_{\rm sys} = 1$ for LOWZ. While these weights have been shown to correct for biases in clustering on large scales, it is not clear whether this approach works well for lensing calculations (see More et al., 2015). $w_{\rm sys}$ is not expected to change the lensing measurements done by stacking procedure, as long as these weights do not alter the overall properties of the sample. These systematic weights do depend on the apparent surface brightness of the galaxies, as those with low surface brightness are more likely to be missed in regions of high stellar density (Ross et al., 2012). However, the dependence on surface brightness is sufficiently mild that their inclusion does not significantly alter the properties of the sample. We checked that using these weights changes the absolute magnitude and redshift distribution of the sample by $\lesssim 0.1\%$ and thus we do not expect any significant changes in the lensing measurement from $w_{\rm sys}$. Still, we do use these weights for all samples and subsamples of CMASS. Redshift failure weights do change both the lensing and clustering measurements by up-weighting the higher density regions. However, they do not mitigate the bias below the fiber collision scale, and even at slightly larger scales, the measurements will be biased since we are stacking on the wrong galaxy. Still, at scales $r_p \gtrsim 2h^{-1}$Mpc, these weights should not lead to any significant change other than changing the effective galaxy bias $b_g$. In Appendix 2.A, we directly show the effect of using these weights using the CMASS sample. For field galaxies, these weights are all equal to one and for volume limited sub-samples of CMASS, it is not guaranteed whether the galaxies missed from redshift failures will pass the magnitude cuts. Hence we omit the results using redshift failure weights for these sub-samples.

### 2.3.3 SDSS shear catalog

For galaxy-galaxy lensing (shear measurements), we use the SDSS re-Gaussianization shape catalog that was introduced in Reyes et al. (2012). Briefly, these shapes are measured using the re-Gaussianization algorithm (Hirata & Seljak, 2003). The algorithm is a modified version of early ones that used "adaptive moments" (equivalent to fitting the light intensity profile to an elliptical Gaussian), determining shapes of the PSF-convolved galaxy image based on adaptive moments and then correcting the resulting shapes based on adaptive moments of the PSF. The re-Gaussianization method involves additional steps to correct for non-Gaussianity of both the PSF and the galaxy surface brightness profiles (Hirata & Seljak, 2003). The components of the PSF-corrected distortion are defined as

$$(e_+, e_\times) = \frac{1-(b/a)^2}{1+(b/a)^2}(\cos 2\phi, \sin 2\phi), \tag{2.49}$$

where $b/a$ is the galaxy minor-to-major axis ratio and $\phi$ is the position angle of the major axis on the sky with respect to the RA-Dec coordinate system. The ensemble average of the distortion is related

to the shear as

$$\widehat{\gamma}_+, \widehat{\gamma}_\times = \frac{\langle e_+, e_\times \rangle}{2R} \tag{2.50}$$

$$\widehat{R} = 1 - \frac{1}{2}\langle e_{+,i}^2 + e_{\times,i}^2 - 2\sigma_i^2 \rangle \tag{2.51}$$

where $\sigma_i$ is the per-component measurement uncertainty of the galaxy distortion, and $\widehat{R} \approx 0.87$ is the shear responsivity representing the response of an ensemble of galaxies with some intrinsic distribution of distortion values to a small shear (Bernstein & Jarvis, 2002).

For this sample, we use photometric redshifts derived from the template fitting code ZEBRA (Feldmann et al., 2006), as described and characterized in Nakajima et al. (2012). Using photometric redshifts can also introduce bias in galaxy-galaxy lensing measurements. Nakajima et al. (2012) showed that this bias can be large, but can be determined to 2 percent accuracy using *representative* spectroscopic calibration samples. Using the calibration method described in Nakajima et al. (2012), we estimate the calibration bias for the galaxy-galaxy lensing by the LOWZ sample to be $\sim -10\%$, and thus multiply our lensing signal by a factor of 1.1 before plotting it or fitting models to it.

When estimating the galaxy lensing-CMB lensing cross-correlations, we do not need redshifts for individual source galaxies. To derive the theoretical predictions for this quantity, we directly use the $\mathrm{d}n/\mathrm{d}z$ obtained from the representative spectroscopic redshift dataset from Nakajima et al. (2012) after applying the same cuts that were applied to data during measurements.

### 2.3.4 Planck Lensing Maps

We use the Planck 2015 lensing map provided by the Planck collaboration (Planck Collaboration et al., 2015b). We convert the provided $\kappa_{l,m}$ values to a convergence map using HEALPY (Górski et al., 2005), with $\mathrm{n_{side}}$= 1024 (pixel size of 3.43′), where $\mathrm{n_{side}}$ determines the resolution of the HEALPY map (higher $\mathrm{n_{side}}$ means smaller pixels; $n_{\mathrm{pix}} = 12 n_{\mathrm{side}}^2$ over the full sky). When constructing the convergence map, we use $\kappa_{l,m}$ in the range $8 < \ell < 2048$, which corresponds to modes from $\sim 25°$ to $\sim 6'$. Planck Collaboration et al. (2015b) found some evidence of systematics in the high $\ell$ range, and used $40 < \ell < 400$ ('conservative') for their main cosmological constraints, though using the 'aggressive' range that is adopted here gives a very similar amplitude of the lensing power spectrum and constraints for cosmological parameters except for $\sigma_8 \Omega_m^{0.25}$, which shifts by $\sim 1\sigma$ between the two $\ell$ ranges.

While our primary results use $\mathrm{n_{side}}$= 1024, we test the effects of changing the pixel size (using $\mathrm{n_{side}}$= 512 and 2048, with pixel sizes of 6.9 and 1.71′ respectively) and applying smoothing on the convergence maps (Gaussian beams with $\sigma$ = 1 and 10′). Since the pixel size with $\mathrm{n_{side}}$= 512 is somewhat greater than the smoothing scale in the lensing map, we do expect to gain some information by using smaller pixels with $\mathrm{n_{side}}$= 1024. Going to even higher resolution with $\mathrm{n_{side}}$= 2048 should not make a very significant difference except at very small scales, in case there is some information left in the lensing maps at those scales. Similarly, smoothing with a Gaussian beam with $\sigma = 1'$ should not significantly affect our results given the resolution of the Planck maps and the scales used for our measurements, though $\sigma = 10'$ should change the signal on scales up to the FWHM ($\approx 25'$) of the smoothing kernel. Hence we will omit the measurements with $\sigma = 1'$ and will show results with $\sigma = 10'$ smoothing for comparison with the main results, which have no additional smoothing applied to maps.

When calculating the cross-correlations, we apply the common galaxy and Planck mask on both the galaxy shear and Planck convergence maps. This reduces the area within the BOSS mask by $\sim 3\%$, primarily driven by the Planck point source mask which selectively masks the very massive clusters. This can change the effective linear bias of the galaxy samples and hence we use the same mask when computing galaxy clustering as well as galaxy-galaxy lensing.

To perform null tests, we generate a map by shuffling the pixels in the convergence map (with the Planck mask applied) and then applying the galaxy mask. We also generate a realization of a noise

map using the noise power spectrum provided with the lensing maps. Throughout this work, $\kappa_{sh}$ and $\kappa_N$ will be used to represent the shuffled map and noise map convergences, respectively.

## 2.4 Results

### 2.4.1 Lensing of the CMB by galaxies

In this section we present the results from cross-correlating the Planck convergence maps with the lens galaxy samples described in Sec. 2.3.2.

Figure 2.3 shows the cross-correlation signal for galaxy position vs. CMB lensing for both the LOWZ and CMASS samples. We show the signal measured around the galaxies and around random points. At small scales, there is significant signal around galaxies, and the the measurements in different $r_p$ bins are uncorrelated. At large scales, the lower signal and the fact that the stack in each bin includes almost all pixels in the map leads to the signal being dominated by the noise in the CMB convergence, and hence the bins are very strongly correlated (we are effectively measuring the mean and standard deviation of the map in every bin). This noise can be removed by subtracting out the signal measured around random points from the signal measured around the galaxies, as shown in Eq. (2.32). Another way to understand the effect of this random subtraction is that similar to the Landy-Szalay estimator in clustering, we want to correlate the CMB map with a mean zero quantity ($\langle D - R \rangle = 0$), so that any additive systematics are removed to first order (Mandelbaum et al., 2006). After subtracting, the different $r_p$ bins exhibit less substantial but still noticeable correlations even at large scales, as shown in Fig. 2.5.

The final measurements after subtracting the signal around random points are consistent with the Planck 2015 ΛCDM model predictions. The solid lines in Fig. 2.3 show the linear theory + halofit correlation functions with the best-fitting galaxy bias, $b_g$; the parameters of these fits are presented in Table 2.1. We only fit this model for scales $r_p > 10h^{-1}$Mpc, as at smaller scales, the effects of non-linear bias and stochasticity in the galaxy-matter cross-correlation are expected to cause deviations between the data and the model. However, as shown in Fig. 2.3, there is no evidence of tension between the theory and the data even down to $r_p \sim 5h^{-1}$Mpc. On the scales used for the fits, $\chi^2_{\rm red} \sim 0.8$ (0.6) for LOWZ (CMASS). Given the large minimum radius for the fits, the constraints on galaxy bias are independent of the pixel size and of the smoothing imposed by the cutoff in the CMB $\kappa$ map at $l_{\rm max} = 2048$. We remind the reader that the minimum radius for the fits is larger than the resolution of the convergence maps, and hence the results from fitting to theoretical models in this section do not have any smoothing applied to the models. We will discuss the small scale signal, where smoothing is necessary, in Sec. 2.4.2.

In Fig. 2.4 we show the comparison of $\Sigma$ and $\Upsilon_{gm}$ with the predictions from Planck theory model using the best fitting bias from $\Sigma$, $\Upsilon_{gm}$, $w_{gg}$and $\Upsilon_{gg}$. As shown in Fig. 2.4 and Table 2.1, there is some discrepancy between the $b_g$ obtained from $\Sigma$ and $\Upsilon_{gm}$. These discrepancies are not very significant ($\lesssim 1.5\sigma$ after accounting for correlations). This is mostly caused by the noise in the measurements, which leads to mild tension between theory and data as they have slightly different scale dependences. Differences in the scale dependence of the theory and data can in general lead to different bias measurements from $\Upsilon$ and $\Sigma$, since $\Upsilon$ at any scale $r_p$ depends on values at scales smaller than $r_p$ (but greater than $r_0$).

The right column of Fig. 2.3 also shows the null tests for these measurements. The signal around random points is consistent with zero at all scales, though the noise at large scales is dominated by the reconstruction noise in CMB lensing. When using the shuffled CMB lensing map, the signal around random points, galaxies, and their difference is consistent with zero. Finally, the signal measured using the noise map is also consistent with zero. The noise map also serves as a diagnostic for our covariance estimates as shown in Fig. 2.5, since it has the same correlated noise properties. The Jackknife covariance and correlation matrices obtained from measurements using the CMB map and noise map are consistent. In Appendix 2.B we show the consistency between the jackknife covariance matrix and the covariance matrix obtained using 100 independent realizations of the noise maps.

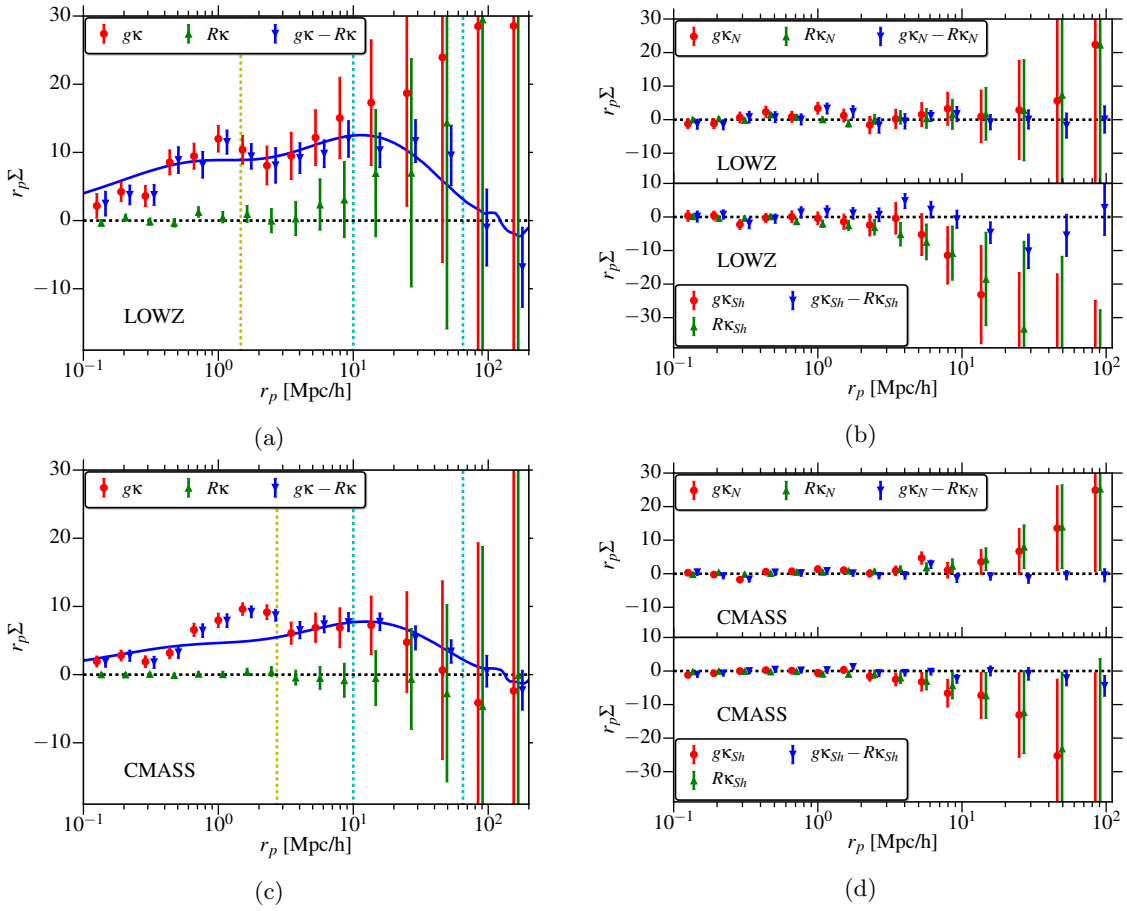



Figure 2.3: $r_p\Sigma$ (in units of $10^6 M_\odot/\mathrm{pc}$) measurements for the LOWZ (top row) and CMASS (bottom row) samples, using maps with $n_{side}$= 1024. The left column shows the measurements around galaxies (red points), random points (green points) and the final signal after taking the difference (blue), along with the Planck ΛCDM model with best-fitting bias. Vertical yellow lines mark the scales corresponding to $6'$ at the maximum redshift of the sample while vertical cyan lines mark the fitting range for the model. At $r_p \gtrsim 30h^{-1}$Mpc ($\gtrsim 2°$), the noise from the convergence map starts dominating the signal around galaxies, but is removed by subtracting out the signal around random points. The right column shows the systematics tests, where the Planck convergence is replaced by the convergence from the noise map $\kappa_N$ and the shuffled map $\kappa_{sh}$. The errors in $\kappa_n$ measurements are very similar to the errors in $\kappa$ measurements, though the errors in $\kappa_{sh}$ show different behavior since the noise correlations are broken by shuffling the measurements. The points at large radius are moderately correlated; see Fig. 2.5.

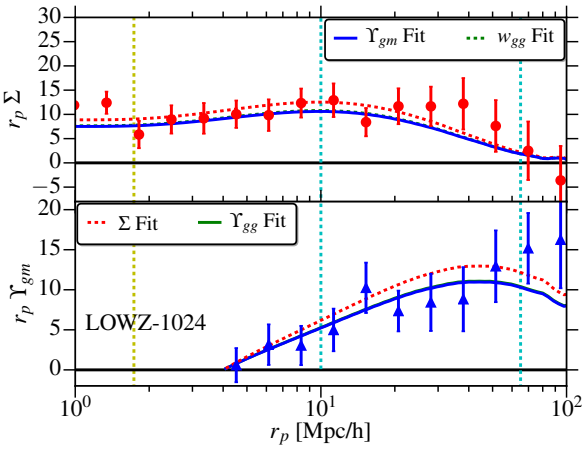



Figure 2.4: Comparison of $\Sigma$ and $\Upsilon_{gm}$ obtained from CMB lensing with Planck theory predictions using best fitting bias from $\Sigma$ and $\Upsilon_{gm}$ as well as from galaxy clustering. Note that the bias from $w_{gg}$ and $\Upsilon_{gg}$ is consistent with the bias from $\Upsilon_{gm}$ and hence the models overlap on the plot. The vertical yellow line marks the smoothing scale $6'$ at $z = 0.36$, and the vertical cyan lines show the range over which models are fitted, $10 < r_p < 70h^{-1}$Mpc.

| Lens sample | $\kappa$ | Pixel size | $b_g(\Upsilon)$ | $b_g(\Sigma)$ | $M_h$ $[10^{12}M_\odot/h]$ $\Sigma_{gm}^{\rm approx}$ |
|---|---|---|---|---|---|
| LOWZ | $\kappa$ | $6.9'$ | 1.80±0.30 | 2.20±0.50 | 13.5±6.5 |
| LOWZ | $\kappa$ | $3.4'$ | 1.80±0.28 | 2.12±0.46 | 13.9±3.9 |
| LOWZ | $\kappa_{\sigma=10'}$ | $3.4'$ | 1.87±0.27 | 2.20±0.40 | 17.1±7.4 |
| LOWZ | $\kappa$ | $1.7'$ | 1.75±0.28 | 2.21±0.46 | 11.8±3.6 |
| LOWZ | $\gamma$ | | | | $10.1 \pm 0.6(\Delta\Sigma)$ |
| Field | $\kappa$ | $3.4'$ | 1.61±0.26 | 1.90±0.40 | 13.5±3.9 |
| Field | $\gamma$ | | | | $8.6 \pm 0.6(\Delta\Sigma)$ |
| CMASS | $\kappa$ | $6.9'$ | 1.50±0.20 | 1.46±0.28 | 24.9±5.7 |
| CMASS | $\kappa_{\sigma=10'}$ | $3.4'$ | 1.75±0.15 | 1.60±0.20 | 34.3±6.4 |
| CMASS | $\kappa$ | $1.7'$ | 1.56±0.19 | 1.51±0.27 | 6.9±2.3 |
| CMASS-v1 | $\kappa$ | $3.4'$ | 1.64±0.29 | 1.80±0.40 | 4±5 |
| CMASS-v2 | $\kappa$ | $3.4'$ | 1.53±0.29 | 1.31±0.39 | 16.3±4.5 |

Table 2.1: Measurement of halo mass (in units of $10^{12}h^{-1}M_\odot$) and linear galaxy bias $b_g$ from lensing alone ($r_{\rm cc} = 1$ fixed), using different estimators. When measuring $b_g$, the signals are fit using $r_p > 10h^{-1}$Mpc, with $r_0 = 4h^{-1}$Mpc for $\Upsilon_{gm}$. We show results for different choice of pixel size and smoothing applied to maps and the rows called "$\gamma$" use the optical galaxy lensing shear instead of CMB lensing convergence maps. Our linear galaxy bias constraints are not significantly impacted by the choice of pixel size or smoothing scale. When measuring halo masses, the NFW profiles are fit between $r_p < 0.3h^{-1}$Mpc (in case of $\Delta\Sigma$) while $\Sigma_{gm}^{\rm approx}$ are fit for $r_p < 5h^{-1}$Mpc. When fitting a smoothed NFW profile, we apply $\sigma = 10'$ smoothing in cases where convergence map has also been smoothed by $10'$ ($\kappa_{\sigma=10'}$).

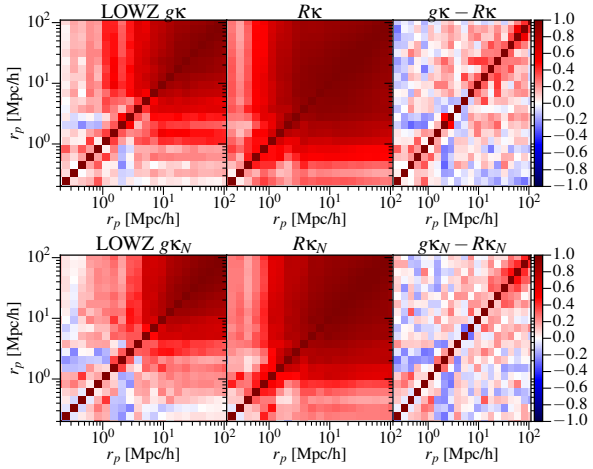



Figure 2.5: Correlation matrix for the galaxy position vs. CMB lensing cross-correlation measurement for the LOWZ sample, with $n_{side}$= 1024. We show correlation matrices for the signal measured around galaxies ($g\kappa$, left column), around random points ($R\kappa$, middle column), and the difference between the two ($g\kappa - R\kappa$, right column) using both CMB ($\kappa$, top row) and noise convergence ($\kappa_N$, bottom row) maps. The correlation and covariance matrices obtained using both maps are consistent. The correlations in $g\kappa$ and $R\kappa$ at large scales are caused by the CMB lensing reconstruction noise, which can be removed by subtracting the signal around random points.

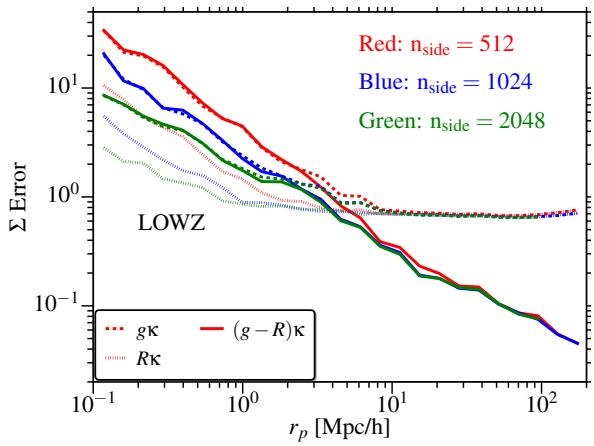



Figure 2.6: Comparison of error bars (square root of the diagonal part of the covariance matrix) in $\Sigma$ measurement for LOWZ. The color indicates the pixel size as indicated on the plot. Solid, dashed, and dotted lines for each $n_{side}$ show errors in $g\kappa - R\kappa$, $g\kappa$, and $R\kappa$ respectively. We use $N_R = 10N_g$, hence errors are smaller for $R\kappa$, and as shown before, taking the difference between $g\kappa$ and $R\kappa$ reduces the errors at large scales. The errors obtained using $\kappa_N$ from noise maps (not shown) are very similar to those using $\kappa$.

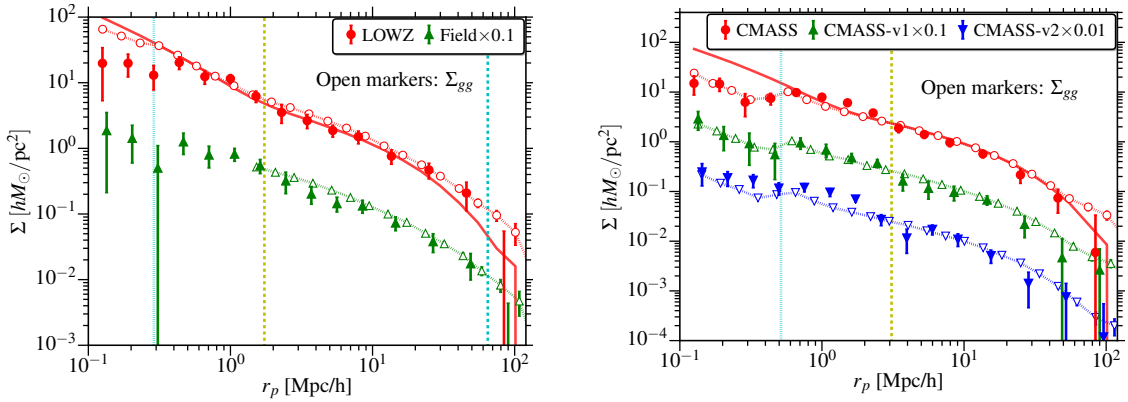



Figure 2.7: Comparison of the surface density $\Sigma$ obtained from CMB lensing (solid markers) with the clustering measurement, $w_{gg}$, converted into $\Sigma_{gg} = \bar{\rho} w_{gg}/b_g$ (open markers). Note that some samples are shifted vertically for easier viewing, with shift factor mentioned in the legend. The solid red lines for LOWZ and CMASS show $\Sigma_{gm}$ measured from simulations (no smoothing applied). The dashed yellow lines mark the $6'$ scale (corresponding to $\ell_{\rm max}$ cutoff) at $z_{\rm max} = 0.36$ (0.7) for LOWZ (CMASS). The dashed cyan lines show the size of the jackknife regions at $z_{\rm min}$ ($r_p \sim 70h^{-1}$Mpc) for LOWZ, off the right side of the plot for CMASS). The dotted cyan lines show the fiber collision scale at $z_{\rm min} = 0.16$ (0.45) for LOWZ (CMASS).

Fig. 2.6 shows the scaling of the errors with $r_p$, and a comparison of errors using different pixel sizes for LOWZ; CMASS results, which are not shown, are similar. Going to smaller pixel size improves the signal-to-noise ratio at small scales, which suggests that there is some information available at these scales In Fig. 2.6, we show results for $\mathrm{n_{side}}$= 512, 1024 and 2048. Using $\mathrm{n_{side}}$= 2048 results in some improvement at small scales, but the results are comparable to those with $\mathrm{n_{side}}$= 1024 at scales $r_p \gtrsim 1h^{-1}$Mpc, motivating our choice of $\mathrm{n_{side}}$= 1024 for our primary results. However, our galaxy bias constraints that are based on large scales are not significantly affected by the choice of $\mathrm{n_{side}}$.

Also in Fig. 2.6, statistical uncertainties on the signals around galaxies and random points saturate for scales $r_p \gtrsim 10h^{-1}$Mpc ($\theta \gtrsim 1^\circ$). This saturation results from the fact that once we have stacked many pixels, we are limited by the reconstruction noise of the CMB lensing maps. This noise is also the reason for the strong bin-to-bin correlations in the first two columns in Fig. 2.5. After subtracting the signal around random points, the final signal is still dominated by noise in the CMB convergence measurements, but there is no evidence of residual systematics from our null tests. Moreover, Fig. 2.5 shows that there are only mild to moderate correlations between the bins.

In Figure 2.7, we show the $\Sigma$ measured using CMB lensing from Fig. 2.3 with the $\Sigma$ measured from simulations and the clustering measurement by converting the clustering into $\Sigma_{gg}$ using Eq. (2.22), where we use the best fit $b_g$ to the clustering signal. The signals are consistent at most scales, though the comparison with clustering is only qualitatively valid. At small scales, the effects of non-linear galaxy bias and the stochasticity in the galaxy-matter cross-correlation can lead to differences between the two. Also, in the BOSS data, the clustering below $\lesssim 1'$ is affected by the incompleteness due to fiber collisions, which biases the signal for $r_p \lesssim 0.3(0.5)h^{-1}$Mpc in the case of LOWZ (CMASS) sample, even when weights are used (even with weights, clustering is not unbiased for $r_p \lesssim 2h^{-1}$Mpc). The lensing signal is also affected since fiber collisions preferentially affect the higher density regions which lead them to be under-weighted in the lensing measurements as well. While fiber collision weights do attempt to correct the bias, the signal is still biased at small scales since we are stacking on the wrong galaxies. The CMB lensing measurement at these scales is also affected by the pixel size and smoothing of the Planck maps. At large scales, the clustering is affected by the residual RSD, while the lensing signal is not to first order. This effect is $\gtrsim 10\%$ above $r_p \gtrsim 70h^{-1}$Mpc. Between

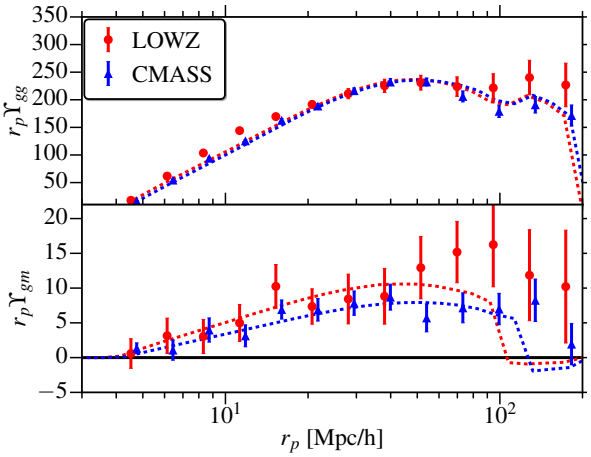



Figure 2.8: $r_p\Upsilon_{gg}$ (top; units of $(\mathrm{Mpc}/h)^2$) and $r_p\Upsilon_{gm}$ (bottom; units of $10^6 M_\odot/\mathrm{pc}$) measurements using CMB lensing, using $r_0 = 4h^{-1}$Mpc. The dashed lines are the Planck ΛCDM model with the best-fitting bias and $r_{\rm cc}$ from our fits.

$10 < r_p < 50h^{-1}$Mpc, we do expect to find a good agreement ($\lesssim 10\%$) between the CMB lensing and clustering measurements, as demonstrated in Fig. 2.7.

In Fig. 2.7, we also show the signals measured for various LOWZ and CMASS subsamples defined in Sec. 2.3.2. The primary motivation for defining these samples was to test for the effects of redshift failure weights (using field galaxies, which should be less affected by fiber collisions) and variations in number densities with redshift (using volume-limited samples CMASS-v1 and CMASS-v2). We do not find any significant tension in results using the subsamples, with the lensing results being largely consistent with the predictions from theory combined with the bias measurements from clustering.

Fig. 2.8 shows the $\Upsilon$ measurements from the BOSS LOWZ and CMASS samples obtained by converting $w_{gg}$ and $\Sigma$ measurements into $\Upsilon$ using methods described in Section 2.2.6 This figure also shows the results from jointly fitting both clustering and lensing measurements with fixed Planck 2015 cosmology to get the galaxy bias $b_g$ and the relative lensing amplitude $r_{\rm cc}$. Due to the use of weighting in lensing measurements, the clustering and lensing measurements are not at the same effective redshift. As stated in Sec. 2.2.3, we integrate the theoretical predictions over redshift using the weights, assuming redshift-independent linear bias and $r_{\rm cc}$. The theory fits the data well within the noise, with $b_g = 1.95 \pm 0.02$ and $r_{\rm cc} = 0.79 \pm 0.13$ for the CMASS sample and $b_g = 1.75 \pm 0.03$ and $r_{\rm cc} = 1.0 \pm 0.2$ for the LOWZ sample (see also Table 2.2). Our $r_{\rm cc}$ measurements are consistent with the theoretical expectation of $r_{\rm cc} = 1$ at linear scales. The galaxy bias measurements for CMASS are consistent with those from Rodríguez-Torres et al. (2016), who measured a scale-dependent bias of $b_g = 1.9$–$2$, using scales 10–$60h^{-1}$Mpc. For LOWZ, our bias is consistent with that measured by Singh et al. (2015) using the DR11 sample ($b_g = 1.77 \pm 0.04$).

In Fig. 2.9, we compare the signal-to-noise ratio ($S/N$) for the clustering and CMB lensing measurements using different estimators. $\Upsilon$ has lower signal-to-noise for $r_p \lesssim 3r_0$ since by definition it has the signal from scales below $r_0$ removed. However, $\Upsilon$ also reduces the impact of cosmic variance and additive systematics in the measurement, and hence improves the $S/N$ at large scales.

### 2.4.2 The small-scale signal

In Figure 2.7, we show the CMB lensing signal at small separations. As discussed in Sec. 2.2.3 the $\ell$ cutoff effectively smoothens the configuration space convergence map with a two-dimensional sinc kernel. In Fig. 2.10 we show the measurements only at $r_p < 5h^{-1}$Mpc along with the smoothed $\Sigma_{gm}$ profile measured from the simulations (blue lines), described in Sec. 2.2.3. We also show the best-fitting smoothed $\Sigma_{gm}^{\rm approx}$ profile, where the NFW halo mass was set as a free parameter in the

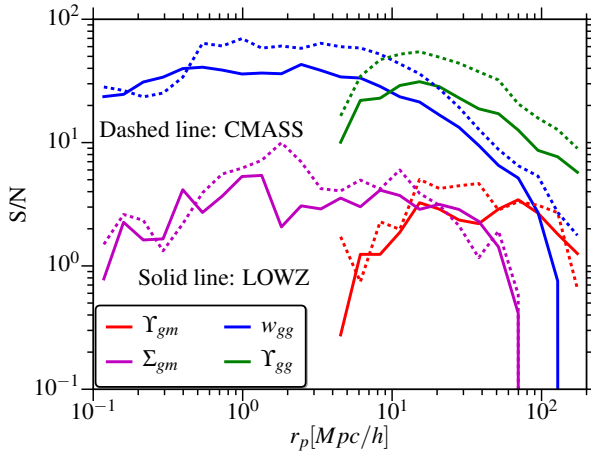



Figure 2.9: Comparison of the signal-to-noise ratio ($S/N$) in the clustering and CMB lensing measurements using different estimators for both LOWZ and CMASS samples. For $\Upsilon$, $r_0 = 4h^{-1}$Mpc. Using $\Upsilon$ decreases the $S/N$ for $r_p \lesssim 3r_0$ since we are removing some of the signal, while at large scales, the $S/N$ improves as $\Upsilon$ reduces the impact of cosmic variance.

| Sample-Planck | $\kappa$ | Pix Area | $b_g$ | $r_{\rm cc}$ |
|---|---|---|---|---|
| LOWZ | $\kappa$ | $3.4'$ | 1.75±0.04 | 1.0±0.2 |
| LOWZ | $\kappa_{\sigma=10'}$ | $3.4'$ | 1.75±0.04 | 1.1±0.2 |
| CMASS | $\kappa$ | $3.4'$ | 1.95±0.02 | 0.78±0.13 |
| CMASS | $\kappa_{\sigma=10'}$ | $3.4'$ | 1.95±0.02 | 0.8±0.1 |
| Field | $\kappa$ | $3.4'$ | 1.47±0.03 | 1.15±0.24 |
| CMASS-v1 | $\kappa$ | $3.4'$ | 2.0±0.03 | 0.9±0.2 |
| CMASS-v2 | $\kappa$ | $3.4'$ | 1.9±0.03 | 0.7±0.2 |

Table 2.2: Results from joint fitting of $\Upsilon_{gg}$ and $\Upsilon_{gm}$, with $r_0 = 10h^{-1}$Mpc and $r_p > 20h^{-1}$Mpc.

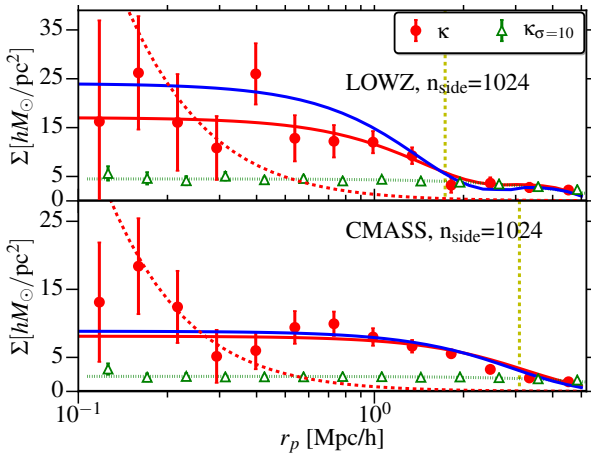



Figure 2.10: $\Sigma$ measurement for the LOWZ (top) and CMASS (bottom) samples with different levels of smoothing applied to the convergence map. The solid red lines show the best-fitting smoothed $\Sigma_{gm}^{\rm approx}$ profile, while dashed lines show the unsmoothed NFW model as a reference (incorrect model given the way the lensing map was produced). The dotted green line shows the $\Sigma_{gm}^{\rm approx}$ profile smoothed with a $\sigma = 10'$ gaussian kernel. The blue lines show the smoothed $\Sigma$ profile measured from $N$-body simulations with an HOD tuned to match the galaxy clustering. The vertical yellow line marks the $6'$ scale (corresponding to $\ell$ cutoff in Planck convergence map) at $z = 0.7$ ($z = 0.36$) for CMASS (LOWZ) sample.

fitting procedure. The halo masses are presented in Table 2.1; these are consistent with the halo mass measured from galaxy-galaxy lensing in the case of the LOWZ and LOWZ-Field subsamples. The halo mass measurement using galaxy-galaxy lensing in this work, $M_h = (1.01 \pm 0.06) \times 10^{13} M_\odot/h$, is different from that of Singh et al. (2015), $M_h = (1.5 \pm 0.2) \times 10^{13} M_\odot/h$, using BOSS DR11 sample. The difference is primarily driven by the different adopted mass-concentration relations, to which the mass estimates are sensitive when fitting scales $r_p < 0.3 h^{-1}$Mpc as described in Sec. 2.2.3. Using the same mass-concentration relation as Singh et al. (2015), we get consistent results ($M_h = (1.68 \pm 0.15) \times 10^{13} M_\odot/h$).

In the case of the LOWZ sample, $\Sigma_{gm}$ from simulations over-predicts the signal. This is because the mean halo mass from the simulations is $M_h \sim 5 \times 10^{13} M_\odot/h$ (median mass $M_h \sim 2.5 \times 10^{13} M_\odot/h$), which is higher than the mass preferred by data, $M_h \sim 10^{13} M_\odot/h$, in the case of both the CMB and galaxy lensing measurements.

The halo mass obtained using the $\Sigma_{gm}^{\rm approx}$ fits for the CMASS sample, $M_h \sim 10^{13} M_\odot/h$, is low compared to the values of $M_h \sim 2 \times 10^{13} M_\odot/h$ measured by Miyatake et al. (2015) and Madhavacheril et al. (2015) using galaxy-galaxy and galaxy-CMB lensing respectively. This discrepancy is in the expected direction since $\Sigma_{gg}$ tends to over-predict the $\Sigma$ profile due to the effects of non-linear galaxy bias and hence the NFW mass will be suppressed. Ultimately, the proper interpretation of the lensing signal at small scales requires proper halo modeling, which we do not do given the noise and resolution of Planck CMB lensing maps. Instead we have presented a simple model to enable easy comparisons, but possible biases in this model should be kept in mind. We note that the profile from simulations (after smoothing) is a reasonable description of the data; in the simulated CMASS sample, the mean $M_h \sim 3.3 \times 10^{13} M_\odot/h$ and median $M_h \sim 1.7 \times 10^{13} M_\odot/h$ (Reid et al., 2014).

### 2.4.3 Lensing calibration: CMB vs. galaxy lensing

Figure 2.11 shows the comparison of $\Upsilon_{gm}$ obtained from galaxy lensing and CMB lensing using LOWZ galaxies as lenses. Note that due to the different weighting used in galaxy lensing and CMB lensing, the two measurements are not at the same effective lens redshift, with $z_{\rm eff} = 0.24$ (0.3) for galaxy

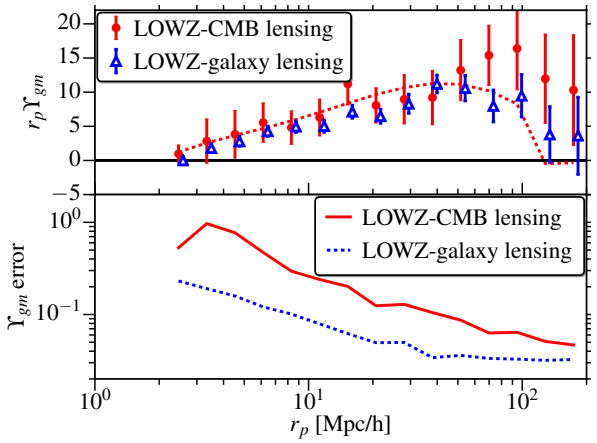



Figure 2.11: Comparison of $r_p\Upsilon_{gm}$ (in units of $10^6 M_\odot/\mathrm{pc}$) and errors in $\Upsilon_{gm}$, obtained from galaxy-CMB lensing and galaxy-galaxy lensing. The dashed red line shows the prediction using Planck 2015 cosmology along with best fitting bias to galaxy-CMB lensing measurement.

(CMB) lensing. In Fig. 2.11, we also compare the measurement uncertainties, with galaxy lensing having a higher signal-to-noise ratio by a factor of 2–5.

In Fig. 2.12, we show the results from jointly fitting the galaxy clustering, CMB lensing and galaxy lensing signals for the LOWZ sample using the MCMC fitting method. We fit for linear galaxy bias $b_g$, $\Omega_m$ and $b_\gamma$, where $b_\gamma$ is the relative calibration bias between CMB lensing and galaxy lensing (CMB lensing amplitude $\propto b_g$, and galaxy lensing amplitude $\propto b_g b_\gamma$). Our result of $b_\gamma = 0.824 \pm 0.15$ is consistent with 1, which would imply no difference in calibration between the two lensing methods, at $\sim 1\sigma$ level. Note that the $b_g$ value from MCMC fits, $b_g = 1.80 \pm 0.06$ differs from the jackknife fits, $b_g = 1.73 \pm 0.04$, shown earlier for two main reasons: the lower value of $\Omega_m$ (fixed to 0.309) in jackknife fits; and we do not use the RSD corrections in the MCMC fits to speed up computation time, moving $b_g$ higher by $\lesssim 1\sigma$ ($b_g = 1.74 \pm 0.04$ for jackknife fitting without RSD correction). The RSD correction is $\lesssim 5\%$ at the scales we use (Baldauf et al., 2010), which is much less than the statistical uncertainties in both $\Omega_m$ and $b_\gamma$. A more detailed cosmological analysis using these measurements with improved modeling on small scales will be presented in a forthcoming paper (Singh et al. *in prep*).

### 2.4.4 Cosmography

In Fig. 2.13, we present the measurement of the distance ratio $\mathcal{R}$ as defined in Eq. (2.42). We present the measurement using two different values of $r_0$: 0.2 and $2h^{-1}$Mpc. As was discussed in Sec. 2.2.7, it is desirable to use $\upsilon_t$ in estimating $\mathcal{R}$, to avoid the information from scales $r_p < 2h^{-1}$Mpc where the smoothing of the CMB lensing map is important. Using $r_0 = 2h^{-1}$Mpc accomplishes this goal, while the $r_0 = 0.2h^{-1}$Mpc case is equivalent to taking ratios using $\gamma_t$ when using scales $r_p > 2 \gg 0.2h^{-1}$Mpc. The sudden drop in measured $\mathcal{R}$ below $r_p < 2h^{-1}$Mpc when using $r_0 = 0.2h^{-1}$Mpc is consistent with expectations from the effects of smoothing. As discussed in Sec. 2.2.7, the effects of non-linear growth also lead to lower $\mathcal{R}$ on small scale by giving higher weights to lower redshifts (where $\mathcal{R}$ is lower), though this is effect is estimated to be much smaller than the statistical uncertainties in our measurements ($\sim 1 - 2\%$ at $r_p < 2h^{-1}$Mpc). Using the "aggressive" range, $2 < r_p < 70h^{-1}$Mpc with $r_0 = 0.2h^{-1}$Mpc, the mean value of $\mathcal{R}$ is $2.68 \pm 0.29$, consistent with the predicted value of $\mathcal{R} = 2.31$ from the Planck 2015 cosmology. Including the effects of noise in our theory prediction as described in Sec. 2.2.7, the prediction becomes $\langle\tilde{R}\rangle = 2.35$ using $> 1000$ realizations with noise and scales from $r_0 = 0.2h^{-1}$Mpc and $2 < r_p < 70h^{-1}$Mpc. Using a more conservative range of scales, $5 < r_p < 70h^{-1}$Mpc with $r_0 = 2h^{-1}$Mpc, we get $\mathcal{R} = 2.74 \pm 0.44$, with the prediction $\langle\tilde{R}\rangle = 2.37$. However, given the strong agreement between our results with conservative and aggressive choices

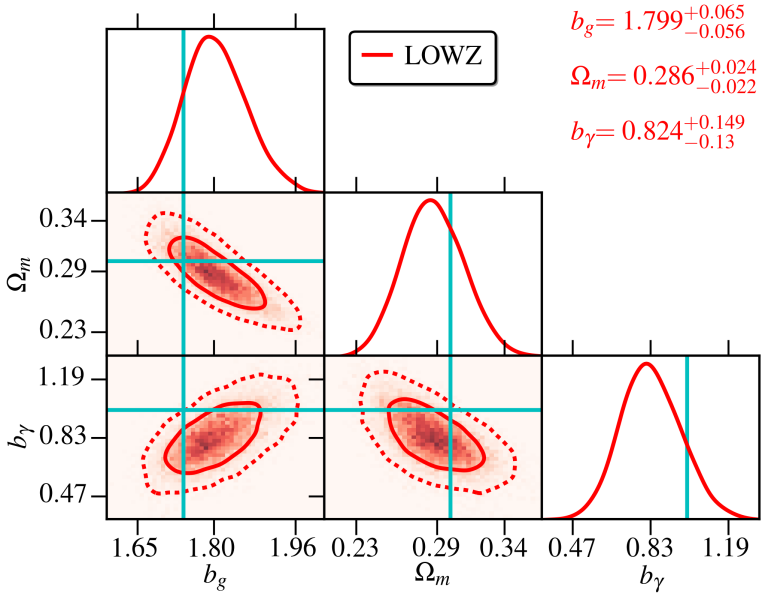



Figure 2.12: Results from MCMC fits to $\Upsilon_{gg}, \Upsilon_{g\gamma}, \Upsilon_{g\kappa}$ for the LOWZ sample, using $r_0 = 10h^{-1}$Mpc and $20 < r_p < 70h^{-1}$Mpc. $b_g$ is the linear galaxy bias and $b_\gamma$ is the relative calibration bias between galaxy lensing and CMB lensing. We use broad uniform priors: $b_g > 0$, $b_\gamma > 0$, and $\Omega_m > 0$. Solid (dashed) contours show $1\sigma$ ($2\sigma$) limits. Cyan lines show the fiducial values: $b_\gamma = 1$, $\Omega_m = 0.309$ and $b_g = 1.74$ ($b_g$ value is obtained with fixed cosmology jackknife best fit and no RSD correction applied).

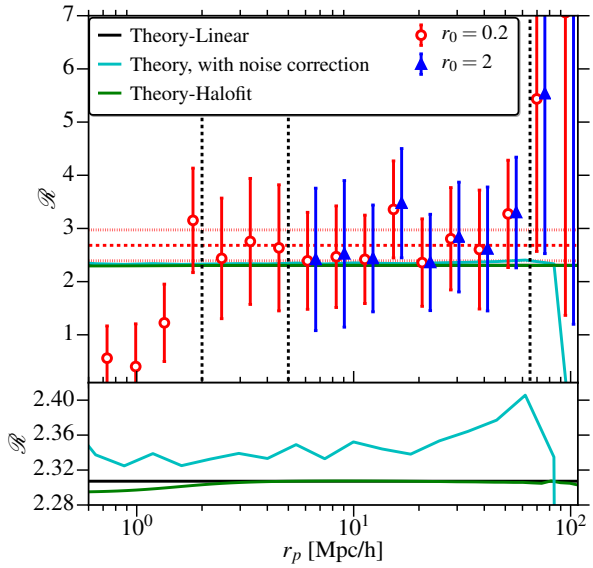



Figure 2.13: Distance ratio $\mathcal{R}$ (see Eq. (2.42)), measured using LOWZ galaxies. Dashed red line marks the mean $\mathcal{R}$ between $2 < r_p < 70h^{-1}$Mpc, while dotted red lines mark the $1\sigma$ limits on the mean. Bottom panel shows the zoomed in comparison of different theory curves. Solid black line is the predicted $\mathcal{R}$ using linear theory with Planck 2015 cosmology while solid cyan line is the prediction using linear linear with the effects of observed measurement noise also included (using $r_0 = 0.2h^{-1}$Mpc measurement). The green line is the predicted $\mathcal{R}$ using linear theory with halofit using Planck 2015 cosmology, accounting for the fact that redshift weighting will emphasize lower redshifts and hence suppress the predicted $\mathcal{R}$ on small scales.

of $r_0$ and scales for the measurement, we quote $2.68 \pm 0.29$ as our primary result. We do not use $r_p > 70h^{-1}$Mpc as these scales are larger than the size of jackknife regions and hence the covariance matrix is not very reliable at these scales.

We do not derive any cosmological constraints using $\mathcal{R}$ since it is not very sensitive to cosmological parameters given the redshift distribution of our lens and source redshifts (Hu et al., 2007b).

### 2.4.5 Lensing-lensing correlations

In this section we present the results of cross-correlating the Planck lensing map with the galaxy shear from the SDSS shape sample (Reyes et al., 2012; Nakajima et al., 2012), using the estimator presented in Sec. 2.2.2.

In Fig. 2.14, we present the lensing-lensing cross-correlations with two different choices of pixel sizes. We fit the signal to a constant $A$ times the predictions for the Planck 2015 cosmology, for $\theta < 2^\circ$ (to avoid scales where noise starts dominating). The best-fitting amplitudes, $A = 0.78 \pm 0.24$ ($n_{side}$=512) and $A = 0.76 \pm 0.23$ ($n_{side}$=1024), which are consistent with the Planck 2015 cosmological parameters ($A = 1$) at $1\sigma$ level, for both pixel sizes. If we relax the fit limits to $\theta < 5^\circ$, the amplitude decreases to $0.63 \pm 0.18$, which is in tension but still consistent with the Planck 2015 cosmological parameters at $2\sigma$ level. The shift between the fits using the more conservative and aggressive ranges of $\theta$ is less than $1\sigma$ after accounting for the correlations between the $A$ values for these two cases, and may be an effect of large-scale systematics at large scales. In the lensing-lensing cross-correlations, $\gamma_t$ is (in principle) also a quantity with zero mean and hence the estimator is less prone to the effects of correlated noise. However, $\gamma_t$ also has some residual additive systematics at large scales (Mandelbaum et al., 2013), making it a quantity with non-zero mean on large scales; it can therefore combine with the correlated noise Two(or systematics) in the CMB convergence maps to give some residual systematics in the cross-correlations. [3] As discussed later in this section, we do not find any evidence of systematics in our null tests, though the uncertainties in our measurements are large. In this work we do not attempt to construct a better estimator given the noise in our measurements. In future works, with better signal-to-noise, it will be worth exploring a better estimator that removes the effects of residual additive systematics.

Fig. 2.15 shows several null tests used to uncover the effects of systematic errors. The first is the 'B-mode' signal $\langle \kappa\gamma_\times \rangle$, which is expected to be zero from parity conservation. As a test, we fit $\langle \kappa\gamma_\times \rangle$ to a model consisting of a constant $A$ times the prediction for $\langle \kappa\gamma_t \rangle$ from the Planck 2015 cosmology; this gives $A = 0.01 \pm 0.21$, consistent with 0 as expected. Similarly, we repeat the measurement by replacing the Planck convergence with the noise map to compute $\langle \kappa\gamma_t \rangle$ and $\langle \kappa\gamma_\times \rangle$, and using the shuffled convergence map (not shown, $A = 0.3 \pm 0.4$). All of these measurements give $A$ consistent with zero.

While the deviations from the ΛCDM predictions using the Planck 2015 cosmology are not statistically significant, it is worth noting that there are several possible systematics that could bias the amplitude of this cross-correlation, for example, intrinsic alignments (IA) and biases in the redshift distributions. Chisari et al. (2015a) estimated the contribution from IA contamination in the CMB vs. galaxy lensing cross-correlations to be around $\sim 10\%$ for the CFHT stripe 82 survey. Since the SDSS source sample is shallower than the stripe 82 sample and hence at lower effective redshift, the fractional IA contamination can be higher. This statement assumes the validity of the linear alignment model (Hirata & Seljak, 2003), which has been shown to describe the IA of red galaxies, and which predicts that the IA signal is approximately constant with redshift while the lensing cross-correlation signal with the CMB decreases at lower redshift. The IA contamination also depends on the galaxy

[3] TwoIt is known that shear has systematics, so that measured shear, $\widehat{\gamma} = \gamma + \gamma_{\rm noise} + \gamma_{\rm sys}$. The measured cross-correlation with galaxies is then $\langle g\widehat{\gamma}_t \rangle = \langle (1+\delta_g)\widehat{\gamma}_t \rangle \approx \langle \delta_g\gamma_t \rangle + \langle \gamma_{\rm sys} \rangle$, where $\langle \gamma_{\rm sys} \rangle$ is the systematics term in galaxy-galaxy lensing and is removed by subtracting measurement around randoms. Now, lets consider simple case in which CMB convergence also has some additive systematic, so that $\widehat{\kappa} = \kappa + \kappa_{noise} + \kappa_{sys}$, where we assume $\kappa_{sys}$ is constant and same in all pixels. In galaxy-CMB lensing, this systematic also gets removed when measurement around randoms is subtracted. The correlation with shear is then, $\langle \widehat{\kappa}\widehat{\gamma}_t \rangle = \langle \kappa\gamma_t \rangle + \kappa_{sys}\langle \gamma_{sys} \rangle$. The $\kappa_{sys}\langle \gamma_{sys} \rangle$ can bias the shear-convergence cross correlation measurements.

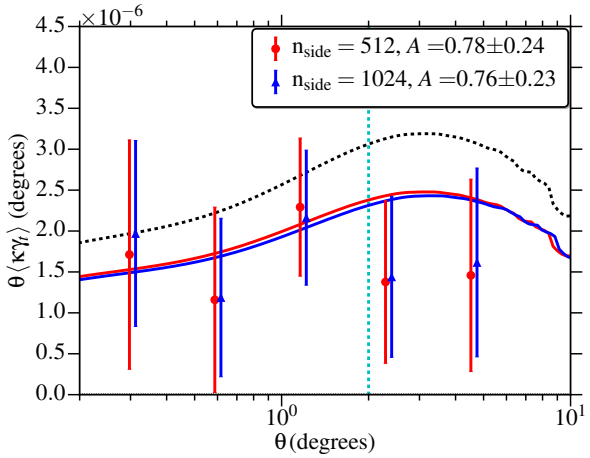



Figure 2.14: The cross-correlation between the Planck CMB lensing convergence ($\kappa$) map and the SDSS shears. The solid lines show the Planck $\Lambda$CDM model obtained using the $\mathrm{d}n/\mathrm{d}z$ from Nakajima et al. (2012), with best fit amplitude and fitting scale $\theta < 2^\circ$ (marked by a vertical line). The dashed black line shows the model with amplitude $A = 1$.

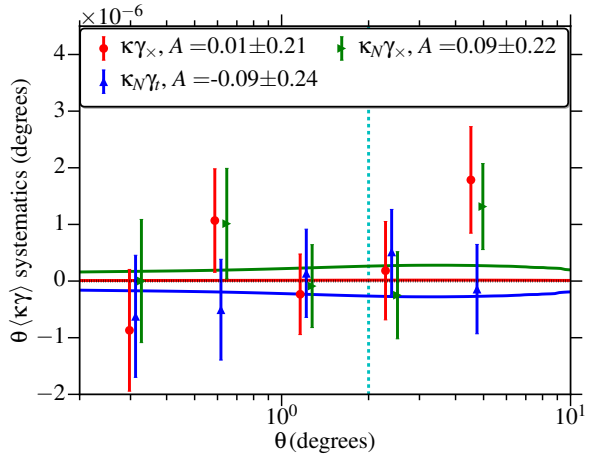



Figure 2.15: Null tests for the lensing-lensing cross-correlations. Points show the measurements (expected to be consistent with zero) while solid lines show the $\langle\kappa\gamma_t\rangle$ model fit to these with a free overall amplitude $A$, for which the best-fitting value is given in the legend. The vertical cyan line marks $\theta = 2^\circ$, the largest scale use in fitting the data. All the null tests we perform are consistent with zero at the $1\sigma$ level.

luminosity (fainter galaxies have lower IA) and on the fraction of red galaxies in the sample (here we ignore alignments for spiral galaxies since all current measurements of IA for spirals are consistent with zero). Assuming a sample with $\sim 20\%$ galaxies to be LOWZ-like LRGs (effective IA amplitude $A_I = 1$), we predict a contamination of $\sim -30\%$. This is a conservative upper limit considering that our source sample is much fainter than the LOWZ sample and is dominated by blue galaxies. See also Blazek et al. (2012) who constrained IA contamination in galaxy-galaxy lensing measurements to be less than 5% using the same source sample as this work.

Regarding photo-$z$ systematics and redshift uncertainties, their impact on galaxy-galaxy lensing for this shear catalogue was quantified using a complete and representative spectroscopic sample in Nakajima et al. (2012), who found $\sim 2$ per cent uncertainties in the mean $\Sigma_c$ and therefore lensing amplitude. To check whether the difference may be more significant here, we considered the difference in the best-fitting amplitude $A$ when we make the theoretical predictions using the redshifts from the calibration sample from Nakajima et al. (2012), vs. when we make them from a smooth parametric fit to the histogram, and find $\sim 1$ per cent uncertainties. We therefore conclude that redshift uncertainty is a subdominant contributor to the error budget for the lensing-lensing correlations. The same may not be true in future datasets for which a representative spectroscopic sample is not available.

## 2.5 Conclusions

In this paper we have presented results from cross-correlating Planck CMB lensing maps with shear from SDSS galaxy lensing and galaxy positions from the SDSS-III BOSS survey using both the LOWZ and higher redshift CMASS sample.

Cross-correlating galaxy positions with the convergence maps, we detect the CMB lensing signal around galaxies out to $\sim 100h^{-1}$Mpc. The measured signals are consistent with $\Lambda$CDM predictions using the Planck 2015 cosmology with bias measurements from clustering. Our null tests do not reveal any significant evidence for systematics in our measurements. The mild tensions between data and theory, e.g., $r_{\rm cc} = 0.78 \pm 0.13$ for CMASS sample, are likely due to noise fluctuations, particularly given that there is no tension observed for the LOWZ sample.

We also detected the CMB lensing signal around galaxies at very small separations, well below the effective smoothing scale of $6'$ in the convergence maps. Combining the clustering measurements with NFW profiles and then applying the smoothing kernel, we are able to constrain halo mass at the 3–4$\sigma$ level for different samples, though the halo masses could be biased given the simple adopted model.

We directly compared the lensing signal around LOWZ galaxies obtained from galaxy-lensing and CMB-lensing. We find that the galaxy lensing has a better signal-to-noise ratio by a factor of 2–5, depending on the scale. Combining these measurements with the galaxy clustering signal, we also performed a basic cosmological analysis jointly fitting for $\Omega_m$, galaxy bias, and the relative calibration bias between galaxy and CMB lensing. We find $\Omega_m = 0.286 \pm 0.024$, consistent with the Planck 2015 cosmology at the $1\sigma$ level. We find the relative calibration bias between galaxy lensing and CMB lensing to be $0.82 \pm 0.15$, consistent with 1 at just over $1\sigma$. In addition, we also measured the distance ratio between the lens and source galaxies and the CMB last scattering surface to within $\sim 10\%$. The ratio is consistent with the $\Lambda$CDM prediction using Planck 2015 cosmology; unfortunately, this ratio is not strongly sensitive to cosmology given the low lens redshift, and hence does not provide competitive cosmological constraints.

For lensing-lensing cross-correlations, we detected the signal at $> 3\sigma$ significance at an effective redshift of 0.35. The amplitude of the signal is consistent with $\Lambda$CDM model predictions using Planck 2015 cosmology. Given the noise in this measurement, we expect systematic errors such as intrinsic alignments and uncertainties in the source galaxy redshift distribution to be subdominant components of the error budget. We also performed null tests, which did not show any evidence for systematics within the errorbars.

To conclude, our results demonstrate how CMB lensing data can be incorporated into and combined with the galaxy lensing analysis using existing lensing surveys. Even though the CMB lensing measurements are noisier than galaxy lensing and will perhaps remain so for the near future, existing

CMB lensing measurements are already good enough to provide strong consistency checks on galaxy lensing measurements. This analysis is an important proof of concept for future surveys that plan to use CMB lensing in conjunction with galaxy lensing, as an additional high-redshift lens plane with completely independent systematics. With better resolution in the upcoming lensing results from current generation and Stage IV CMB surveys, CMB lensing can also develop into a unique tool to study dark matter at higher redshifts, to which it is most sensitive, and which will remain beyond the reach of currently planned galaxy lensing surveys.

## Acknowledgments

We thank François Lanusse, Anthony Pullen, Alex Geringer-Sameth, Sébastien Fromenteau and Shirley Ho for useful discussions related to this work. We also thank Uroš Seljak, Emmanuel Schaan, David Spergel and the anonymous referee for helpful feedback on this work. We thank Martin White and Beth Reid for providing us halo catalog from simulations. We also thank the SDSS-I/II/III and Planck collaboration for their efforts in providing the datasets used in this work.

RM acknowledges the support of the Department of Energy Early Career Award program. SS acknowledges support from John Peoples Jr. Presidential Fellowship from Carnegie Mellon University.

Some of the results in this paper have been derived using the HEALPix package (Górski et al., 2005).

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.SDSS3.org/.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

*Author Contributions*: SS and RM contributed to the analysis and writing the paper. JB is an SDSS-III BOSS architect who contributed to the development of the BOSS survey.

## 2.A Effect of weights in the CMASS sample

In this section we briefly discuss the effect of using weights in measurements involving the CMASS sample. Fig. 2.16 shows the difference in clustering measurements done with and without using weights. The effect of the systematics weights is to shift $w_{gg}$ lower by an additive factor that has only a weak scale dependence; this effect is the dominant impact of the weights on large scales. Since clustering varies very strongly with scale, the fractional change in clustering increases very strongly with scale. The weights for redshift failure up-weight the higher density regions and hence change the clustering strongly at small scales, while the change in the large-scale bias is small ($\lesssim 1\sigma$).

$\Upsilon_{gg}$ by definition will not be very strongly affected by additive changes in $w_{gg}$. Fig. 2.16 demonstrates the fact that the fractional change in $\Upsilon_{gg}$ from systematics weights is $< 5\%$, even on scales where the fractional change in $w_{gg}$ is $\geq 20\%$. At the scales that dominate our constraints in $b_g$ ($r_p \lesssim 50h^{-1}$Mpc), the fractional change in $\Upsilon_{gg}$ without vs. with systematics weights is $< 1\%$, below the statistical errors. As a result, our $b_g$ constraints from $\Upsilon_{gg}$ do not depend on the choice of whether or not to use systematics weights. The redshift failure weights change the clustering by a strongly

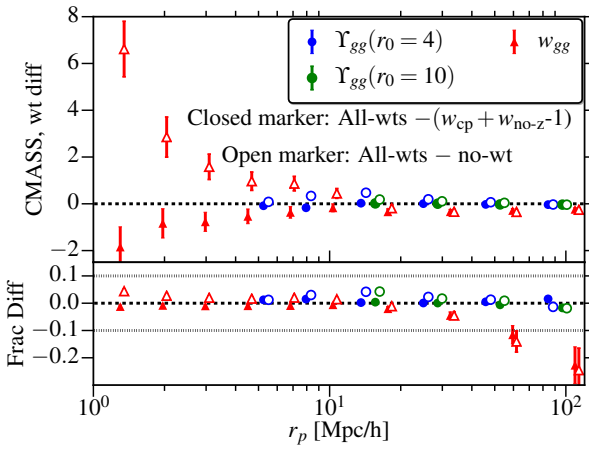



Figure 2.16: Effect of using weights in CMASS clustering measurement. In the top panel we show the difference between using all weights vs. using some or no weights. The bottom panel shows the fractional difference; the y-axis is $\frac{Y_{\rm all\text{-}wt} - Y_{\rm w}}{\langle Y_{\rm all\text{-}wt}\rangle}$, where $Y_w$ refers to calculations with no weights or with only incompleteness weights ($w_{cp}$ and $w_{no-z}$). The systematic weights lower the clustering by an approximately scale independent additive factor, while the redshift incompleteness weights primarily change the clustering at small scales by up-weighting the higher density regions.

scale-dependent factor, so $\Upsilon_{gg}$ is affected in nearly the same way as $w_{gg}$. However, the effect on scales we use for the fits for large-scale bias is small, so the change in linear galaxy bias is again $\lesssim 1\sigma$.

In Fig. 2.17 we show the effect of using weights on the lensing measurements with Planck convergence maps. Given the noise in the lensing measurements, the systematics weights do not affect the measurements very significantly and the fractional change in lensing is $< 10\%$ at all scales, less than the statistical errors in the measurements. The redshift incompleteness weights change the small scale signal in a similar way as in the clustering measurement, by up-weighting the higher density regions, though the shift at large scales is small and negligible given the larger noise in the lensing measurements.

## 2.B Comparison of different error estimates

Here we present a brief comparison of errors in the CMB lensing measurements obtained using the jackknife method with the errors obtained using the scatter between 100 random realizations of the noise map $\kappa_N$.

In Fig. 2.18, we show the comparison of the error bars (square root of diagonal elements of covariance matrix, $\delta\Sigma$) in $\Sigma$ from these two methods. The errors obtained using the two methods are consistent to within 20% ($< 10\%$ on most scales), with the jackknife errors being larger on most scales. The relative uncertainty in the errors, $\delta(\delta\Sigma)/\delta\Sigma \sim \sqrt{2/99} \sim 0.14$ (Taylor et al., 2013), which predicts $\sim 20\%$ scatter (assuming two estimates are independent) when taking the ratio of errors obtained using the two methods. Thus we can conclude that the errors obtained using the two methods are consistent. In Fig. 2.19, we show the correlation matrix obtained from the two estimation methods; they are consistent within the noise.

We also tested the covariance matrices by varying the number of random points. We find correlation matrices that are consistent when using a number of random points $N_R = n \times N_G$, with $n = 5, 10, 20$; the variation in the error on $\Sigma$ (square root of diagonal of covariance) is $\lesssim 5\%$ between these cases. We use $N_R = 10N_g$ for our main results.

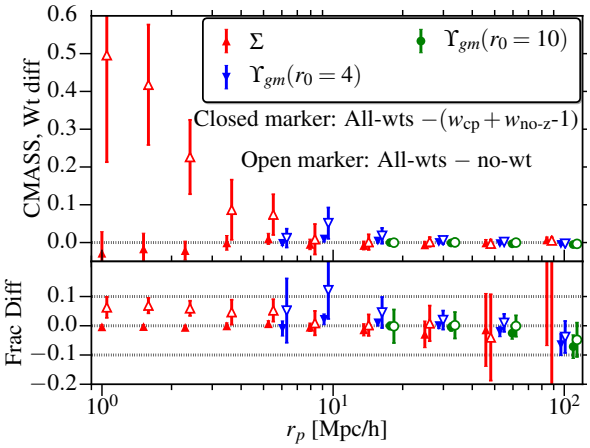



Figure 2.17: Effect of using systematic weights in CMASS lensing measurement using Planck CMB lensing map. y- axis definition is same as in fig. 2.16. The systematics weights do not cause any significant change in lensing measurement, while redshift failure weights change signal primarily at small scales by up-weighting higher density regions.

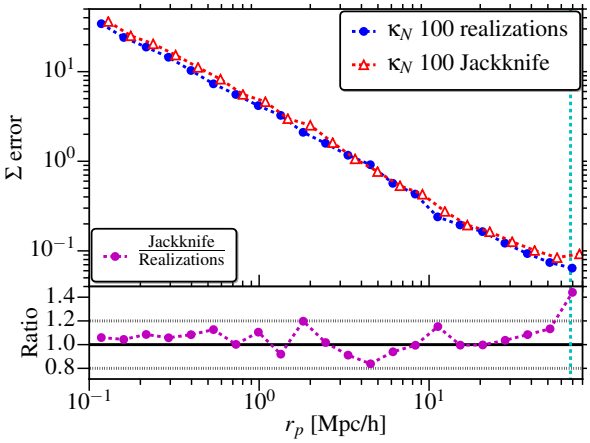



Figure 2.18: Comparison of errors (square root of diagonal of covariance matrix) in Σ measurements using simulated noise maps, estimated using 100 jackknife regions and separately using 100 independent realizations of the map. Up to ∼ 20% scatter in ratio is expected from the noise in estimation of errors.

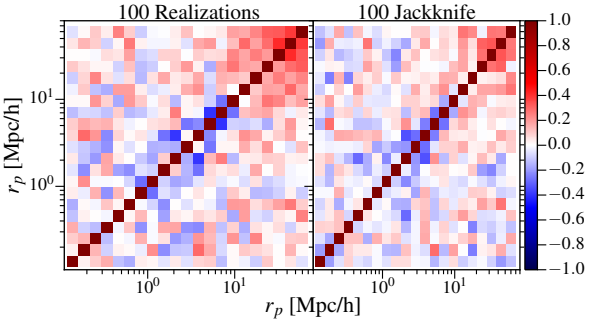



Figure 2.19: Comparison of the correlation matrix estimated using the two methods.

# 3 Galaxy-galaxy lensing estimators and their covariance properties

Sukhdeep Singh[1], Rachel Mandelbaum[1],Uroš Seljak[2,3], Anže Slosar[4], Jose Vazquez Gonzalez[4]

[1]McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[2]Department of Physics, University of California at Berkeley, Berkeley, CA 94705, USA
[3]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4]Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA

**Abstract**

We study the covariance properties of real space correlation function estimators – primarily galaxy-shear correlations, or galaxy-galaxy lensing – using SDSS data for both shear catalogs and lenses (specifically the BOSS LOWZ sample). Using mock catalogs of lenses and sources, we disentangle the various contributions to the covariance matrix and compare them with a simple analytical model. We show that not subtracting the lensing measurement around random points from the measurement around the lens sample is equivalent to performing the measurement using the lens density field instead of the lens over-density field. While the measurement using the lens density field is unbiased (in the absence of systematics), its error is significantly larger due to an additional term in the covariance. Therefore, this subtraction should be performed regardless of its beneficial effects on systematics. Comparing the error estimates from data and mocks for estimators that involve the over-density, we find that the errors are dominated by the shape noise and lens clustering, that empirically estimated covariances (jackknife and standard deviation across mocks) are consistent with theoretical estimates, and that both the connected parts of the 4-point function and the super-sample covariance can be neglected for the current levels of noise. While the trade-off between different terms in the covariance depends on the survey configuration (area, source number density), the diagnostics that we use in this work should be useful for future works to test their empirically-determined covariances.

## 3.1 Introduction

Galaxy-galaxy lensing, the measurement of the tangential shape distortion ("shear") due to gravitational lensing by galaxies, has emerged as an important cosmological probe to study the dark matter distribution around galaxies and the growth of large scale structure (Bartelmann & Schneider, 2001; Weinberg et al., 2013). Since lensing is sensitive to all matter, galaxy-galaxy lensing provides a unique way to map the matter distribution around galaxies (or galaxy clusters) and has been measured to good precision by many surveys (e.g., Fischer et al., 2000; Hoekstra et al., 2004; Sheldon et al., 2004; Mandelbaum et al., 2006; Heymans et al., 2006; van Uitert et al., 2012; Leauthaud et al., 2012; Velander et al., 2014; Viola et al., 2015; Hudson et al., 2015; Clampitt et al., 2016; Buddendiek et al., 2016). Several studies have used galaxy-galaxy lensing to study the halo mass of galaxies and understand the galaxy-halo connection (e.g., Hoekstra et al., 2004; Mandelbaum et al., 2006; Heymans et al., 2006; Tinker et al., 2012; Leauthaud et al., 2012; van Uitert et al., 2012; Gillis et al., 2013; Velander et al., 2014; Sifón et al., 2015a; Hudson et al., 2015; van Uitert et al., 2016). In combination with galaxy clustering, galaxy-galaxy lensing can also be used to recover the underlying matter correlation function, which can then be used to constrain cosmology (Seljak et al., 2005; Baldauf et al., 2010;

Mandelbaum et al., 2013; More et al., 2015; Kwan et al., 2016; Buddendiek et al., 2016) and to test the theory of gravity (Zhang et al., 2007; Reyes et al., 2010; Blake et al., 2016).

With the increasing precision of cosmological surveys, there has been an increasing focus on estimating the covariances of the measurements more accurately as well, since the uncertainties in covariance matrices can lead to incorrect estimation of uncertainties in cosmological parameters (e.g., Hartlap et al., 2007; Dodelson & Schneider, 2013; Taylor et al., 2013; Mohammed et al., 2016). Estimating the non-Gaussian or connected part of the covariance matrix, which has two contributions, is especially challenging. The first contribution is due to mode couplings between small-scale (or in-survey) modes from the non-linear evolution of structure (Scoccimarro et al., 1999; Hu & White, 2001; Cooray & Hu, 2001; Mohammed et al., 2016). The second term is the super-sample variance contribution from the couplings of modes within the survey to the modes corresponding to length scales that are larger than the survey size (eg. Hu & Kravtsov, 2003; Hamilton et al., 2006; Takada & Hu, 2013; Takada & Spergel, 2014; Schaan et al., 2014; Li et al., 2014). Current prescriptions for estimating covariances include using numerical simulations (e.g., Dodelson & Schneider, 2013; Manera et al., 2013; Li et al., 2014), using a physically-motivated halo model (e.g., Takada & Hu, 2013; Schaan et al., 2014) and using analytical estimates based on perturbation theory (e.g., Mohammed et al., 2016).

While many studies have explored this issue of covariance matrices for two-point functions in general, galaxy-galaxy lensing covariances have been relatively less well-studied. When addressing this question, one must also address the question of which estimator is used for the measurement. Several estimators for the galaxy-galaxy lensing signal can be found in the literature. One estimator uses the average tangential shear of background galaxies with respect to the lens galaxies. Another estimator also includes the subtraction of tangential shear around random points, which has an expectation value of zero in the absence of systematics and which can be used to remove the impact of coherent additive shear systematics. Subtraction of the lensing shear around random points is often argued to be beneficial primarily due to the way it removes these additive systematic errors (e.g., Mandelbaum et al., 2005, 2013). However, another motivation for the latter estimator can be found in the work on optimal estimators of galaxy clustering: Landy & Szalay (1993) illustrated that the estimators that are constructed using mean-zero quantities (over-density), while having an expectation value that is the same as the simple estimator, have better covariance properties[1]. For example, in the case of galaxy surveys, random points ($R$) that follow the area coverage of the lenses are typically used to estimate the mean of the galaxy field ($D$) in the presence of complicated survey masks. They are then used to convert the galaxy field into the normalized over-density (mean-zero) field $(D - R)/R$, the auto-correlation of which is the standard Landy-Szalay estimator for galaxy clustering with improved covariance properties. Likewise, for galaxy-galaxy lensing, estimating the mean tangential shear around lens galaxies corresponds to correlating the galaxy density field (nonzero mean) with the shear field, while subtraction of the mean tangential shear around random points results in correlating the mean-zero galaxy over-density with the shear.

In general, most galaxy-galaxy lensing studies either compute the covariance matrices analytically assuming shape noise and measurement noise only (see for example, Viola et al., 2015), or use the jackknife method, which has the advantage that it includes all observational effects, though it is noisier and also limits the scales which can be used in the analysis (e.g., see Hildebrandt et al., 2016; Blake et al., 2016, for comparison of theoretical and jackknife covariance). It is also not clear how well the jackknife method can capture the super-sample covariance, though since galaxy-galaxy lensing is dominated by shape noise in current generation surveys, super-sample covariance is expected to be subdominant. Recently Shirasaki et al. (2016) (see also Blake et al., 2016) did a detailed study of the galaxy-galaxy lensing covariance matrix using realistic $N$-body and ray tracing simulations. In version 1 of their paper, they found that once the scales are of similar order as the jackknife division size, the

[1]Sometimes the reverse claim is made in the literature – i.e., that the estimator with the signal around random points subtracted has increased variance. This claim is typically made in cases where not enough random points are used, in which case there is indeed some added variance. Our argument that the covariance properties of this estimator are superior is true in the limit of infinite random points: they are nothing other than a Monte Carlo method to determine the survey volume and hence the mean density. We explore this issue in more realistic cases in this work.

jackknife method overestimates the errors compared to errors obtained from the standard deviation across different simulation realizations, even in the presence of shape noise. This overestimation in jackknife errors was interpreted as increased contribution from super-sample covariance, since the jackknife method has effectively divided the survey into several small survey realizations, and the super-sample covariance grows with the square of the mass variance within the survey volume (Takada & Hu, 2013), which can scale differently from the usual inverse-volume scaling of the covariance terms. The analysis by Shirasaki et al. (2016) applies to the galaxy-galaxy lensing estimators without subtracting the measurement around the randoms lens sample, as in e.g. Leauthaud et al. (2012); Viola et al. (2015); Hudson et al. (2015); van Uitert et al. (2016); Blake et al. (2016) (Blake et al. (2016) subtracted the measurement around random points from the signal, but this was not done for the covariance estimation). Other galaxy-galaxy lensing studies subtract out the signal around the randoms (e.g., Sheldon et al., 2004; Mandelbaum et al., 2005, 2006; Clampitt et al., 2016; Kwan et al., 2016). In an updated version of their paper, Shirasaki et al. (2016) show that after subtracting out the measurement around randoms, the covariance decreases and the covariance from the jackknife method is consistent with the covariance obtained using different mock realizations (for scales smaller than the size of the jackknife regions).

In this work, we explore the covariance properties of these two galaxy-galaxy lensing estimators both in the presence and the absence of systematic errors. We show that there is a theoretical reason to believe that the estimator with the mean shear around random points subtracted should have more optimal covariance properties, and we explore the impact of this difference in practice for one particular survey. Aside from the issue of removing systematics, we demonstrate the correlated noise term between measurements around galaxies and randoms, which results in more optimal variance properties after subtracting the shear around random points. We also study the differences in the covariance matrices obtained from the jackknife method and standard deviations across several mock realizations, similar to Shirasaki et al. (2016). We demonstrate several methods of empirically estimating specific covariance contributions, and interpret the results of those methods in terms of which galaxy-galaxy lensing covariance terms they include.

This work is organized as follows. In Sec. 3.2 we briefly review the theoretical formalism and estimators, and in Sec. 3.3 we present the data used. Results are presented in Sec. 3.4, and we conclude in 3.5. In appendix 3.A we derive the expressions for covariance when cross-correlating non-zero mean quantities and in appendix 3.B we present comparisons of different estimators in the case of galaxy clustering measurements.

Throughout this work we use the Planck 2015 cosmological parameters (Planck Collaboration et al., 2015a) with $\Omega_m = 0.309$, $n_s = 0.967$, $A_s = 2.142 \times 10^{-9}$, $\sigma_8 = 0.82$. Theory predictions are computed using the linear theory + halofit (Smith et al., 2003; Takahashi et al., 2012) power spectrum generated with the CAMB (Lewis & Bridle, 2002b) software. We use $h = 1$ when computing distances and hence our $\Delta\Sigma$ measurements are in units of $hM_\odot/\mathrm{pc}^2$.

## 3.2 Formalism and Methodology

### 3.2.1 Galaxy lensing

Here we briefly review the formalism of galaxy-galaxy lensing. For a general review of gravitational lensing we refer the reader to Bartelmann & Schneider (2001); Weinberg et al. (2013); Kilbinger (2015).

In galaxy-galaxy lensing, we measure the projected surface mass density $\Sigma$ around the lens galaxies. In the case of a spherically symmetric lens, we can write the convergence and shear as

$$\kappa(r_p) = \frac{\Sigma(r_p)}{\Sigma_c} \tag{3.1}$$

$$\gamma_t(r_p) = \frac{\bar{\Sigma}(< r_p) - \Sigma(r_p)}{\Sigma_c}. \tag{3.2}$$

$\bar{\Sigma}(< r_p)$ is the mean surface mass density within the transverse separation $r_p$, and the critical surface density is defined as

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{f_k(\chi_s)}{(1+z_l) f_k(\chi_l) f_k(\chi_s - \chi_l)}, \tag{3.3}$$

where $f_k(\chi)$ is the transverse comoving distance ($f_k(\chi) = \chi$ in a flat universe). $1 + z_l$ converts the $c^2/G$ factor to comoving space.

We can write $\Sigma$ in terms of the projected galaxy-matter correlation function as

$$\Sigma(r_p) = \bar{\rho}_m \int \mathrm{d}\Pi\, \xi_{gm}(r_p, \Pi) = \bar{\rho}_m w_{gm}(r_p), \tag{3.4}$$

where $\Pi$ denotes the line-of-sight separation from the halo center, and we have ignored the effects of lensing window function, which depends on $\Pi$. Nominally the definition for $\Sigma$ should include a factor of $1 + \xi_{gm}$ within the integral (rather than just $\xi_{gm}$), but the constant term does not contribute to $\gamma_t$ because it gets removed by subtraction of the $\bar{\Sigma}(< r_p)$ term. In the linear bias regime, the galaxy-matter projected correlation function can be derived from the matter power spectrum as

$$w_{gm}(r_p) = b_g r_{cc} \int \mathrm{d}z W(z) \int \frac{\mathrm{d}^2\mathbf{k}}{(2\pi)^2} P_{\delta\delta}(\mathbf{k}, z) e^{i(\mathbf{r}_p \cdot \mathbf{k})}, \tag{3.5}$$

where $b_g$ is the galaxy bias and $r_{cc}$ is the galaxy-matter cross correlation coefficient, both of which are assumed to be independent of redshift in this equation. $P_{\delta\delta}(k, z)$ is the matter power spectrum (linear+halofit) at redshift $z$. To lowest order, lensing measurements are not affected by redshift space distortions, and hence we do not include any corrections for them. The weight function $W(z)$ depends on the redshift distribution of the source galaxies and on the weights used in the estimators when measuring the signal (see Sec. 3.2.2). We explicitly include these weights when computing the effective redshift $z_{\rm eff}$ for the theory calculations.

### 3.2.2 Estimator

Our observable quantity for the galaxy-galaxy lensing measurement is $\Delta\Sigma$, which is estimated in bins of $r_p$ as

$$\widehat{\Delta\Sigma}_{gR}(r_p) = \frac{\sum_{ls} w_{ls} \gamma_t^{(ls)} \Sigma_c^{(ls)}}{\sum_{Rs} w_{Rs}} - \frac{\sum_{Rs} w_{Rs} \gamma_t^{(Rs)} \Sigma_c^{(Rs)}}{\sum_{Rs} w_{Rs}}. \tag{3.6}$$

The summation is over all lens-sources ("ls") pairs. $\gamma_t$ is the tangential shear measured in the lens-source frame. $\Sigma_c$ is the geometric factor defined in Eq. (3.3), and the optimal weight $w_{ls}$ for each lens-source pair ($w_{Rs}$ is defined analogously for random-source pairs) is defined as (see Mandelbaum et al., 2005)

$$w_{ls} = \frac{\Sigma_c^{-2}}{\sigma_\gamma^2 + \sigma_{SN}^2}. \tag{3.7}$$

The $\Sigma_c^{-2}$ enters the inverse variance weight because we defined the $\Delta\Sigma$ in Eq. (3.6) as the maximum likelihood estimator (Sheldon et al., 2004). Note that the denominator in Eq. (3.6) has a sum over weights $w_{Rs}$, measured by using random lenses rather than lens galaxies. Division by $\sum_{Rs} w_{Rs}$ rather than $\sum_{ls} w_{ls}$ corrects for the dilution of the shear signal by source galaxies that are physically associated with the lens but appear to be behind the lens due to photo-$z$ scatter. These galaxies do not contribute any shear but are counted in the total weights (sum over $w_{ls}$). The correction factor for this effect $\sum_{ls} w_{ls} / \sum_{Rs} w_{Rs}$ (properly normalized to account for different number of random and real lenses) is usually called the boost factor (Sheldon et al., 2004; Mandelbaum et al., 2005) and is $\sim 1$ for the scales we use in this work $r_p \gtrsim 1h^{-1}$Mpc. Finally, we subtract the shear signal measured around the random points to remove any systematics that may contribute a spurious shear signal at large scales, and to construct a more optimal estimator. Throughout this paper, the subscript 'gR'

is used to indicate that the measurement around random points is subtracted from the measurement around the lenses:

$$\widehat{\Delta\Sigma}_{gR}(r_p) = \widehat{\Delta\Sigma}_{g}(r_p) - \widehat{\Delta\Sigma}_{R}(r_p) \tag{3.8}$$

One of the main goals of this paper is to test how the subtraction of the signal measured around random points impacts the covariance matrix of the final measurement. Hence, we will study the signals measured around galaxies and randoms separately as well. We will refer to the signal measured around galaxies by $\widehat{\Delta\Sigma}_g$ and around random points by $\widehat{\Delta\Sigma}_R$. The ratio of the number of random points used to the number of lens galaxies is $N_R$:

$$N_R = \frac{\text{Number of random lenses}}{\text{Number of lens galaxies}} \tag{3.9}$$

In case of $N_R = 0$, $\widehat{\Delta\Sigma}_{gR} \equiv \widehat{\Delta\Sigma}_{g}$.

To estimate jackknife errors, we use 100 approximately-equal area (∼10 degrees on a side) jackknife regions to obtain the jackknife mean and errors for each $r_p$ bin.

### 3.2.3 Covariance: theoretical expectations

As is derived in appendix 3.A, the covariance for $\Delta\Sigma_g$ is given by

$$\begin{aligned}\mathrm{Cov}(\Delta\Sigma_g)(|\mathbf{r}_{p,i}|,|\mathbf{r}_{p,j}|) = &\left[\frac{\mathcal{A}_W(\mathbf{r}_{p,i}-\mathbf{r}_{p,j})}{\mathcal{A}_W(\mathbf{r}_{p,i})\mathcal{A}_W(\mathbf{r}_{p,j})}\frac{1}{L_W}\int\frac{\mathrm{d}k\,k}{2\pi}J_2(kr_{p,i})J_2(kr_{p,j})\right.\\ &\left.\left(\Sigma_c^2\left(P_{gg}(k)+N_g\right)\left(P_{\kappa\kappa}(k)+N_\gamma\right)+\Delta\Pi_2\overline{\rho}P_{g\delta}^2+T_{g\gamma g\gamma}\right)\right]\\ +&\left\{\frac{1}{\mathcal{A}_W(\mathbf{r}_{p,i})\mathcal{A}_W(\mathbf{r}_{p,j})L_W}\int\frac{\mathrm{d}k\,k}{2\pi}J_2(kr_{p,i})J_2(kr_{p,j})\tilde{W}(k)^2\Sigma_c^2\left(P_{\kappa\kappa}(k)+N_\gamma\right)\right\}.\end{aligned} \tag{3.10}$$

Here the lens galaxy power spectrum can be written as $P_{gg} = b_g^2 P_{\delta\delta}(k)$ in the linear bias regime, the lens galaxy shot noise power spectrum is $N_g = \frac{1}{\overline{n}_g}$, the shape noise term is $N_\gamma = \frac{\sigma_\gamma^2}{\overline{n}_s}$, the galaxy-shear cross-power spectrum is $P_{g\gamma} = \overline{\rho} b_g r_{cc} P_{\delta\delta}(k)$, and the convergence power spectrum $P_{\kappa\kappa}$ is given in Eq. (3.51). We compute the $\langle\Sigma_c\rangle$ when performing the measurements, and use $\langle\Sigma_c\rangle \sim 4.7\times10^3 \frac{h M_\odot}{\mathrm{pc}^2}$, $n_s \sim 8h^2\mathrm{Mpc}^{-2}$ (after accounting for weights) in theoretical covariance calculations. $J_2$ is the second order spherical Bessel function, $\sigma_\gamma \sim 0.36/2\mathcal{R}$ is the shape noise, $\Delta\Pi_2 \approx 700h^{-1}\mathrm{Mpc}$ is the line-of-sight integration length using the lensing window function, $W(k)$ is the projected lens window function in Fourier space (see appendix 3.A for the expressions for the window function) and $L_W$ is the line-of-sight length of the lens window function. $T_{g\gamma g\gamma}$ is the connected part of the covariance, which we will ignore in numerical calculations. $\mathcal{A}_W$ (defined in Eq. 3.53) is the window function-dependent effective area covered by each bin, and accounts for the edge effects due to the survey window. For scales much smaller than the survey window, $\mathcal{A}_W \approx A_W$, where $A_W$ is the survey area. In the $\Delta\Sigma$ measurements in this work, we only divide the lens sample into jackknife regions, but the source sample stays the same and hence the edge effects are small. When calculating numerical predictions for the jackknife errors, we set the window function $\mathcal{A}_W \approx A_W$ (ideally we should set $\mathcal{A}_W$ for jackknife to be same as that of full sample, but in the case of an idealized LOWZ-sized window, $\mathcal{A}_W \approx A_W$ for the scales of interest). In appendix 3.B, we show the effects of $\mathcal{A}_W$ on the jackknife covariance in the case of clustering measurements.

The covariance for $\Delta\Sigma_{gR}$ is similar to what is shown in Eq. (3.10), except that it does not contain the last term in curly brackets, {}. This term arises because of the non-zero mean value of the lens density (here the lens sample is assumed to be normalized and hence its mean is 1). This term is independent of the lens over-density and only depends on the window function of the lens sample. Hence it get removed when the measurement around random points is subtracted from the measurement around galaxies.

### 3.2.4 Covariance matrix estimation methods

To estimate the covariance matrix, we use two different methods. The first is the jackknife method, in which we divide the whole survey into $N_{\rm Jk} = 100$ approximately equal-area regions ($\sim 90$ degrees$^2$ $\equiv 76^2, 125^2, 163^2$ [Mpc/h]$^2$ at $z = 0.16$, 0.27, and 0.36 respectively). We then make $N_{\rm Jk}$ measurements by dropping one region at a time, so that each measurement contains $N_{\rm Jk} - 1$ regions. The jackknife variance estimate (diagonals of the covariance matrix) is then

$$\mathrm{Var}_{\rm Jk}(\widehat{\Delta\Sigma}) = \frac{N_{\rm Jk}-1}{N_{\rm Jk}} \sum_{i=1}^{N_{\rm Jk}} (\Delta\Sigma_i - \overline{\Delta\Sigma})^2 \tag{3.11}$$

Our second method is to measure $\Delta\Sigma$ using $N_M$ mock realizations of the lens sample and then compute the standard deviation ("Std") of the measurement across all realizations:

$$\mathrm{Var}_{\rm Std}(\widehat{\Delta\Sigma}) = \frac{1}{N_M-1} \sum_{i=1}^{N_M} (\Delta\Sigma_i - \overline{\Delta\Sigma})^2 \tag{3.12}$$

Finally, for comparison, we also show error estimates using subsamples of the survey. We use the same subsampling as in the jackknife method, but in this case we perform the measurements using one subsample at a time. The variance in this case (error on the mean) is

$$\mathrm{Var}_{\rm subsample}(\widehat{\Delta\Sigma}) = \frac{1}{N_{\rm Jk}(N_{\rm Jk}-1)} \sum_{i=1}^{N_{\rm Jk}} (\Delta\Sigma_i - \overline{\Delta\Sigma})^2 \tag{3.13}$$

## 3.3 Data

### 3.3.1 SDSS

The SDSS (York et al., 2000) imaged roughly $\pi$ steradians of the sky, and the SDSS-I and II surveys followed up approximately one million of the detected objects spectroscopically (Eisenstein et al., 2001; Richards et al., 2002; Strauss et al., 2002). The imaging was carried out by drift-scanning the sky in photometric conditions (Hogg et al., 2001; Ivezić et al., 2004), in five bands (*ugriz*) (Fukugita et al., 1996; Smith et al., 2002) using a specially-designed wide-field camera (Gunn et al., 1998) on the SDSS Telescope (Gunn et al., 2006). These imaging data were used to create the catalogues of shear estimates that we use in this paper. All of the data were processed by completely automated pipelines that detect and measure photometric properties of objects, and astrometrically calibrate the data (Lupton et al., 2001a; Pier et al., 2003; Tucker et al., 2006). The SDSS-I/II imaging surveys were completed with a seventh data release (Abazajian et al., 2009), though this work will rely as well on an improved data reduction pipeline that was part of the eighth data release, from SDSS-III (Aihara et al., 2011); and an improved photometric calibration ('ubercalibration', Padmanabhan et al., 2008).

### 3.3.2 SDSS-III BOSS

Based on the SDSS photometric catalog, galaxies were selected for spectroscopic observation (Dawson et al., 2013), and the BOSS spectroscopic survey was performed (Ahn et al., 2012) using the BOSS spectrographs (Smee et al., 2013). Targets were assigned to tiles of diameter 3° using an adaptive tiling algorithm (Blanton et al., 2003), and the data were processed by an automated spectral classification, redshift determination, and parameter measurement pipeline (Bolton et al., 2012).

We use SDSS-III BOSS data release 12 (DR12; Alam et al., 2015a; Reid et al., 2016) LOWZ galaxies in the redshift range $0.16 < z < 0.36$. The LOWZ sample consists of Luminous Red Galaxies (LRGs) at $z < 0.4$, selected from the SDSS DR8 imaging data and observed spectroscopically in the BOSS survey. The sample is approximately volume-limited in the redshift range $0.16 < z < 0.36$,

with a number density of $\bar{n} \sim 3 \times 10^{-4}\ h^3\mathrm{Mpc}^{-3}$ (Manera et al., 2015; Reid et al., 2016). We use the same sample as used by Singh et al. (2016a), who mask out certain regions on the sky which have higher galactic extinction or poor imaging quality (Reyes et al., 2012), which leaves 225,181 galaxies in the sample.

### 3.3.3 Re-Gaussianization Shapes and Photometric redshifts

The shape measurements for the source sample used in this work are described in more detail in Reyes et al. (2012). Briefly, these shapes are measured using the re-Gaussianization technique developed by Hirata & Seljak (2003). The algorithm is a modified version of ones that use "adaptive moments" (equivalent to fitting the light intensity profile to an elliptical Gaussian), determining shapes of the PSF-convolved galaxy image based on adaptive moments and then correcting the resulting shapes based on adaptive moments of the PSF. The re-Gaussianization method involves additional steps to correct for non-Gaussianity of both the PSF and the galaxy surface brightness profiles (Hirata & Seljak, 2003). The components of the distortion are defined as

$$(e_+, e_\times) = \frac{1-(b/a)^2}{1+(b/a)^2}(\cos 2\phi, \sin 2\phi), \tag{3.14}$$

where $b/a$ is the minor-to-major axis ratio and $\phi$ is the position angle of the major axis on the sky with respect to the RA-Dec coordinate system. The ensemble average of the distortion is related to the shear as

$$\gamma_+, \gamma_\times = \frac{\langle e_+, e_\times \rangle}{2\mathcal{R}} \tag{3.15}$$

$$\mathcal{R} = 1 - \frac{1}{2}\langle e_{+,i}^2 + e_{\times,i}^2 - 2\sigma_i^2 \rangle \tag{3.16}$$

where $\sigma_i$ is the per-component measurement uncertainty of the galaxy distortion, and $\mathcal{R} \approx 0.87$ is the shear responsivity representing the response of an ensemble of galaxies with some intrinsic distribution of distortion values to a small shear (Kaiser et al., 1995; Bernstein & Jarvis, 2002). A discussion of corrections for shear-related systematic biases and the residual systematic uncertainties can be found in Mandelbaum et al. (2013). These estimates are based on a combination of null tests using the shear catalog and external image simulations.

The photometric redshifts for the catalog were estimated using the template-fitting code ZEBRA (Feldmann et al., 2006). Using photometric redshifts for the source sample introduces a bias in galaxy-galaxy lensing through misestimation of the $\Sigma_c$ factor (with the most severe misestimation arising due to the inclusion of some lens-foreground "source" pairs due to scatter in photometric redshifts). Nakajima et al. (2012) showed that this bias can be large, but can be estimated to within 2% using a representative calibration sample with spectroscopic redshifts. We compute these correction factors using the method of Nakajima et al. (2012) with the LOWZ lens redshift distributions to be $\sim 10\%$ and then multiply our measurements with a calibration factor of 1.1.

### 3.3.4 Mock source catalog

We generate 100 mock catalogs of the shape sample by randomly rotating the shapes of galaxies in the real source sample, while keeping their positions (RA, Dec, $z$) fixed. Random rotations remove any coherent shear (cosmological or due to systematics) in the sample while maintaining the shape noise and measurement noise in each realization. As a result, $\Delta\Sigma$ measurements using rotated (mock) sources will not have any coherent signal and their covariance matrix will only have contributions from shape noise and measurement noise. The comparison of the covariance matrix of mocks with the covariance from real sources will allow us to study the contribution of shape noise and measurement noise to the covariance in the real data.

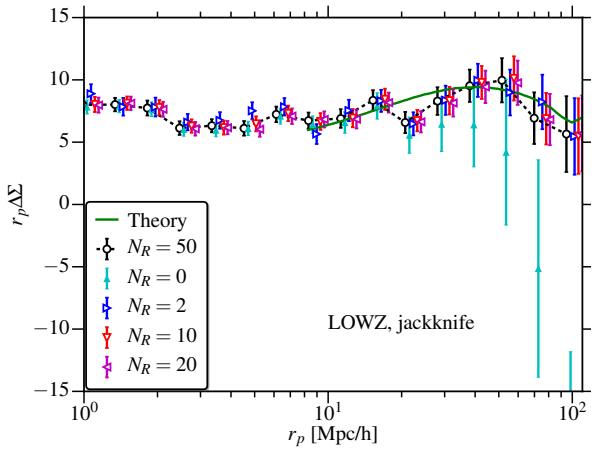



Figure 3.1: $\Delta\Sigma$ measured for the LOWZ sample, with different numbers of random catalogs used. The errors shown are from the jackknife method. The signal without subtraction of the lensing signal around random points ($N_R = 0$, cyan points) shows the presence of additive systematics in the SDSS source sample. These systematics are removed with the subtraction of the signal measured around random points. Theory predictions use the linear theory+halofit power spectrum with fixed cosmology along with the best-fitting linear bias and $r_{cc} = 1$, and fitting was done for $10h^{-1}\text{Mpc} < r_p < 65h^{-1}\text{Mpc}$.

### 3.3.5 QPM mocks

To estimate the galaxy-galaxy lensing covariance matrix using a mock lens sample, we use the QPM mocks (White, 2014) which have been used in several BOSS analyses (e.g., Cuesta et al., 2016; Grieb et al., 2016; Gil-Marín et al., 2016a). QPM mocks are constructed using the quick particle mesh method (White, 2014) to mimic the large-scale clustering properties of BOSS galaxies. In this work, we use 100 QPM mocks with the same sky coverage, mask and jackknife splitting as in the LOWZ sample.

## 3.4 Results

In this section we present our results from measuring galaxy-galaxy lensing using different estimators and different combinations of lens and source galaxies. We perform several tests to study the effects on the estimated covariance by using different covariance estimation methods, varying $N_R$, varying the clustering properties of lens sample, and varying the source sample (without and with systematics). A summary of the various terms in the covariance that contribute for different combinations of lens and source samples is presented in Table 3.1, and a summary of the results is in Figure 3.7.

### 3.4.1 LOWZ lensing results

We begin by showing the galaxy-galaxy lensing measurements using the LOWZ lens sample. Fig. 3.1 shows $\Delta\Sigma$ measured using LOWZ lens galaxies and different numbers of random points. When using no randoms (the $N_R = 0$ case), there is evidence for a spurious systematic signal at large scales. This spurious signal arises because the PSF correction method used to measure the galaxy shapes is unable to fully remove all of the PSF anisotropy. The SDSS survey strategy results in large-scale coherent PSF anisotropy which, when improperly removed, causes a large-scale coherent galaxy shape alignment (see Mandelbaum et al., 2005, 2013, for a detailed discussion). The fact that this spurious signal gets removed when the measurement around randoms is subtracted has been the primary motivation for the subtraction of the signal around random points in SDSS galaxy-galaxy lensing measurements.

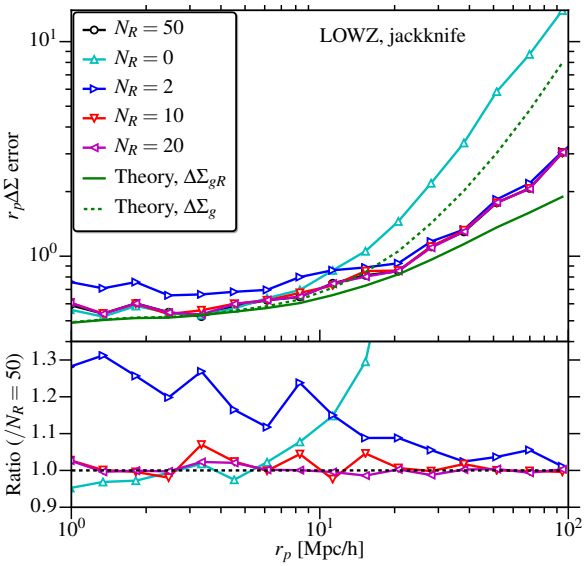



Figure 3.2: Jackknife errors in the LOWZ $\Delta\Sigma$ measurement (square root of diagonal elements of the covariance matrix), for different numbers of random catalogs. The errors generally follow the $\propto 1/r_p$ scaling expected from shape noise and logarithmic binning in $r_p$, though there is some saturation at large scales due to the correlated shape noise and systematics. Also shown are the theory predictions, which are consistent with data at small scales, though there are differences at large scales due to systematics that are not included in the theory.

However, the subtraction of the signal around random points also reduces the errors in the measurements, especially at large scales (the noise in the $\Delta\Sigma_g$ term has contributions from systematics as well as shape noise, as we will show in later sections.). In Fig. 3.2, we show the variations in the error estimates (square root of the diagonal covariance matrix elements) with different numbers of random catalogs $N_R$. At small scales, where the errors follow the expected scaling for shape noise ($\propto 1/r_p$ in logarithmic $r_p$ bins), subtracting the signal around random points increases the error estimates, though with $N_R \gtrsim 10$, the errors converge to $N_R = 0$ case. The errors in this regime should scale with $N_R$ as

$$\left(\frac{\delta\Delta\Sigma(N_R)}{\delta\Delta\Sigma(N_R=0)}\right)^2 = 1 + \frac{1}{N_R}. \tag{3.17}$$

Given that the jackknife error estimates using 100 regions have uncertainty of order $\sim 15\%$ ($\sqrt{2/99}$; Taylor et al., 2013), using $N_R = 10$ is sufficient and henceforth our results will use $N_R = 10$ unless a different value is explicitly given. However, note that when using large numbers of mocks for error estimates, more randoms might be required.

At large scales, contributions to the noise from systematics and the correlated shape noise ($P_{gg}(P_{\gamma\gamma}+N_{\gamma\gamma})$ term) start dominating and hence the error estimates diverge from the $1/r_p$ scaling. The errors are mostly consistent with the theoretical predictions calculated using Eq. (3.10). At large scales there is a contribution from the systematics that is not included in the theory predictions, hence the errors diverge from those predictions especially for the $N_R = 0$ case where systematics are most important. We distinguish between the different terms in the variance in the following sections.

Fig. 3.3 shows the correlation and cross-correlation matrices for $\Delta\Sigma_g$, $\Delta\Sigma_R$ and $\Delta\Sigma_{gR}$, both from theory and data and their difference. The measurements of $\Delta\Sigma$ around galaxies ($\Delta\Sigma_g$) and randoms ($\Delta\Sigma_R$) are highly correlated for $r_p \gtrsim 10h^{-1}$Mpc. When we subtract the measurement around randoms, this correlated noise gets removed and hence the noise in $\Delta\Sigma_{gR}$ decreases compared to that in $\Delta\Sigma_g$ at large scales. The bin-to-bin correlations also decrease, though there are still some residual correlations due to the clustering of the lens sample and the effects of systematics. Since the theory prediction

does not include systematics, the residuals after subtracting the theory correlation matrix from the jackknife are not consistent with zero. A cleaner test of the theoretical expressions will use randomly rotated sources, which do not have any systematic shear correlations, in the next subsection.

### 3.4.2 Mock Sources

In this section we quantify the effects of additive shear systematics on the covariance estimation, especially on the differences in the errors with and without $\Delta\Sigma_R$ subtracted. We create 100 mock realizations of the source sample by randomly rotating the source galaxies. The resulting source catalogs should exhibit no coherent signals of cosmological origin or due to systematics. When measuring $\Delta\Sigma$ around the LOWZ galaxies with these randomly rotated source catalogs, we should observe realistic levels of correlated shape noise, but no systematics, cosmic variance, or super-sample covariance. All terms involving shear correlations – $P_{\gamma\gamma}$, $P_{g\gamma}$, and $T_{g\gamma g\gamma}$ – are zero and hence do not contribute to the covariance.

Fig. 3.4 shows the jackknife and standard deviation errors obtained with and without subtracting $\Delta\Sigma_R$. Subtracting $\Delta\Sigma_R$ reduces the errors, and the results are consistent with the theory predictions. However, the magnitude of the difference in errorbars for $\Delta\Sigma_{gR}$ vs. $\Delta\Sigma_g$ shown here (factor of $\sim 2$ at the largest scale) is lower compared to what was seen with real SDSS sources (factor of $\sim 5$), which suggests that a bit more than half the contribution to the errorbars for $\Delta\Sigma_g$ with real sources came from shear correlations, mostly caused by the systematics rather than cosmic shear given the low redshift of this sample. In the case of $\Delta\Sigma_{gR}$, the errors computed using the standard deviation (Std) across the realizations are consistent with the jackknife errors with $\Delta\Sigma_R$ subtraction. In the case of $\Delta\Sigma_g$, the Std errors are lower than the jackknife errors because of the much larger effective window for Std (full survey window) compared to the jackknife ($1/N_{jk}$ of the survey window).

As shown in Figs. 3.3 and 3.4, in the case of the jackknife window, the predictions from theory are consistent with the data at the $\sim 10\%$ level for both $\Delta\Sigma_g$ and $\Delta\Sigma_{gR}$, which is within the noise in the jackknife errors. In the case of $\Delta\Sigma_g$ with the full survey window, the theory predictions for the errors are lower than the jackknife errors. This is likely because when computing the theory predictions, we assume an idealized geometry (see Appendix 3.A), which underestimates the window function effects from the realistically complicated window in the data.

Fig. 3.4 and the theoretical predictions demonstrate that even in a survey with no known additive systematic errors, measurements of $\Delta\Sigma$ in the $r_p$ range where correlated shape noise is important will have substantially better $S/N$ when using the more optimal $\Delta\Sigma_{gR}$ estimator.

### 3.4.3 Lens Mocks

In this section, we vary the lens properties to examine how the covariance depends on the lens sample properties.

#### QPM mocks

In this section we measure $\Delta\Sigma$ around the galaxies in the QPM mocks using the real and mock source sample. In both cases shear-galaxy correlations will be absent, i.e., $P_{g\gamma} = T_{g\gamma g\gamma} = 0$, while the former will include $P_{\gamma\gamma}$ terms and the latter will not. Even though the QPM mocks have somewhat different clustering at small scales than the real LOWZ sample (see Fig. 3.10), the typical separation between galaxies ($\sim$ 1–2 $h^{-1}$Mpc) is very similar between the mocks and the LOWZ sample and hence the QPM mocks are adequate to test the effects of lens clustering, $P_{gg}$, on the galaxy-galaxy lensing covariance. The signal around the QPM mocks should include large-scale systematics and realistic levels of shape noise.

The left panel of Fig. 3.5 shows the error estimates in the $\Delta\Sigma$ measurements with the $\Delta\Sigma_{gR}$ estimator using $N_R = 10$, and with the $\Delta\Sigma_g$ estimator (without the measurement around random points subtracted). The jackknife errors with the $\Delta\Sigma_R$ subtraction are consistent with the error estimates using the standard deviation across 200 QPM mocks. Also the jackknife errors for the

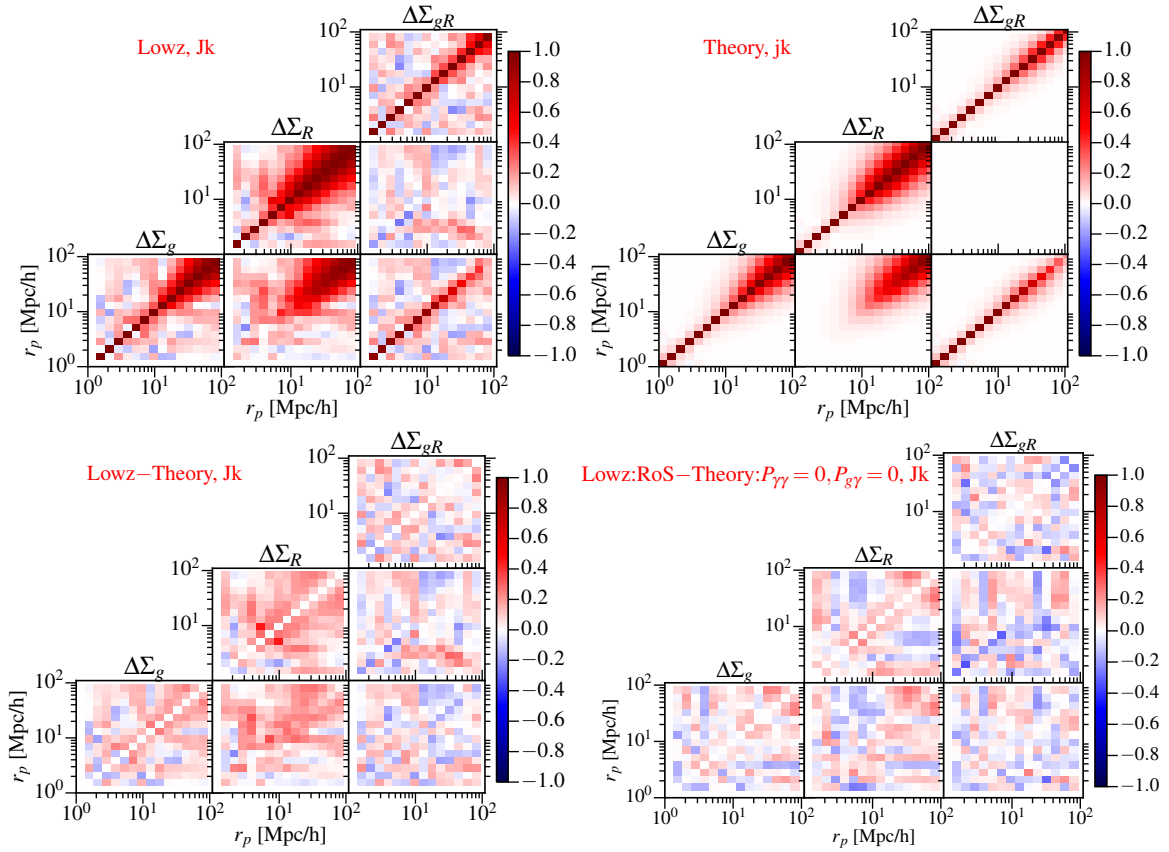



Figure 3.3: *Upper row*: Correlation and cross-correlation matrices for $\Delta\Sigma$ measured around galaxies ($\Delta\Sigma_g$), randoms ($\Delta\Sigma_R$) and the difference of the two ($\Delta\Sigma_{gR}$), both from theory and data (jackknife). Due to the shape noise, there are strong auto- and cross-correlations at large scales in both $\Delta\Sigma_g$ and $\Delta\Sigma_R$. Subtracting the measurement around the random points removes most of the correlated noise (from systematics and correlated shape noise), though there are still some residual bin-to-bin correlations in $\Delta\Sigma_{gR}$, primarily due to the clustering of lens galaxies. *Bottom row*: The difference between the correlation matrices (note we do not take the difference of the covariances here) from the jackknife and the theoretical predictions (*left*) and from the mocks and the theoretical predictions (*right*). In the case of the jackknife matrices, there are systematics that are not included in the theory predictions, so the latter are under-predicated. In the case of the mock sources (or rotated sources, RoS), shear systematics are removed and the theory predictions are consistent with the data within the noise in the jackknife covariances.

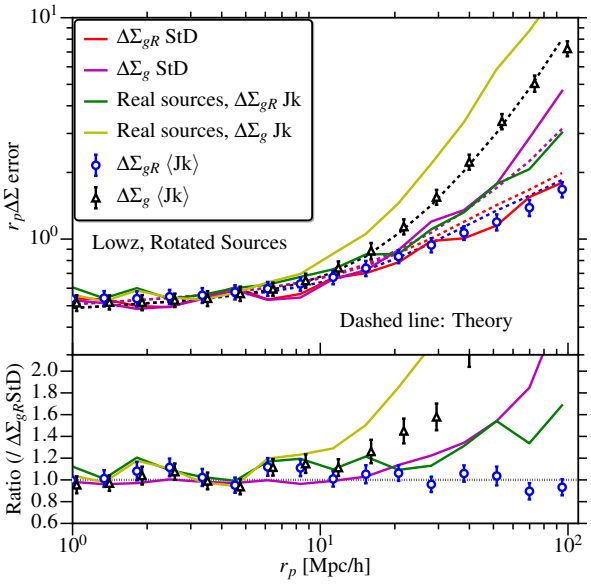



Figure 3.4: $\Delta\Sigma$ errors using different estimators with LOWZ galaxies as lenses and 100 mock realizations of the source sample obtained by randomly rotating the SDSS source galaxies. Open black and blue markers show the mean and standard deviation of the jackknife errors. Solid red and magenta lines show the errors from the standard deviation ("StD") across different realizations of $\Delta\Sigma_g$ and $\Delta\Sigma_{gR}$. Dashed lines are the theory predictions for the curves with corresponding colors. In the bottom panel we plot the ratio of the different errors with respect to the StD errors of $\Delta\Sigma_{gR}$. Jackknife errors for the real LOWZ sample are also plotted for comparison ($N_R = 10$ for LOWZ). Note that in the $\Delta\Sigma_g$ case, the correlated noise in these mocks is lower than when using real sources due to the removal of the contribution from systematics.

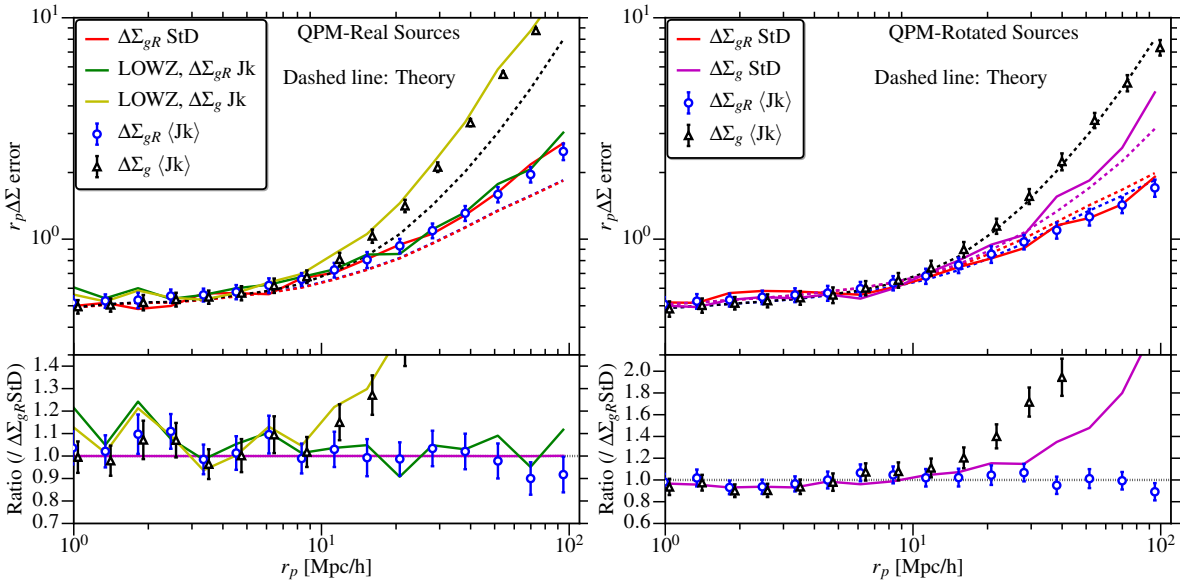



Figure 3.5: Same as Fig. 3.4, using the QPM mocks as the lens sample with real sources (left panel) and rotated (mock) sources (right panel). As in Fig. 3.4, the errors with real sources are higher due to contributions from systematics. The theoretical predictions are consistent except at large scales due to systematics and the effects of the LOWZ window function, which are not well captured by the idealized window function assumed in the theoretical calculations. Also in the case of the real sources, the LOWZ jackknife errors are consistent with those for the QPM mocks, suggesting that errors from the connected part of the covariance are subdominant.

LOWZ sample are consistent with the jackknife and standard deviation errors computed from the QPM mocks. This consistency confirms that the errors are dominated by the shape noise, $N_\gamma(P_{gg}+N_g)$ with some contributions from systematics, while contributions from cosmic variance and super-sample covariance (not included in the signal with the QPM mocks) are subdominant.

As in the case of the LOWZ lens sample, the theoretical predictions for the errors in the left panel of Fig. 3.5 are lower than the actual errors. This is due to the effect of systematics (there are no connected terms in this case). In the right panel, we show the errors estimated by using the rotated (or mock) sources. In this case there are no shear correlations (either cosmological or due to systematics) and the theoretical predictions for the errors are consistent with the measured errors in the case of $\Delta\Sigma_{gR}$. The discrepancies in the case of $\Delta\Sigma_g$ are due to the idealized window function used for the extra term in the theory calculations.

### Randoms

In this section, we measure $\Delta\Sigma$ by replacing the LOWZ galaxies with random lens catalogs. In this case, the covariance only has contributions from terms with lens shot noise, $N_g(P_{\gamma\gamma}+N_\gamma)$, as there is no lens clustering, $P_{gg}=0$. We use 75 random samples that are the same size as the LOWZ sample, along with 10 additional random samples, which are used to compute $\Delta\Sigma_R$. In this section we only show results using the mock source sample, so $P_{\gamma\gamma}=0$.

Fig. 3.6 shows the errors in the $\Delta\Sigma$ measurements using randoms lenses, with and without $\Delta\Sigma_R$ subtracted out. Also shown are the error estimates using the standard deviation of the signal measured across all 75 independent realizations.

In the case of $\Delta\Sigma_{gR}$ errors from jackknife, the errors from the standard deviation and theory are consistent. The errors also follow the expected $1/r_p$ scaling (no lens clustering in this case), except at the largest scales where there are some deviations, possibly due to small amounts of large-scale power

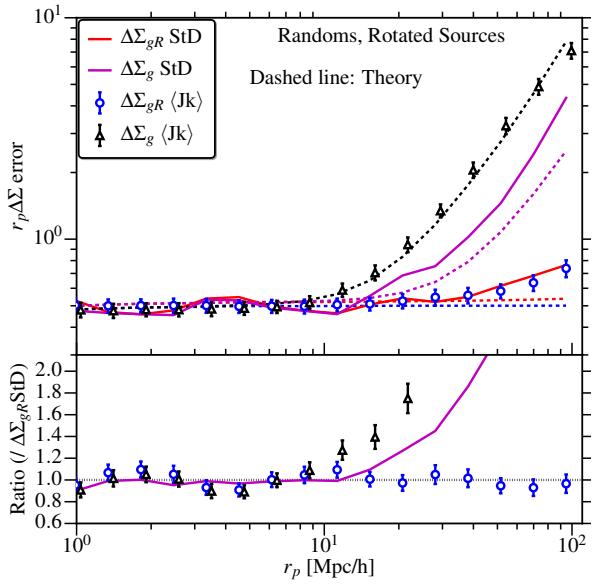



Figure 3.6: Same as Fig. 3.4, now using 75 random realizations of the LOWZ lens catalog (no clustering) with rotated (mock) sources. Since the lens sample has no clustering and sources have no shear correlations, the errors only include shot noise terms and hence scale as $1/r_p$, except at the largest scales where the random points have some clustering as they match the selection function in LOWZ.

in the distribution of the random catalogs that enables them to match the selection function of the LOWZ sample. Also, the errors in the case of the random lenses are in general lower than those for the LOWZ sample or the QPM mocks at large scales due to the effects of lens clustering. Finally, for $\Delta\Sigma_g$ the errors do not follow the typical $1/r_p$ scaling because of the additional $W^2N_\gamma$ term.

### 3.4.4 Putting it all together

Using the results of the previous subsections, we can now understand the contributions of various terms in the covariance, using both theoretical predictions and errors estimated using data and mocks.

In Fig. 3.7 we show the error estimates from various combinations of data and mocks (left panel) and theory calculations using various terms in Eq. (3.10) (right panel). In Table 3.1 we also show various terms that contribute to various combinations of data and mocks.

As shown in table 3.1, we can use the random lenses with the rotated (mock) sources to compute the contributions of lens shot noise and source shape noise to the covariance. Including the lens samples with clustering (LOWZ or QPM) then provides the contribution from the clustering of the lenses. Substituting the real sources with mock lenses (QPM) provides the contributions from shear correlations (systematics or cosmological). Thus we can study all terms except for those arising from lens-source correlations ($P_{g\gamma}$ and $T_{g\gamma g\gamma}$) using the mocks we have used in this paper. Using more realistic simulations as in Shirasaki et al. (2016) will further allow a study of these lens-source correlations terms, though as shown in previous sections, contributions from these terms are subdominant when using the SDSS shape sample.

As demonstrated in Fig. 3.7, at small scales the errors are dominated by the shot noise terms $N_gN_\gamma$, where $N_g$ is the galaxy shot noise power spectrum and $N_\gamma$ is the shape noise power spectrum. At larger scales ($r_p \gtrsim 20h^{-1}$Mpc), the term involving the lens clustering, $P_{gg}N_\gamma$, starts dominating (in the literature this term is commonly referred to as "correlated shape noise"). The contributions from the shear power spectrum terms are in general small, with $P_{\gamma\gamma}$ and $P_{g\gamma}$ terms only contributing $\sim 10\%$ of the error even at $r_p \sim 100h^{-1}$Mpc for this particular survey. In this work we did not compute the trispectrum terms ($T_{g\gamma g\gamma}$), but based on the comparison between the LOWZ and QPM

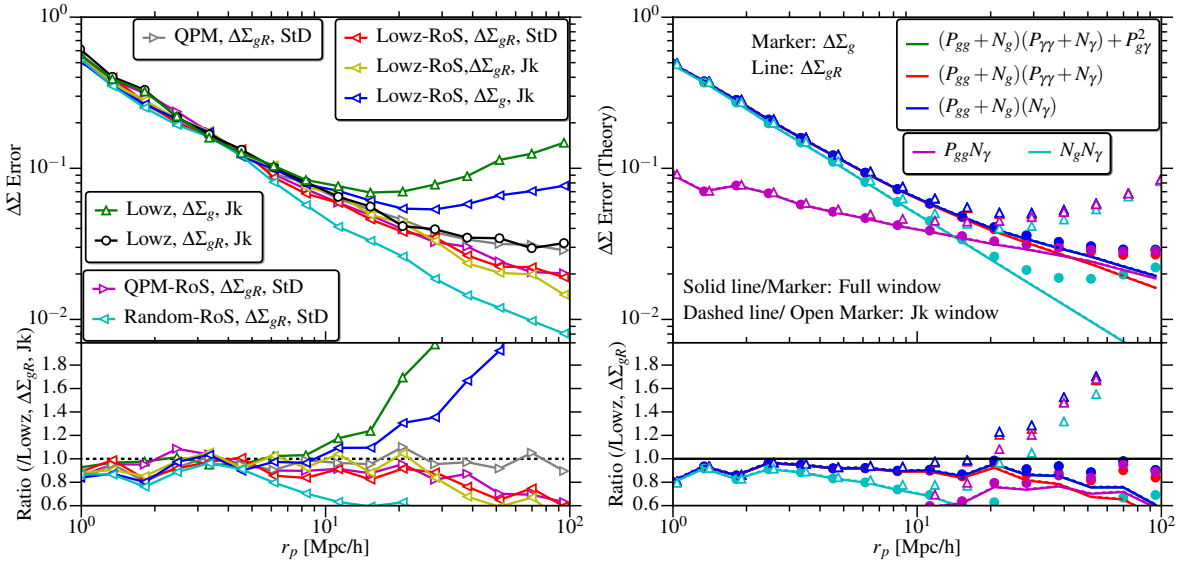



Figure 3.7: Figure summarizing the main error estimates discussed in this paper. Note that unlike in other plots, the y-axis here is the error on $\Delta\Sigma$ error without a factor of $r_p$. In the left panel, we show the errors estimated using different combinations of the data and mock catalogs (RoS stands for "Rotated sources" (or mock sources); but unless explicitly mentioned, the curves use the real sources). In the right panel, we show the error estimates from different terms that contribute to the theoretical covariance using the notation of Eq. (3.10). In the bottom panels, the curves are divided by the jackknife errors on the real LOWZ $\Delta\Sigma_{gR}$, which contains contributions from systematics that are not included in the theoretical expressions, hence the ratios are systematically below 1. For different curves, the power spectra terms not mentioned in the legend are set to zero, e.g. for the cyan curves $(N_gN_\gamma)$ all auto-power spectra are zero. There is still an $\{N_\gamma\}$ term in this case in $\Delta\Sigma_{gR}$.

| Lens-Shape Sample | Error Term |
|---|---|
| LOWZ-SDSS | $(P_{gg}+N_g)(P_{\gamma\gamma}+N_\gamma)+P_{g\gamma}^2+T_{g\gamma g\gamma}$ $+\{W^2(P_{\gamma\gamma}+N_\gamma)\}$ |
| QPM-SDSS | $(P_{gg}+N_g)(P_{\gamma\gamma}+N_\gamma)+\{W^2(P_{\gamma\gamma}+N_\gamma)\}$ |
| LOWZ-Mock | $(P_{gg}+N_g)N_\gamma+\{W^2N_\gamma\}$ |
| QPM-Mock | $(P_{gg}+N_g)N_\gamma+\{W^2N_\gamma\}$ |
| Randoms-Mock | $N_gN_\gamma+\{W^2N_\gamma\}$ |

Table 3.1: Table showing the main sources of statistical uncertainty for different combinations of lens, source and estimators in this work. We use the notation from Eq. (3.10), with $P_{gg}=b_gP_{\delta\delta}$, $N_g$ is the galaxy shot noise term, $N_\gamma$ is the shape noise and the terms in curly brackets $\{\}$ involving $W^2$ are the window function-dependent covariance contributions to $\Delta\Sigma_g$.

lens samples (see Fig. 3.5), we find no evidence that such terms are important at any scale considered in this work for SDSS.

We also show the contributions of the window function-dependent terms in $\Delta\Sigma_g$. As the size of the window function increases, $W(k)$ approaches a delta function, $\delta_D(k)$. As a result, the contribution of these terms to the covariance decreases with increasing window size, which is the reason why the jackknife errors have higher contributions than the standard deviation from the mocks using full survey window. Once these terms are removed, the errors in $\Delta\Sigma_{gR}$ are consistent from both the jackknife and the full window. Normally we do expect $\Delta\Sigma_{gR}$ to be different between the jackknife and the full window due to the edge effects (see results for clustering in appendix 3.B). However, in this work we only apply the jackknife to the lens sample while using the full source sample at all times. Hence the edge effects in the jackknife and full window cases are the same, though our theory curves under-predict these edge effects since we assumed an idealized window function with circular symmetry and no holes (see Appendix 3.A).

### 3.4.5 Comparison of different error estimates

In this section we compare the error estimates from the jackknife method with 100 regions against those from taking the mean and error on the mean from 100 subsamples. The primary motivation for this comparison is to test for edge effects and to check whether the jackknife method underestimates the errors once the scales are close to the size of the subsamples. The subsamples were defined in the same way for both methods and the division was only done on the lens sample. Each subsample/jackknife region is cross- correlated with the entire shape sample. Subsampling on the lens sample alone is sufficient in the shape noise-dominated regime, since the shape noise for different subsamples will be uncorrelated. In the case of other measurements, e.g. clustering, the measurement across different subsamples will get correlated once the length scale approaches the size of the subsample, and the errors will be underestimated in both cases.

Fig. 3.8 shows the comparison of three different error estimates: jackknife, subsampling and standard deviation for the QPM mocks. All error estimates are consistent with each other (within the uncertainties), though the scatter in the subsampling errors is somewhat higher than the jackknife errors. In Fig. 3.9 we also show the correlation matrix for the three different error estimates; they are all consistent with each other.

## 3.5 Conclusions

In this work we have studied the behavior of covariances in galaxy-galaxy lensing measurements using mock catalogs and theoretical predictions, including a comparison of two different galaxy-galaxy lensing estimators with different covariance properties. The mock catalogs include randomly distributed

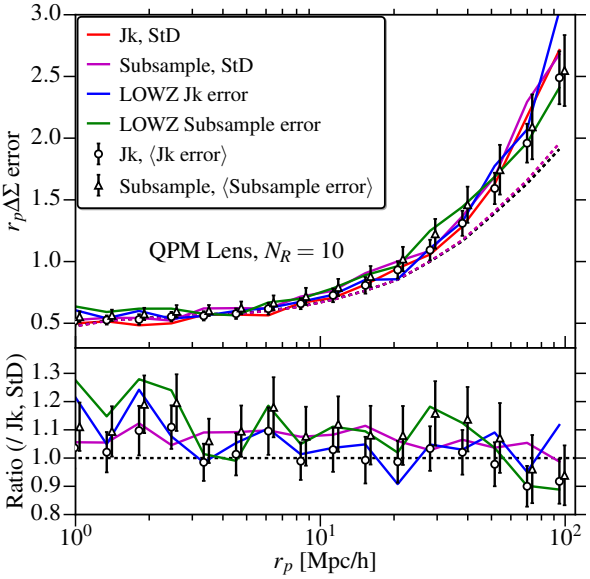



Figure 3.8: Comparison of the error estimates obtained using the jackknife method and the standard deviation across 100 subsamples of the lens samples (QPM except for the two lines labelled "LOWZ"). The same subsamples were used for both methods. All calculations use $\Delta\Sigma_g - \Delta\Sigma_R$ with $N_R = 10$. "Jk, Std" refer to the errors from the jackknife and from the standard deviation of the mock samples as defined in Sec. 3.2.4 for $\Delta\Sigma_{gR}$ (signal, not the noise), while "Subsample, Std" refers to errors from the standard deviations of $\Delta\Sigma_{gR}$ measured as the mean of subsamples in each realization. ⟨Jk error⟩ (⟨Subsample error⟩) is the mean of the jackknife (subsample) error across the realizations. Dashed lines are the corresponding theory predictions.

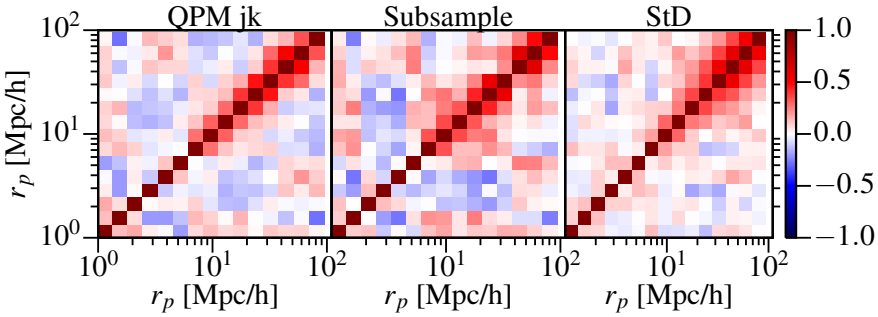



Figure 3.9: Comparison of the correlation matrices for the QPM mocks from the jackknife, subsampling and standard deviation methods across 200 realizations. All calculations use $\Delta\Sigma_g - \Delta\Sigma_R$ with $N_R = 10$.

lenses, QPM mocks that have similar clustering as LOWZ galaxies, and mock source catalogs obtained by randomly rotating the real SDSS source galaxies. Our main results are summarized in Fig. 3.7 and Table 3.1.

For the SDSS, at small scales the covariance is dominated by the pure uncorrelated shape noise, which is white, but at larger scales, contributions from lens and shear correlations also matter. Using mock source catalogs obtained by randomly rotating the sources, we show that the errors are dominated by the terms involving the shape noise, $N_\gamma(P_{gg} + N_g)$. When using the real sources from SDSS, we also found evidence of contributions to the covariance from the systematics to the covariance, even when subtracting random signal (i.e. using lens over-density). While our theory calculations did not include the contributions from systematics or the connected term (including the super-sample variance), the consistency of the covariances when using LOWZ and QPM mocks as lenses demonstrates that contributions from the connected terms are subdominant and the differences between the theoretical predictions and the measurements in Fig. 3.2 arise primarily from the systematics. This conclusion in general depends on the survey configuration, and for different surveys the tradeoffs between shape noise and other covariance contributions must be re-evaluated.

We also demonstrated that the additional variance seen on large scales when not subtracting the shear around random points is only partially contributed by systematics. Even without systematics, on scales where correlated shape noise is important, using the sub-optimal estimator (lens density instead of over-density) for galaxy-galaxy lensing can reduce the per-bin $S/N$ by a substantial factor (up to a factor of 2 on the largest scales considered, which are $\sim 10$ times the correlation length of the lens galaxies). Our covariance calculations suggest that this reduction in the covariance primarily arises from the removal of shear correlations (including shape noise), which only depends on the window function of the lens sample. Covariance estimation methods using subsamples (including the jackknife) have a smaller survey window and hence have a higher contribution from this term compared to the full survey window. The tests using mock lens and mock source catalogs are consistent with this explanation. Our results suggest that the discrepancy between the jackknife error estimates and standard deviation across different realizations observed by Shirasaki et al. (2016) (version 1) can be explained by the fact that they used the lens density instead of the lens over-density in the galaxy-galaxy lensing measurements, given that the contributions from the additional covariance terms due to use of density depend on the survey window. This assertion has been confirmed in the updated work (version 2) of Shirasaki et al. (2016).

In our calculations of covariances, we also identified the effects of the window function, which can be important when comparing the covariance estimations from empirical methods such as jackknife or subsampling, which divide the survey window into smaller parts. Since in our jackknife estimates we only split the lens sample, these effects are not important in our lensing measurements. In appendix 3.B, using clustering measurements, we show that the window function effects can alter the covariance by up to 40-50% on scales approaching the size of the subsamples. Finally, for both clustering (appendix 3.B) and lensing (section 3.4.5), we also demonstrated using mocks that the jackknife errors are consistent with the errors from the subsampling methods for the scales that are smaller than the subsample size at the effective redshift of the sample.

Our results emphasize the importance of using the optimal galaxy-galaxy lensing estimator $\Delta\Sigma_g - \Delta\Sigma_R$ even in the absence of systematics, for g-g lensing estimates that extend to scales above a few Mpc, to obtain better covariance properties and to enable use of internal error estimates like the jackknife. Our conclusions are also applicable to galaxy-CMB lensing cross correlations, as were done by Singh et al. (2016a). Finally we recommend the use of the tests of covariances demonstrated in this work and those in Shirasaki et al. (2016) for ongoing and future surveys, to better understand which terms are dominating the covariances.

## Acknowledgements

We thank François Lanusse, Reiko Nakajima, Ben Wibking, Erin Sheldon and Masahiro Takada for useful discussions related to this work. We also thank Martin White and Jeremy Tinker for providing

the QPM mocks and the SDSS-I,II,III collaborations for providing the datasets used in this work.

SS acknowledges the support from John Peoples Jr. Presidential Fellowship from Carnegie Mellon University. RM acknowledges the support of the Department of Energy Early Career Award program. US acknowledges support of NASA grant NNX15AL17G.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## 3.A Covariance

### 3.A.1 General case

Here we derive the expression for the covariance of the cross-correlation function of two fields with non-zero mean. The results depend on the estimator used for that cross-correlation function, as we will show explicitly below (see also Landy & Szalay 1993), and directly motivate the use of estimators that involve subtraction of the mean density for both fields. While we will use the example of clustering in this section, the results are in general true for any tracer of large-scale structure. In Section 3.A.3 we will use the results from this section to compute the covariance for the galaxy-shear cross-correlation function.

We are interested in the cross-correlation function of two (biased) tracer fields, $g_X, g_Y$, of the matter density field ($\rho_m$, not just $\delta\rho_m$)

$$g_i = (1 + \delta_i + n_i)M_i \tag{3.18}$$

where $i = X$ or $Y$, $M_i$ is the mean value of the field (mean number density in the case of galaxies) and $n_i$ is the noise in the tracer field (shot noise in the case of galaxies, shape noise in the case of shear). Hereafter, in this section we will assume that the field is normalized so that $M_i = 1$. For notational compactness, we also define

$$\widehat{\delta}_i = \delta_i + n_i \tag{3.19}$$

$$g_i = 1 + \widehat{\delta}_i \tag{3.20}$$

In Fourier space

$$\tilde{g}_i(\mathbf{k}) = \delta_D(\mathbf{k}) + \tilde{\delta}_i(\mathbf{k}) + \tilde{n}_i(\mathbf{k}) = \delta_D(\mathbf{k}) + \tilde{\widehat{\delta}}_i(\mathbf{k}), \tag{3.21}$$

where $\delta_D$ is the Dirac delta function.

We can write the cross-correlation function of two fields as (analogous to normalized $\frac{DD}{RR} - 1$)

$$\widehat{\xi}_{XY}(\mathbf{r}) = \widehat{\xi}_{g_1 g_2}(\mathbf{r}) = \frac{1}{\mathcal{V}_W(\mathbf{r})} \int d^3\mathbf{r}' W(\mathbf{r}'+\mathbf{r}) W(\mathbf{r}') \left[g_1(\mathbf{r}') g_2(\mathbf{r}'+\mathbf{r}) - 1\right] \tag{3.22}$$

$$= \frac{1}{\mathcal{V}_W(\mathbf{r})} \int d^3\mathbf{r}' W(\mathbf{r}'+\mathbf{r}) W(\mathbf{r}') \left[\widehat{\delta}_1(\mathbf{r}')\widehat{\delta}_2(\mathbf{r}'+\mathbf{r}) + \widehat{\delta}_1(\mathbf{r}') + \widehat{\delta}_2(\mathbf{r}'+\mathbf{r}) + 1\right] - 1 \tag{3.23}$$

$$= \frac{1}{\mathcal{V}_W(\mathbf{r})} \int d^3\mathbf{r}' W(\mathbf{r}'+\mathbf{r}) W(\mathbf{r}') \widehat{\delta}_1(\mathbf{r}') \widehat{\delta}_2(\mathbf{r}'+\mathbf{r}) \tag{3.24}$$

$$= \frac{1}{\mathcal{V}_W(\mathbf{r})} \int d^3\mathbf{r}' W(\mathbf{r}'+\mathbf{r}) W(\mathbf{r}') \left[\delta_1(\mathbf{r}') \delta_2(\mathbf{r}'+\mathbf{r}) + n_1(\mathbf{r}') n_2(\mathbf{r}'+\mathbf{r})\right] \tag{3.25}$$

$$\widehat{\xi}_{XY}(\mathbf{r}) = \widehat{\xi}_{g_1 g_2}(\mathbf{r}) = \xi_{XY}(\mathbf{r}) + \xi_{n_X n_Y}(\mathbf{r}) \tag{3.26}$$

where $g_1$ belongs to field $X$ and $g_2$ to field $Y$. $W(\mathbf{r})$ is the survey window function. We have assumed that the noise and $\delta_i$ have zero mean and are also uncorrelated with each other on all scales. The normalization factor is the integral over window functions

$$\mathcal{V}_W(\mathbf{r}) = \int d^3\mathbf{r}' W(\mathbf{r}'+\mathbf{r}) W(\mathbf{r}') = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{r}} \tilde{W}(\mathbf{k}) \tilde{W}(-\mathbf{k}) \tag{3.27}$$

The covariance of the correlation function is given as

$$\mathrm{Cov}(\widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j)) = \left\langle \widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j) \right\rangle - \left\langle \widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \right\rangle \left\langle \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j) \right\rangle, \tag{3.28}$$

where $g_1, g_3$ belong to field $X$ and $g_2, g_4$ belong to $Y$. Using Eq. (3.22)

$$\mathrm{Cov}(\widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j)) = \Bigg\langle \frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int d^3\mathbf{r} \int d^3\mathbf{r}' W(\mathbf{r}) W(\mathbf{r}') W(\mathbf{r}+\mathbf{r}_i) W(\mathbf{r}'+\mathbf{r}_j)$$
$$\left[g_1(\mathbf{r}) g_2(\mathbf{r}+\mathbf{r}_i) g_3(\mathbf{r}') g_4(\mathbf{r}'+\mathbf{r}_j) - g_1(\mathbf{r}) g_2(\mathbf{r}+\mathbf{r}_i) - g_3(\mathbf{r}') g_4(\mathbf{r}'+\mathbf{r}_j) + 1\right] \Bigg\rangle$$
$$- \left\langle \widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \right\rangle \left\langle \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j) \right\rangle \tag{3.29}$$

$$\mathrm{Cov}(\widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j)) = \Bigg\langle \frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int d^3\mathbf{r} \int d^3\mathbf{r}' W(\mathbf{r}) W(\mathbf{r}') W(\mathbf{r}+\mathbf{r}_i) W(\mathbf{r}'+\mathbf{r}_j)$$
$$g_1(\mathbf{r}) g_2(\mathbf{r}+\mathbf{r}_i) g_3(\mathbf{r}') g_4(\mathbf{r}'+\mathbf{r}_j) \rangle - \frac{\left\langle \widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \right\rangle}{\mathcal{V}_W(\mathbf{r}_i)} - \frac{\left\langle \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j) \right\rangle}{\mathcal{V}_W(\mathbf{r}_j)} - 1$$
$$- \left\langle \widehat{\xi}_{g_1 g_2}(\mathbf{r}_i) \right\rangle \left\langle \widehat{\xi}_{g_3 g_4}(\mathbf{r}_j) \right\rangle \tag{3.30}$$

We use $\langle g_1 g_2 g_3 g_4 \rangle_{ij}$ as short-hand for the first term in Eq. (3.30), which we would like to simplify.

$$\langle g_1 g_2 g_3 g_4 \rangle_{ij} = \left\langle \frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int d^3\mathbf{r} \int d^3\mathbf{r}' g_1(\mathbf{r}) g_2(\mathbf{r}+\mathbf{r}_i) g_3(\mathbf{r}') g_4(\mathbf{r}'+\mathbf{r}_j) W(\mathbf{r}) W(\mathbf{r}') W(\mathbf{r}+\mathbf{r}_i) W(\mathbf{r}'+\mathbf{r}_j) \right\rangle \tag{3.31}$$

Writing the $g_i$ in terms of its Fourier space counterpart $\tilde{g}_i$, we get

$$\langle g_1 g_2 g_3 g_4 \rangle_{ij} = \frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int d^3\mathbf{r} \int d^3\mathbf{r}' \iiiint \prod_{n=1}^{4} \left[\frac{d^3\mathbf{k}_n}{(2\pi)^3}\right] \iiiint \prod_{m=1}^{4} \left[\frac{d^3\mathbf{q}_m}{(2\pi)^3}\tilde{W}(\mathbf{q}_m)\right] \times e^{i(\mathbf{k}_1-\mathbf{q}_1)\cdot\mathbf{r}} e^{i(\mathbf{k}_2-\mathbf{q}_2)\cdot(\mathbf{r}+\mathbf{r}_i)} e^{i(\mathbf{k}_3-\mathbf{q}_3)\cdot\mathbf{r}'} e^{i(\mathbf{k}_4-\mathbf{q}_4)\cdot(\mathbf{r}'+\mathbf{r}_j)} \langle \tilde{g}_1(\mathbf{k}_1)\tilde{g}_2(\mathbf{k}_2)\tilde{g}_3(\mathbf{k}_3)\tilde{g}_4(\mathbf{k}_4)\rangle \tag{3.32}$$

$$\langle g_1 g_2 g_3 g_4 \rangle_{ij} = \frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \iint \frac{d^3\mathbf{k}_1}{(2\pi)^3}\frac{d^3\mathbf{k}_3}{(2\pi)^3} \iiiint \prod_{m=1}^{4} \left[\frac{d^3\mathbf{q}_m}{(2\pi)^3}\tilde{W}(\mathbf{q}_m)\right] e^{-i(\mathbf{k}_1-\mathbf{q}_1)\cdot\mathbf{r}_i} e^{-i(\mathbf{k}_3-\mathbf{q}_3)\cdot\mathbf{r}_j} \times \langle \tilde{g}_1(\mathbf{k}_1)\tilde{g}_2(-\mathbf{k}_1+\mathbf{q}_1+\mathbf{q}_2)\tilde{g}_3(\mathbf{k}_3)\tilde{g}_4(-\mathbf{k}_3+\mathbf{q}_3+\mathbf{q}_4)\rangle\,. \tag{3.33}$$

We have integrated over $d^3\mathbf{r}$ and $d^3\mathbf{r}'$ and then over $d^3\mathbf{k}_2$ and $d^3\mathbf{k}_4$, to obtain the last expression.

We now expand the four-point function into two separable parts: the connected or non-Gaussian component $\left\langle \tilde{\delta}_1\tilde{\delta}_2\tilde{\delta}_3\tilde{\delta}_4 \right\rangle'$ and the Gaussian component, which using Wick's theorem can be expanded as the sum of the product of two-point functions.

$$\begin{aligned}\langle g_1 g_2 g_3 g_4 \rangle_{ij} = \frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \iint \frac{d^3\mathbf{k}_1}{(2\pi)^3}\frac{d^3\mathbf{k}_3}{(2\pi)^3} \iiiint \prod_{m=1}^{4} \left[\frac{d^3\mathbf{q}_m}{(2\pi)^3}\tilde{W}(\mathbf{q}_m)\right] e^{-i(\mathbf{k}_1-\mathbf{q}_1)\cdot\mathbf{r}_i} e^{-i(\mathbf{k}_3-\mathbf{q}_3)\cdot\mathbf{r}_j} \Big[ \left\langle \tilde{\delta}_1\tilde{\delta}_2\tilde{\delta}_3\tilde{\delta}_4 \right\rangle' + \\ \left\langle \tilde{\delta}_1\tilde{\delta}_2 \right\rangle \left\langle \tilde{\delta}_3\tilde{\delta}_4 \right\rangle + [\langle \delta_{D,1}\delta_{D,2}\rangle + \langle \tilde{n}_1\tilde{n}_2\rangle] \left\langle \tilde{\delta}_3\tilde{\delta}_4 \right\rangle + \left\langle \tilde{\delta}_1\tilde{\delta}_2 \right\rangle [\langle \delta_{D,3}\delta_{D,4}\rangle + \langle \tilde{n}_3\tilde{n}_4\rangle] + \\ \left\langle \tilde{\delta}_1\tilde{\delta}_3 \right\rangle \left\langle \tilde{\delta}_2\tilde{\delta}_4 \right\rangle + [\langle \delta_{D,1}\delta_{D,3}\rangle + \langle \tilde{n}_1\tilde{n}_3\rangle] \left\langle \tilde{\delta}_2\tilde{\delta}_4 \right\rangle + \left\langle \tilde{\delta}_1\tilde{\delta}_3 \right\rangle [\langle \delta_{D,2}\delta_{D,4}\rangle + \langle \tilde{n}_2\tilde{n}_4\rangle] + \\ \left\langle \tilde{\delta}_1\tilde{\delta}_4 \right\rangle \left\langle \tilde{\delta}_2\tilde{\delta}_3 \right\rangle + [\langle \delta_{D,1}\delta_{D,4}\rangle + \langle \tilde{n}_1\tilde{n}_4\rangle] \left\langle \tilde{\delta}_2\tilde{\delta}_3 \right\rangle + \left\langle \tilde{\delta}_1\tilde{\delta}_4 \right\rangle [\langle \delta_{D,2}\delta_{D,3}\rangle + \langle \tilde{n}_2\tilde{n}_3\rangle] + \\ \langle \delta_{D,1}\delta_{D,2}\tilde{n}_3\tilde{n}_4\rangle + \text{all perms} + \\ \langle \tilde{n}_1\tilde{n}_2\tilde{n}_3\tilde{n}_4\rangle + \langle \delta_{D,1}\delta_{D,2}\delta_{D,3}\delta_{D,4}\rangle \Big]\end{aligned} \tag{3.34}$$

We have omitted the positional arguments for $\tilde{\delta}_i$ and $\tilde{n}_i$, which are the same as for $g_i$ in Eq. (3.33). We defined $\delta_{D,i} = \delta_D(\mathbf{k}_i)$ and $\left\langle \tilde{\delta}_i\tilde{\delta}_j \right\rangle = P_{ij}(\mathbf{k}_i)\delta_D(\mathbf{k}_i+\mathbf{k}_j)$, where $P_{ij}(\mathbf{k})$ is the power spectrum.

Simplifying, the terms involving $\langle \xi_{12}(\mathbf{r}_i)\rangle \langle \xi_{34}(\mathbf{r}_j)\rangle$ cancel out, and using the fact that $g_1$ and $g_3$ belonged to field $X$ and $g_2$ and $g_4$ belonged to field $Y$, we can write the covariance as

$$\begin{aligned}\text{Cov} = &\left[\frac{\mathcal{V}_W(\mathbf{r}_i-\mathbf{r}_j)}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{r}_i} e^{i\mathbf{k}\cdot\mathbf{r}_j} \widehat{P}_{XX}(\mathbf{k})\widehat{P}_{YY}(\mathbf{k}) + \frac{\mathcal{V}_W(\mathbf{r}_i+\mathbf{r}_j)}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{r}_i} e^{-i\mathbf{k}\cdot\mathbf{r}_j} \right. \\ &\left. \widehat{P}_{XY}(\mathbf{k})\widehat{P}_{XY}(\mathbf{k}) + T_{XYXY}\right] \\ &+ \left\{\frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{r}_i} e^{i\mathbf{k}\cdot\mathbf{r}_j} \tilde{W}(\mathbf{k})\tilde{W}(-\mathbf{k}) \left(\widehat{P}_{YY}(\mathbf{k}) + \widehat{P}_{XX}(\mathbf{k})\right)\right\} \\ &+ \left\{\frac{1}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)} \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{r}_i} e^{-i\mathbf{k}\cdot\mathbf{r}_j} \tilde{W}(\mathbf{k})\tilde{W}(\mathbf{k}) \left(\widehat{P}_{XY}(\mathbf{k}) + \widehat{P}_{XY}(\mathbf{k})\right)\right\}\end{aligned} \tag{3.35}$$

Here $\widehat{P}_{ij} = P_{ij} + P_{ij,N}$, where $P_{ij,N}$ is the noise power spectrum. $T_{XYXY}$ is the connected term. To simplify expressions, we have assumed that the power spectrum is a slowly varying function of $k$ and that we are working with modes much smaller than the survey size, so that $P(\mathbf{k}-\mathbf{q}) \approx P(\mathbf{k})$ and then $P(\mathbf{k})$ can be moved out of the window function integrals. For scales much smaller than the survey size, $\mathcal{V}_W(\mathbf{r}) \to V_W$, where $V_W$ is the physical volume of the survey, the expression simplifies to the more familiar form

$$\begin{aligned}\text{Cov} = &\left[\frac{1}{V_W}\left(\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{-i\mathbf{k}\cdot\mathbf{r}_i}e^{i\mathbf{k}\cdot\mathbf{r}_j}\widehat{P}_{XX}(\mathbf{k})\widehat{P}_{YY}(\mathbf{k})+\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{-i\mathbf{k}\cdot\mathbf{r}_i}e^{-i\mathbf{k}\cdot\mathbf{r}_j}\widehat{P}_{XY}(\mathbf{k})\widehat{P}_{XY}(\mathbf{k})\right)+T_{XYXY}\right]\\&+\left\{\frac{1}{V_W^2}\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{-i\mathbf{k}\cdot\mathbf{r}_i}e^{i\mathbf{k}\cdot\mathbf{r}_j}\tilde{W}(\mathbf{k})\tilde{W}(-\mathbf{k})\left(\widehat{P}_{YY}(\mathbf{k})+\widehat{P}_{XX}(\mathbf{k})\right)\right\}\\&+\left\{\frac{1}{V_W^2}\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{-i\mathbf{k}\cdot\mathbf{r}_i}e^{-i\mathbf{k}\cdot\mathbf{r}_j}\tilde{W}(\mathbf{k})\tilde{W}(\mathbf{k})\left(\widehat{P}_{XY}(\mathbf{k})+\widehat{P}_{XY}(\mathbf{k})\right)\right\}\end{aligned} \tag{3.36}$$

The terms in square brackets ([ ]) are the usual covariance terms while the terms in braces ({}) arise when the means of the fields are not subtracted. These additional contributions depend on the survey window function and become less important as the survey size increases. In the case of a large uniform survey, $\lim_{V_W\to\infty}\tilde{W}(\mathbf{k})=\delta_D(\mathbf{k})$. As a result, the terms in braces ({}) approach zero faster (under the assumption $\widehat{P}(\mathbf{k}=0)=0$) and the two estimators (correlating mean zero field or correlating mean non-zero fields) are equivalent. However, in case $\widehat{P}(\mathbf{k}=0)\neq 0$, e.g., due to shot noise in case of galaxies, the ({}) terms approach the value of $P(\mathbf{k}=0)$. We also emphasize that this additional contribution to the covariance will be present in the analysis in Fourier space as well.

### 3.A.2 Projected Case

The projected correlation function is defined as the integral of the 3-d correlation function over the line of sight separation, Π.

$$\widehat{w}(r_p)=\int_{\Pi_{\min}}^{\Pi_{\max}}d\Pi W(\Pi)\widehat{\xi}(r_p,\Pi) \tag{3.37}$$

where $W(\Pi)$ is the line-of-sight weight function (not necessarily the same as the window function).

To compute the covariance, we start with Eq. (3.34), carry out the line-of-sight integrals assuming the integration length is long (getting delta functions of the form $\delta_D(\mathbf{k}_{\parallel,i}-\mathbf{q}_{\parallel,i})$). Then carrying out integrals involving the line-of-sight window functions, we assume that the relevant line-of-sight modes are small, such that power spectrum is only dependent on the projected modes ($k_\parallel \ll k_\perp$, $P(\mathbf{k})\approx P(\mathbf{k}_\perp)$)

$$\begin{aligned}\text{Cov} = &\int d\Pi W_Y(\Pi)W_Y(\Pi)\left[\frac{\mathcal{V}_W(\mathbf{r}_i-\mathbf{r}_j)}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\widehat{P}_{XX}(\mathbf{k}_\perp)\widehat{P}_{YY}(\mathbf{k}_\perp)\right.\\&\left.+\frac{\mathcal{V}_W(\mathbf{r}_i+\mathbf{r}_j)}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\widehat{P}_{XY}(\mathbf{k}_\perp)\widehat{P}_{XY}(\mathbf{k}_\perp)+T_{XYXY}\right]\\&+\int d\Pi W_Y(\Pi)W_Y(\Pi)\left\{\frac{L_W}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\tilde{W}_X(\mathbf{k}_\perp)\tilde{W}_X(-\mathbf{k}_\perp)\right.\\&\left(\widehat{P}_{YY}(\mathbf{k}_\perp)+\widehat{P}_{XX}(\mathbf{k}_\perp)\right)\\&\left.+\frac{L_W}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\tilde{W}_X(\mathbf{k}_\perp)\tilde{W}_X(\mathbf{k}_\perp)\left(\widehat{P}_{XY}(\mathbf{k}_\perp)+\widehat{P}_{XY}(\mathbf{k}_\perp)\right)\right\}\end{aligned} \tag{3.38}$$

Here we distinguished between the window functions of tracers X and Y, $L_W$ is the line-of-sight length of the window function (of $X$) and we ignore the edge effects along the line-of-sight. Thus the volume element can be written as

$$\mathcal{V}_W(\mathbf{r}_p)=\mathcal{A}_W(\mathbf{r}_p)L_W \tag{3.39}$$

$\mathcal{A}_W$ is the physical survey area at the lens redshift.

Note that $\widehat{P}_{YY}$ can in principle be evaluated at a different epoch as $Y_1$ and $Y_2$ are at separation $\Pi_i$ and $\Pi_j$, i.e., $P_{YY}(k_\perp) \sim P_{YY}(k_\perp \frac{\chi_z}{\chi_z+\Pi})$ where $\chi_z$ is the line-of-sight distance to the mean redshift where we are evaluating the covariance. Under the assumption that the power spectrum evolution within the $\Pi_{\rm max}$ limits is small, we keep $P_{YY}(k_\perp)$ (ignoring its $\Pi$ dependence), and simplify the expression as

$$\begin{aligned}
{\rm Cov} = &\left[\frac{\mathcal{A}_W(\mathbf{r}_{p,i}-\mathbf{r}_{p,j})}{\mathcal{A}_W(\mathbf{r}_{p,i})\mathcal{A}_W(\mathbf{r}_{p,j})}\frac{\Delta\Pi_2}{L_W}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\widehat{P}_{XX}(\mathbf{k}_\perp)\widehat{P}_{YY}(\mathbf{k}_\perp)\right.\\
&\left.+\frac{\mathcal{A}_W(\mathbf{r}_i+\mathbf{r}_j)}{\mathcal{A}_W(\mathbf{r}_i)\mathcal{A}_W(\mathbf{r}_j)}\frac{\Delta\Pi_2}{L_W}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\widehat{P}_{XY}(\mathbf{k}_\perp)\widehat{P}_{XY}(\mathbf{k}_\perp)+T_{XYXY}\right]\\
&+\frac{\Delta\Pi_2 L_W^2}{\mathcal{V}_W(\mathbf{r}_{p,i})\mathcal{V}_W(\mathbf{r}_{p,j})}\left\{\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^3}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\tilde{W}(\mathbf{k}_\perp)\tilde{W}(-\mathbf{k}_\perp)\left(\widehat{P}_{YY}(\mathbf{k}_\perp)+\widehat{P}_{XX}(\mathbf{k}_\perp)\right)\right.\\
&\left.+\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^3}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\tilde{W}(\mathbf{k}_\perp)\tilde{W}(\mathbf{k}_\perp)\left(\widehat{P}_{XY}(\mathbf{k}_\perp)+\widehat{P}_{XY}(\mathbf{k}_\perp)\right)\right\} \qquad (3.40)
\end{aligned}$$

where we defined

$$\Delta\Pi_2 = \int d\Pi W(\Pi)W(\Pi) \qquad (3.41)$$

$$\Delta\Pi_1 = \int d\Pi W(\Pi) \qquad (3.42)$$

For the case of galaxy clustering, we assume $W(\Pi)$ is a top-hat function for $\Pi \in [-100, 100]$, which leads to $\Delta\Pi_2 = \Delta\Pi_1 = 200h^{-1}$Mpc.

### 3.A.3 Galaxy lensing case

We now use the formalism of Appendix 3.A.1 and 3.A to derive the covariance for the galaxy-galaxy lensing case. We will assume the same sky coverage for the lens and shape samples. Note that the shear is a mean-zero field since lensing is only sensitive to the matter density contrast, and hence some of the terms in Eq. (3.35) will drop out.

We begin by defining the observed shear as the sum of the true shear and noise.

$$\widehat{\gamma} = \gamma + \gamma_N \qquad (3.43)$$

We also assume that the mean shear around random points is not subtracted. In that case, the galaxy-shear (projected) cross-correlation can be written as

$$\Delta\Sigma(\mathbf{r}_p) = \frac{1}{\mathcal{V}_W(\mathbf{r})}\int d^3\mathbf{r}\Sigma_c(z_l,z_s)g(\mathbf{r})\widehat{\gamma}(\mathbf{r}+\mathbf{r}_p)W_\gamma(\mathbf{r}+\mathbf{r}_p)W_g(\mathbf{r}) \qquad (3.44)$$

$$= \frac{1}{\mathcal{V}_W(\mathbf{r})}\int d^3\mathbf{r}\Sigma_c(z_l,z_s)(1+\widehat{\delta})(\mathbf{r})\widehat{\gamma}(\mathbf{r}+\mathbf{r}_p)W_\gamma(\mathbf{r}+\mathbf{r}_p)W_g(\mathbf{r}) \qquad (3.45)$$

$$\Delta\Sigma(\mathbf{r}_p) = \frac{1}{\mathcal{V}_W(\mathbf{r})}\int d^3\mathbf{r}\Sigma_c(z_l,z_s)\widehat{\delta}(\mathbf{r})\widehat{\gamma}(\mathbf{r}+\mathbf{r}_p)W_\gamma(\mathbf{r}+\mathbf{r}_p)W_g(\mathbf{r}) \qquad (3.46)$$

The covariance of two $\Delta\Sigma$ is

$${\rm Cov}(\Delta\Sigma_{g_1\gamma_2}(\mathbf{r}_{p,i})\Delta\Sigma_{g_3\gamma_4}(\mathbf{r}_{p,j})) = \langle\Delta\Sigma_{g_1\gamma_2}(\mathbf{r}_{p,i})\Delta\Sigma_{g_3\gamma_4}(\mathbf{r}_{p,j})\rangle - \langle\Delta\Sigma_{g_1\gamma_2}(\mathbf{r}_{p,i})\rangle\langle\Delta\Sigma_{g_3\gamma_4}(\mathbf{r}_{p,j})\rangle \qquad (3.47)$$

Following the derivation in appendix 3.A.1 and 3.A and noting that the shear has a mean of zero, and

$\Delta\Sigma$ is a projected galaxy-matter correlation function, the full covariance analogous to Eq. (3.38) is

$$\begin{aligned}\mathrm{Cov} = &\left[\int d\Pi W(\Pi)W(\Pi)\frac{\mathcal{V}_W(\mathbf{r}_i-\mathbf{r}_j)}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_i}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_j}\widehat{P}_{gg}(\mathbf{k}_\perp)\widehat{P}_{\delta\delta}(\mathbf{k}_\perp)\right]\\ &+\left[\int d\Pi W(\Pi)W(\Pi)\frac{\mathcal{V}_W(\mathbf{r}_i+\mathbf{r}_j)}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\widehat{P}_{g\delta}(\mathbf{k}_\perp)\widehat{P}_{g\delta}(\mathbf{k}_\perp)+T_{g\gamma g\gamma}\right]\\ &+\int d\Pi W(\Pi)W(\Pi)\left\{\frac{L_W^2}{\mathcal{V}_W(\mathbf{r}_i)\mathcal{V}_W(\mathbf{r}_j)}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^3}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_i}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_j}\tilde{W}(\mathbf{k}_\perp)\tilde{W}(-\mathbf{k}_\perp)\widehat{P}_{\delta\delta}(\mathbf{k}_\perp)\right\}\end{aligned} \tag{3.48}$$

where the lensing window function is

$$W(\Pi)=\overline{\rho}\frac{\Sigma_c(\chi_s,\chi_l)}{\Sigma_c(\chi_s,\chi_l+\Pi)} \tag{3.49}$$

The line-of-sight integral in the terms involving $P_{\delta\delta}$ leads to the shear auto-correlation function, and the final expression is

$$\begin{aligned}\mathrm{Cov} = &\left[\frac{\mathcal{V}_W(\mathbf{r}_{p,i}-\mathbf{r}_{p,j})}{\mathcal{V}_W(\mathbf{r}_{p,i})\mathcal{V}_W(\mathbf{r}_{p,j})}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_i}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\cos 2\phi_{k,i}\cos 2\phi_{k,j}\widehat{P}_{gg}(\mathbf{k}_\perp)\Sigma_c^2\left(\frac{\sigma_\gamma^2}{n_s}+P_{\kappa\kappa}\right)\right]\\ &+\left[\frac{\mathcal{V}_W(\mathbf{r}_{p,i}+\mathbf{r}_{p,j})\Delta\Pi_2}{\mathcal{V}_W(\mathbf{r}_{p,i})\mathcal{V}_W(\mathbf{r}_{p,j})}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^2}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\cos 2\phi_{k,i}\cos 2\phi_{k,j}\overline{\rho}^2\widehat{P}_{g\delta}(\mathbf{k}_\perp)\widehat{P}_{g\delta}(\mathbf{k}_\perp)+T_{g\gamma g\gamma}\right]\\ &+\left\{\frac{L_W^2}{\mathcal{V}_W(\mathbf{r}_{p,i})\mathcal{V}_W(\mathbf{r}_{p,j})}\int\frac{d^2\mathbf{k}_\perp}{(2\pi)^3}e^{-i\mathbf{k}_\perp\cdot\mathbf{r}_{p,i}}e^{i\mathbf{k}_\perp\cdot\mathbf{r}_{p,j}}\cos 2\phi_{k,i}\cos 2\phi_{k,j}\tilde{W}(\mathbf{k}_\perp)\tilde{W}(-\mathbf{k}_\perp)\Sigma_c^2\left(\frac{\sigma_\gamma^2}{n_s}+P_{\kappa\kappa}\right)\right\}\end{aligned} \tag{3.50}$$

where the convergence power spectrum is

$$P_{\kappa\kappa}(k)=\int_0^{\chi_s}d\chi\frac{\overline{\rho}}{\Sigma_c(\chi,\chi_s)}\frac{\overline{\rho}}{\Sigma_c(\chi,\chi_s)}P_{\delta\delta}\left(k\frac{\chi_l}{\chi}\right) \tag{3.51}$$

For lensing, using the full lens and source redshift distribution, we compute $\Delta\Pi_1 \approx 900h^{-1}$Mpc and $\Delta\Pi_2 \approx 700h^{-1}$Mpc. Again as in appendix 3.A.1, the terms in square brackets ([ ]) are the usual covariance terms for the mean zero fields and the terms in curly brackets ({ }) are additional contribution from terms involving $\delta_{D,i}$, arising from the sub-optimal estimator without the mean subtracted.

Note that this additional contribution only depends on the window function of the lens sample and is independent of the clustering or number density of the lens sample. Hence, this noise term is consistent across the real lens galaxy sample and uniformly-distributed random points, which is why the subtraction of the shear around random points removes this contribution to the covariance. Also the window function-dependence of this term is the reason why the jackknife/subsample methods of estimating errors shows increased contribution from this term compared to the standard deviation across independent mock catalogs, since the effective window function for the subsamples is smaller.

### 3.A.4 Numerical estimates

#### Clustering

To compute numerical estimates, we assume angular symmetry for both the power spectra and the window function. Further, in the case of galaxy clustering, $X \equiv Y$ and $P_{XX} = b_g^2 P_{\delta\delta}$. After carrying out the angular and line-of-sight integrals (for the projected correlation function) in Eq. (3.40), we

get

$$\mathrm{Cov}(w_{gg}) = \left[2\frac{\mathcal{A}_W(|\mathbf{r}_{p,i}-\mathbf{r}_{p,j}|)}{\mathcal{A}_W(r_{p,i})\mathcal{A}_W(r_{p,j})}\frac{\Delta\Pi_2}{L_W}\int\frac{dkk}{2\pi}J_0(kr_{p,i})J_0(kr_{p,j})\left(b_g^2P_{\delta\delta}(k)+\frac{1}{n_g}\right)^2+T_{gggg}\right] + 4\left\{\frac{2\Delta\Pi_2}{\mathcal{A}_W(r_{p,i})\mathcal{A}_W(r_{p,j})}\int\frac{dkk}{2\pi}J_0(kr_{p,i})J_0(kr_{p,j})\tilde{W}(k)\tilde{W}(k)\left(b_g^2P_{\delta\delta}(k)+\frac{1}{n_g}\right)\right\} \quad (3.52)$$

where $J_n$ is the bessel function of order $n$, $\Pi_{\max}$ is the line-of-sight integration length and

$$\mathcal{A}_W(r) = \int\frac{dk\,k}{2\pi}J_0(kr)\left(W(k)\right)^2 \quad (3.53)$$

$$\mathcal{A}_W(|\mathbf{r}_i-\mathbf{r}_j|) = \int\frac{dk\,k}{2\pi}J_0(kr_i)J_0(kr_j)\left(W(k)\right)^2. \quad (3.54)$$

For the window function, we assume a circular geometry with a survey area of 9000 degree$^2$, with a mean redshift of $z=0.27$. $W(k)$ is defined as

$$W(k) = 2\pi R^2\frac{J_1(kR)}{kR} \quad (3.55)$$

where $R\approx 1275h^{-1}$Mpc is the physical scale corresponding to 95 degrees at $z=0.27$.

Finally, to get the covariance for bins in $r_p$, $\mathrm{COV}_{\rm bin}$, we integrate the covariance in Eq. (3.52) as

$$\mathrm{COV}_{\rm bin} = \frac{\int_{r_{p,i,l}}^{r_{p,i,h}}dr'_{p,i}r'_{p,i}\int_{r_{p,j,l}}^{r_{p,j,h}}dr'_{p,j}r'_{p,j}\mathrm{COV}(r'_{p,i},r'_{p,j})}{\int_{r_{p,i,l}}^{r_{p,i,h}}dr'_{p,i}r'_{p,i}\int_{r_{p,j,l}}^{r_{p,j,h}}dr'_{p,j}r'_{p,j}} \quad (3.56)$$

where $r_{p,i,l}, r_{p,i,h}$ are the lower and upper limits of the bins, respectively.

**Galaxy Lensing**

We carry out the angular integrals in Eq. (3.50), to get

$$\mathrm{Cov}(\Delta\Sigma) = \left[\frac{\mathcal{A}_W(\mathbf{r}_{p,i}-\mathbf{r}_{p,j})}{\mathcal{A}_W(\mathbf{r}_{p,i})\mathcal{A}_W(\mathbf{r}_{p,j})}\frac{1}{L_W}\int\frac{dkk}{2\pi}J_2(kr_{p,i})J_2(kr_{p,j})\Sigma_c^2\left(b_g^2P_{\delta\delta}(k)+\frac{1}{n_g}\right)\left(P_{\kappa\kappa}(k)+\frac{\sigma_\gamma^2}{n_s}\right)\right] + \left[\frac{\mathcal{A}_W(\mathbf{r}_{p,i}-\mathbf{r}_{p,J})}{\mathcal{A}_W(\mathbf{r}_{p,i})\mathcal{A}_W(\mathbf{r}_{p,J})}\frac{\Delta\Pi_2}{L_W}\int\frac{dkk}{2\pi}J_2(kr_{p,i})J_2(kr_{p,j})\left(b_gr_{cc}\overline{\rho}P_{\delta\delta}(k)\right)^2+T_{g\gamma g\gamma}\right] + \left\{\frac{1}{\mathcal{A}_W(\mathbf{r}_{p,i})\mathcal{A}_W(\mathbf{r}_{p,j})}\int\frac{dkk}{2\pi}J_2(kr_{p,i})J_2(kr_{p,j})\tilde{W}(k)\tilde{W}(k)\Sigma_c^2\left(P_{\kappa\kappa}(k)+\frac{\sigma_\gamma^2}{n_s}\right)\right\} \quad (3.57)$$

$\mathrm{Cov}(\Delta\Sigma)$ is then integrated to get the covariance in bins as described in Eq. (3.56).

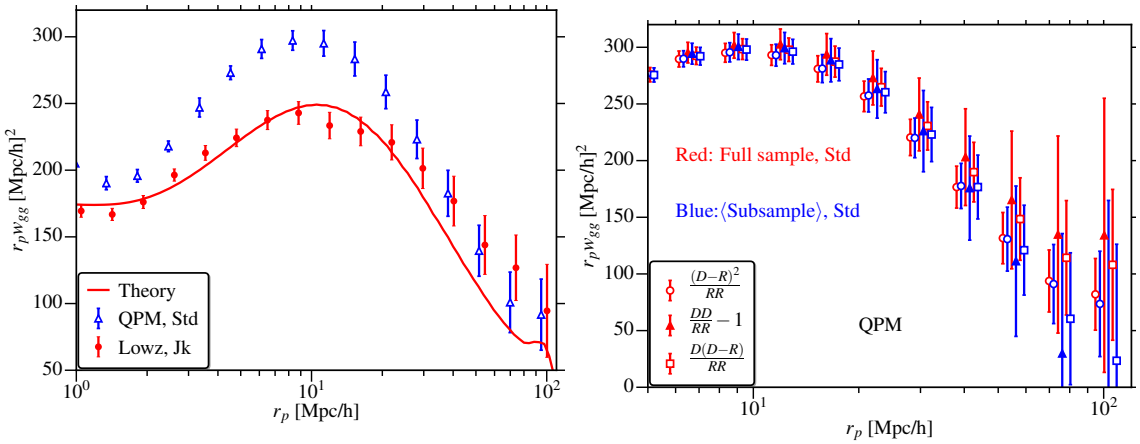
## 3.B Clustering results

In this appendix we present the comparison of different estimators and error estimation methods for the galaxy clustering measurements.

We begin by defining the standard Landy-Szalay (LS) estimator for clustering (Landy & Szalay, 1993)

$$\widehat{\xi}_{LS}(r_p,\Pi) = \frac{(D-R)^2}{RR} = \frac{DD-2DR+RR}{RR}, \quad (3.58)$$

where $\xi$ is the three dimensional correlation function, $r_p$ is the projected separation on the sky, and $\Pi$ is the line-of-sight separation between the pair of galaxies. The use of $D-R$ indicates that we are




(a) (b)

Figure 3.10: a) Comparison of the projected clustering for the LOWZ sample and the QPM mocks. The LOWZ errors are from the jackknife method, while the QPM values are the mean and standard deviation across 150 realizations. The red line is the Planck 2015 ΛCDM prediction along with the best-fitting bias from Singh et al. (2016a); note that the points on large scales have correlated errors. b) Comparison of clustering measurements using different estimators (note different $r_p$ range). ⟨Subsample⟩ refers to the mean signal across the subsamples, in each realization. We then take the mean and standard deviation of ⟨Subsample⟩ across realizations.

correlating over-density fields. The estimator can then be expanded into its standard pair counting form. $DD$ denotes summation over all galaxy-galaxy pairs within the bin, $DR$ are the cross pairs between galaxies and randoms, and $RR$ are the random-random pairs.

In addition, we define the basic estimator from pair counting

$$\widehat{\xi}_{s2}(r_p, \Pi) = \frac{DD}{RR} - 1 \tag{3.59}$$

Motivated by the galaxy-galaxy lensing estimator without subtraction of the mean galaxy density, we also define the estimator correlating an over-density field $(D - R)$ with a density field $(D)$

$$\widehat{\xi}_{s1}(r_p, \Pi) = \frac{D(D-R)}{RR} = \frac{DD - DR}{RR} \tag{3.60}$$

We want to work with projected correlation functions, analogous to galaxy-galaxy lensing. Thus we integrate $\xi_{gg}$ over the line-of-sight to obtain the projected correlation function $w_{gg}$.

$$\widehat{w}_{gg}(r_p) = \int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} \widehat{\xi}(r_p, \Pi) \mathrm{d}\Pi. \tag{3.61}$$

The choice of $\Pi_{\rm max}$ depends on two considerations: we want to choose large $\Pi_{\rm max}$ to capture the full correlation function and to mitigate the effect of redshift space distortions (Kaiser, 1987). However, in a survey of finite redshift window, the bins at large $\Pi$ are also noisier which increases the noise in the projected correlation function as well. In this work we use $\Pi_{\rm max} = 100 h^{-1}$Mpc with linear bins of size $d\Pi = 10 h^{-1}$Mpc.

In Fig. 3.10 we show the clustering measurement for the LOWZ sample as well as for one realization of the QPM mocks, with jackknife errors for both. At small scales, the clustering between the mocks and data does not agree very well, with a maximum difference of order $\sim 30\%$. This is expected, since the QPM mocks are generated using low-resolution simulations, which only resolve the large-scale

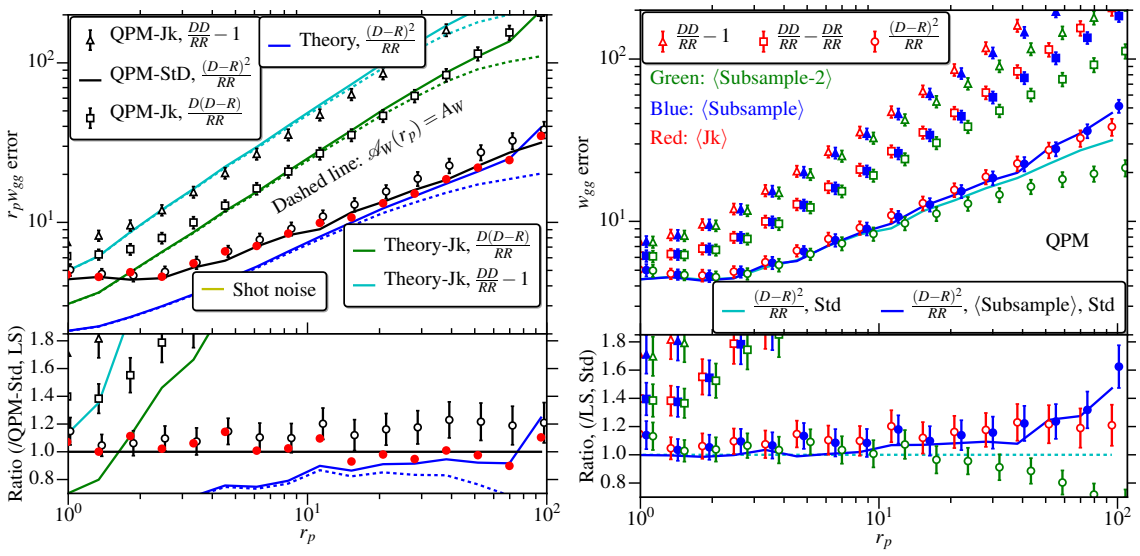



Figure 3.11: *Left panel*: Comparison of errors for different clustering estimators defined in Eqs. (3.58)–(3.60). The Landy-Szalay ($\frac{(D-R)^2}{RR}$) estimator, Eq. (3.58), gives the lowest error followed by the estimator in Eq. (3.60) and then the estimator in Eq. (3.59). Also shown are the jackknife errors for the LS estimator, with the jackknife overestimating the errors by $\sim 10-20\%$ at all scales. We additionally show the estimated errors from the theory predictions. Note that the theory estimates use the linear theory+halofit matter power spectrum, and do not include contributions from non-linear galaxy bias and connected parts of the covariance, hence the theory errors are underestimated at small scales. The difference between the solid and dashed theory lines are due to the edge effects as estimated by $\mathcal{A}_W$. *Right panel*: Comparison between the different error estimation methods in mocks. ⟨Subsample⟩ refers to the mean and standard deviation across subsamples. ⟨Subsample-2⟩ is similar to ⟨Subsample⟩, except for each subsample we also count the cross terms with other subsamples. This reduces edge effects but also leads to correlations between different subsamples at large scales.

density field (White, 2014). Here we compare the estimators and error estimations self-consistently from the QPM mocks, and thus the failure to exactly match the LOWZ sample clustering is not important.

In Fig. 3.11 we compare the clustering error estimates using different estimators. The LS estimator yields the lowest errors followed by the estimator in Eq. (3.60). The relative trends between the estimators are consistent with the theory estimates from expressions in Eq. (3.35) and Eq. (3.52) and, more generally, with the idea that each time you substitute a zero-mean field with a field that has a non-zero mean, the variance increases. We caution that for the clustering measurements, we have not completely explored the consistency between theory and empirical error estimates. Our theory estimates do not capture the full contributions from non-linear bias, redshift space distortions, and the connected part of the covariance. Also in Eq. (3.35) $\widehat{P}_{ij}(\mathbf{k}) = \widehat{P}(\mathbf{k})$ and some terms from Eq. (3.35) will be removed in the case of the estimator in Eq. (3.60); see Eq. (3.40).

In Fig. 3.11 we also compare the error estimates from jackknife and standard deviation across different mock realizations, using different estimators. Our results suggest that the jackknife overestimates the errors at all scales, even when the scales are larger than the jackknife region size. This is contrary to the expectations, since at scales larger than the jackknife region size, the assumption that the regions are independent is violated and thus the errors are expected to be underestimated. However, since the jackknife regions are much smaller than the survey size, the contribution from super-sample variance and other window function-dependent terms is expected to be larger in the jackknife. In the case of the LS estimator, our theory estimates (the difference between solid and

dashed green lines) suggest that the increase in error from edge effects is also important at large scale and can lead to the jackknife errors being over-estimated by $10 - 20\%$. For the non-optimal estimators, the increased contribution from the additional terms identified in Appendix 3.A dominates and hence the increase in jackknife errors compared to the standard deviation is substantially more, when compared to the LS estimator.

# 4 Probing gravity with a joint analysis of galaxy and CMB lensing and SDSS spectroscopy

Sukhdeep Singh[1], Shadab Alam[1,2], Rachel Mandelbaum[1], Uroš Seljak[3], Sergio Rodriguez-Torres[4], Shirley Ho[1,5]

[1] McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[2] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ , UK

**Abstract**

We present measurements of $E_G$, a probe of gravity from large-scale structure, using BOSS LOWZ and CMASS spectroscopic samples, with lensing measurements from SDSS (galaxy lensing) and Planck (CMB lensing). Using SDSS lensing and the BOSS LOWZ sample, we measure $\langle E_G \rangle = 0.405 \pm 0.047$ (stat), consistent with the predicted value, $E_G = 0.47$, from the Planck ΛCDM model at $\sim 1.5\sigma$ level. While statistically not significant, the discrepancy in measurement is largely driven by the relatively low amplitude of lensing measurement compared to clustering, which can potentially be due to systematics in the lensing map or systematics arising from the selection function of LOWZ galaxies. Using CMB lensing, we measure $\langle E_G \rangle = 0.45 \pm 0.07$ (stat) for LOWZ (statistically consistent with galaxy lensing and Planck predictions) and $\langle E_G \rangle = 0.39 \pm 0.04$ (stat) for the CMASS sample, consistent with predictions (predicted $E_G = 0.40$). We also study the redshift evolution of $E_G$ by splitting the LOWZ sample into two samples based on redshift, with results being consistent with model predictions at $2\sigma$ or better. We correct for the effects of non-linear physics and different window functions for clustering and lensing using analytical model or simulations, and demonstrate that these corrections are effective to $\sim$ 1–2%, well within our statistical uncertainties.

## 4.1 Introduction

The standard ΛCDM model of cosmology has been successful in explaining a wide array of cosmological measurements (see Weinberg et al., 2013, for a review), from the early Universe (e.g., Steigman, 2010; Komatsu et al., 2011; Planck Collaboration et al., 2015a) down to $z \lesssim 1$ (e.g., Riess et al., 1998; Perlmutter et al., 1999; Kilbinger et al., 2013; Mandelbaum et al., 2013; Betoule et al., 2014; Alam et al., 2016a), though there are some mild tensions between different probes (see for example Planck Collaboration et al., 2015a; Riess et al., 2016). General relativity (GR) lies at the core of this model but it requires additional matter and energy components (dark matter and dark energy) to explain the structure formation and cosmic acceleration. The nature of these components, especially dark energy, is not very well understood and this leaves open the possibility that the laws of gravity may require modifications as well (e.g., Jain & Khoury, 2010).

The theory of general relativity has been remarkably successful in explaining results over a wide range of scales (see Will, 2014, for a review of experimental tests of GR). On cosmological scales, it is possible to test the nature of gravity through several observables since gravity determines the dynamics and the growth of structure.

One important probe is the large scale velocity field, particularly the redshift space distortions (see Hamilton, 1997, for review). Observationally, local motion (or peculiar velocities) of galaxies introduces errors in the distances inferred using the cosmological redshift-distance relation. Due to the coherent nature of the velocities, these errors then lead to detectable distortions in the otherwise isotropic correlation function (or power spectrum in Fourier space) of galaxies. These distortions

in the redshift space correlation function depend on the strength of gravitational attraction and are parameterized through the redshift (or time) dependent growth rate factor, $f(z)$. Several surveys have detected RSD and constrained $f(z)$ at different redshifts (e.g., Beutler et al., 2012; de la Torre et al., 2013; Alam et al., 2015b, 2016a) and used it to constrain gravity (Jennings et al., 2011; Alam et al., 2016b).

Gravitational lensing is another probe of the large scale structure. Gravitational lensing results from deflections in the path of photons by the gravitational potential of intervening matter, as they travel from the source to the observer (Bartelmann & Schneider, 2001; Kilbinger, 2015). In the weak regime, lensing introduces small but coherent distortions in the shapes of galaxies. Correlations between the shapes of background galaxies can thus be used to study the gravitational potential of foreground matter. Similarly, in the case of the CMB, lensing remaps the background anisotropies, leading to cross-correlations between different modes that can be exploited to recover the matter potential (Zaldarriaga & Seljak, 1999; Hu & White, 2001; Lewis & Challinor, 2006). Cross-correlations between galaxies and the lensing maps from CMB (galaxy-CMB lensing) or background galaxies (galaxy-galaxy lensing) can be used to study the galaxy-matter cross-correlations as well as the evolution of the structure over cosmic time scales (e.g., Massey et al., 2007; Mandelbaum et al., 2013; Kilbinger et al., 2013; Heymans et al., 2013a). These measurements can then also be used to test the laws of gravity as well, since growth of structure and the lensing effect itself depends on the nature of gravity (e.g., Simpson et al., 2013).

Zhang et al. (2007) suggested the probe, $E_G$, as a consistency check on the theory of gravity by combining RSD measurements with the galaxy-lensing cross correlations (see also Leonard et al., 2015). $E_G$ is sensitive to the "gravitational slip" or the ratio of Newtonian potential and curvature potential, which are equal within GR in the absence of any anisotropic stress. $E_G$ has been measured by Reyes et al. (2010), Blake et al. (2016) and Alam et al. (2017) using galaxy-galaxy lensing and by Pullen et al. (2016) using galaxy-CMB lensing. Measurements are largely consistent with the ΛCDM predictions, though Pullen et al. (2016) measured $E_G$ to be $\sim 2.6\sigma$ lower than predictions, with most of the discrepancy coming from the low lensing amplitude at large scales.

In this work, we measure $E_G$ using the BOSS galaxy samples and SDSS galaxy lensing as well as Planck CMB lensing maps. In the case of galaxy lensing, due to limitations of the SDSS sample we only use the BOSS low redshift sample (LOWZ) while for CMB lensing we use both LOWZ and CMASS samples.

Throughout, we use the Planck 2015 cosmology (Planck Collaboration et al., 2015a), with $\Omega_m = 0.309$, $n_s = 0.967$, $A_s = 2.142 \times 10^{-9}$, $\sigma_8 = 0.82$. To get predictions for matter correlation function, we use the linear power spectrum with halofit (Smith et al., 2003; Takahashi et al., 2012), generated using the CAMB software (Lewis & Bridle, 2002b).

## 4.2 Formalism and Estimators

In this section we provide a brief review of the theoretical formalism and the estimators used in this work.

### 4.2.1 Weak Lensing

As photons travel from source to observer, their paths are deflected by the structure in the intervening matter distribution (see Bartelmann & Schneider, 2001; Kilbinger, 2015, for review). The lensing potential of a given lensing mass depends on the lens potential and the geometric factors involving distances between the lens, source and observer, and is given by

$$\Phi_L = \int \mathrm{d}\chi_l \frac{f_k(\chi_s - \chi_l)}{f_k(\chi_s) f_k(\chi_l)} \Psi(f_k(\chi_l)\theta, \chi_l) \tag{4.1}$$

$\chi_l$ and $\chi_s$ are line-of-sight distances to lens and source respectively ($\chi_s > \chi_l$), $\theta$ is the angular separation between the lens and source on the sky and $f_k(\chi)$ is the transverse comoving distance

($f_k(\chi) = \chi$ in a flat universe). The Weyl potential $\Psi$ is given by

$$\Psi = \psi + \phi \tag{4.2}$$

$\psi$ and $\phi$ are the Newtonian and curvature potentials. Within $\Lambda$CDM, $\psi = \phi$ in the absence of any anisotropic stress. The main focus of this paper is to test this equality of the two potentials through the measurement of the $E_G$ parameter as defined in Section 4.2.5.

In the case where the angular extent of source is much smaller than angular scales over which lens potential varies, the distortion matrix is given by

$$A = \begin{bmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{bmatrix}.$$

where $\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{i2\phi}$ is the shear in the observer frame and can be rotated to the lens-source frame to give $\gamma = \gamma_t + i\gamma_\times$. For circularly symmetric lens, the convergence $\kappa$ and the tangential shear $\gamma_t$ are given by ( $\gamma_\times = 0$ from symmetry)

$$\kappa(r_p) = \frac{\Sigma(r_p)}{\Sigma_c} \tag{4.3}$$

$$\gamma_t(r_p) = \frac{\Delta\Sigma}{\Sigma_c} = \frac{\bar{\Sigma}(< r_p) - \Sigma(r_p)}{\Sigma_c}. \tag{4.4}$$

$\Sigma$ is the projected surface mass density and $\bar{\Sigma}(< r_p)$ is the mean $\Sigma$ within radius $r_p$ from lens center. For non-circularly symmetric lens distributions, the equation is true when averaging within annuli at fixed $r_p$. The geometric factor $\Sigma_c$ is given by

$$\Sigma_c = \frac{c^2}{4\pi G}\frac{f_k(\chi_s)}{(1+z_l)f_k(\chi_l)f_k(\chi_s-\chi_l)}. \tag{4.5}$$

$1+z_l$ converts the $c^2/G$ factor to comoving space. The projected surface mass density can be written in terms of the 2-point galaxy-matter cross correlation function (lensing is sensitive to density fluctuations not the mean density)

$$\Sigma(r_p) = \bar{\rho}_m \int \mathrm{d}\Pi\, \xi_{gm}(r_p, \Pi) = \bar{\rho}_m w_{gm}(r_p), \tag{4.6}$$

$\bar{\rho}_m$ is the mean matter density in comoving coordinates. Under the assumption of large projection length $\Pi$, the projected galaxy-matter cross correlation function $w_{gm}$ can be derived from matter power spectrum as

$$w_{gm}(r_p) = b_g A_l r_{\rm cc} \int \mathrm{d}z\, W(z) \int \frac{\mathrm{d}^2k}{(2\pi)^2} P_{\delta\delta}(\mathbf{k}, z) e^{i(\mathbf{r}_p\cdot\mathbf{k})} \tag{4.7}$$

$W(z)$ is the lens weight function, that we compute using the weights defined in Sections 4.2.1 and 4.2.1. $b_g$ is the galaxy bias, $r_{cc}$ is the galaxy-matter cross correlation coefficient and $A_l$ is the scale independent lensing amplitude. Since $r_{cc} \sim 1$ at the scales we use to fit lensing measurements, we fix $r_{cc} = 1$ when computing the model and encapsulate the deviations from the model in a scale-independent lensing amplitude, $A_l$. When fitting the correlation function, we also assume a scale independent bias, $b_g$. We do apply correction for scale dependent bias and $r_{cc}$ when computing $E_G$(see section 4.2.6).

**Estimator: Galaxy-galaxy lensing**

We measure the $\Delta\Sigma$ using tangential shear as

$$\widehat{\Delta\Sigma}(r_p) = \frac{\sum_{ls} w_{ls}\gamma_t^{(ls)}\Sigma_c^{(ls)}}{\sum_{rs} w_{rs}} - \frac{\sum_{rs} w_{rs}\gamma_t^{(rs)}\Sigma_c^{(rs)}}{\sum_{rs} w_{rs}} \tag{4.8}$$

where the summations are over all the lens-source (ls) or random lens-source (rs) pairs. The signal measured around randoms is subtracted to remove the spurious signal from additive systematics (Mandelbaum et al., 2005) and also to reduce the noise (Singh et al., 2016b). Note that the normalization factor uses weights computed using randoms. This is to account for the source galaxies which are associated with the lens and are not lensed, but enter the estimator due to the scatter in photometric redshifts. The lens-source pair weights are given by

$$w_{ls} = \frac{\Sigma_c^{-2}}{\sigma_\gamma^2 + \sigma_{SN}^2}. \tag{4.9}$$

$\Sigma_c^{-2}$ enters because we have defined $\widehat{\Delta\Sigma}$ as the maximum likelihood estimator (Sheldon et al., 2004); $\sigma_{SN}$ is the shape noise and $\sigma_\gamma$ is the measurement noise.

**Estimator: Galaxy-CMB lensing**

Using CMB lensing, we measure the projected surface mass density as (Singh et al., 2016a)

$$\widehat{\Sigma}(r_p) = \frac{\sum_{lp} w_{lp}\kappa_p\Sigma_{c,*}}{\sum_{lp} w_{lp}} - \frac{\sum_{Rp} w_{Rp}\kappa_p\Sigma_{c,*}}{\sum_{Rp} w_{Rp}} \tag{4.10}$$

where the summation is over all the lens-pixel (pixels of CMB convergence map) pairs at separations $r_p \in [r_{p,min}, r_{p,max}]$ at the lens redshift and the signal around randoms is subtracted to remove the effects of the correlated convergence, in the measurement (Singh et al., 2016a). As demonstrated in Singh et al. (2016b) using non-zero mean field (normalized galaxy field, $(1+\delta_g)$) contributes additional mean and window function dependent terms (mean multiplied with 2-point convergence correlation) to the covariance. Randoms subtraction removes the mean and hence leads to better covariance properties.

The weight for lens-pixel pair is given by

$$w_{lp} = \Sigma_{c,*}^{-2}. \tag{4.11}$$

We have defined $\widehat{\Sigma}$ as the maximum likelihood estimator, similar to galaxy-galaxy lensing and $\Sigma_{c,*}$ is the $\Sigma_c$ with CMB as the source.

### 4.2.2 Projected galaxy clustering

The two point galaxy correlation function in redshift space can be written as

$$\xi_{gg}(r_p, \Pi, z) = b_g^2 \int \frac{\mathrm{d}^2k_\perp \mathrm{d}k_z}{(2\pi)^3} P_{\delta\delta}(\mathbf{k}, z) \left(1 + \beta\mu^2\right)^2 e^{i(\mathbf{k}_\perp \cdot \mathbf{r}_p + k_z\Pi)} \tag{4.12}$$

where $P_{\delta\delta}$ is the matter power spectrum. $b_g$ is the linear galaxy bias and the $(1+\beta\mu^2)$ factor accounts for the linear redshift space distortions (Kaiser, 1987, see also Section 4.2.3), where $\beta = f(z)/b_g$ and $f(z) \approx \Omega_m(z)^{0.55}$ is the linear growth rate. The projected correlation function is then

$$w_{gg}(r_p) = \int \mathrm{d}z\, W(z) \int_{-\Pi_{\max}}^{\Pi_{\max}} \mathrm{d}\Pi\, \xi_{gg}(r_p, \Pi, z) \tag{4.13}$$

$\Pi_{\max}$ is the maximum line-of-sight separation for which the correlation function is computed. The weight function accounts for the differential contribution from different redshifts to the correlation function. It depends on total volume as well as redshift distribution of the galaxies, $p(z)$ and is given by (Mandelbaum et al., 2011)

$$W(z) = \frac{p(z)^2}{\chi^2(z)\mathrm{d}\chi/\mathrm{d}z} \left[\int \frac{p(z)^2}{\chi^2(z)\mathrm{d}\chi/\mathrm{d}z}\mathrm{d}z\right]^{-1} \tag{4.14}$$

Finally the projected correlation function can be written as

$$w_{gg}(r_p) = \frac{b^2}{\pi^2} \int \mathrm{d}z W(z) \int_0^\infty \mathrm{d}k_z \int_0^\infty \mathrm{d}k_\perp \frac{k_\perp}{k_z} P(\mathbf{k}, z) \\ \sin(k_z \Pi_{\max}) J_0(k_\perp r_p) \left(1 + \beta\mu^2\right)^2 \tag{4.15}$$

**Estimators**

To measure the galaxy clustering, we use the generalized Landy-Szalay estimator (Landy & Szalay, 1993)

$$\xi_{gg} = \frac{(D-R)(D-R)}{RR} = \frac{DD - 2DR + RR}{RR}, \tag{4.16}$$

where $D$ is the galaxy sample and $R$ is for randoms sample.

The projected correlation function can then be obtained by integrating the correlation function along the line-of-sight

$$w_{gg} = \int_{-\Pi_{\max}}^{\Pi_{\max}} \xi_{gg}(r_p, \Pi)\, \mathrm{d}\Pi. \tag{4.17}$$

$\Pi_{\max}$ denotes the size of top hat window function along the line-of-sight. It is desirable to use larger $\Pi_{\max}$ to mitigate the effects of redshift space distortions in the projected correlation function, though very large line-of-sight ($\Pi$) values add little signal and mostly contribute noise (for a survey of finite window size in redshift). We use $\Pi_{\max} = 100h^{-1}$Mpc, though we also test our results using $\Pi_{\max} = 50h^{-1}$Mpc and $\Pi_{\max} = 200h^{-1}$Mpc, finding consistent results.

To get the errors and covariance matrix, we split the sample into 100 jackknife regions with approximately equal area on the sky (Singh et al. (2016b) show that the jackknife errors are consistent with theoretical expectations when using mean zero quantities, i.e., galaxy field with mean subtracted using randoms). Throughout, the quoted errors and uncertainties will be jackknife errors. Whenever fitting models (or computing derived quantities such as a mean), we do so for each jackknife region separately and then quote the jackknife mean and errors of model parameters.

### 4.2.3 Redshift space distortions

The measured redshift ($z$) of the galaxy is a combination of Hubble recession and the peculiar velocity[1]. The peculiar velocity component of the redshift affects the line-of-sight distance to a galaxy determined using the cosmological distance-redshift relation, introducing an anisotropy in the two point correlation function. Within linear theory, the real space power spectra can be converted to redshift space power spectra as (Kaiser, 1987)

$$P^s_{gg}(\mathbf{k}) = P^r_{gg}(\mathbf{k})(1 + \mu_k^2\beta)^2. \tag{4.18}$$

where $P^r_{gg}$ is the real space galaxy power spectrum, $P^s_{gg}$ is the redshift space galaxy power spectrum, $\mu$ is the cosine of the angle made by the wave-vector $\mathbf{k}$ from the line of sight and $\beta$ is the RSD parameter. Hamilton (1992) extended this approach to the two-point correlation function and showed that the two-point correlation in redshift space up to linear theory is given by

$$\xi_{gg}(\mathbf{s}) = \left[1 + f(\partial/\partial z)^2(\nabla^2)^{-1}\right]^2 \xi_{gg}(\mathbf{r}) \tag{4.19}$$

where $\xi_{gg}(\mathbf{r})$ and $\xi_{gg}(\mathbf{s})$ are the galaxy correlation function in real and redshift space. The development in Lagrangian perturbation theory has produced better models of redshift space correlation which work even in the quasi-linear scale. They enable interpretation of RSD measurements to smaller separations, and hence result in more precise measurements of RSD (Samushia et al., 2014). We use Convolution Lagrangian Perturbation Theory (CLPT) to predict the real space correlation function and velocity

[1]There is also a small contribution from gravitational redshift (Cappi, 1995) , which we will ignore.

moments, which are then combined with the Gaussian Streaming Model (GSM) to predict the redshift-space correlation function (Carlson et al., 2013; Wang et al., 2014). We used tools developed in (Alam et al., 2015b) which have been tested on several mocks and N-body simulations including the completed cosmological analysis of DR12 from BOSS (Satpathy et al., 2016; Alam et al., 2016a).

**Measuring the growth rate**

We estimate $\beta$ using the monopole and quadruple moments of the galaxy auto-correlation function, obtained by projecting the correlation function onto the Legendre basis.

$$\xi_{gg,l}(s) = \int_{-1}^{1} \mathrm{d}\mu\, \mathcal{P}_l(\mu)\xi_{gg}(s,\mu) \tag{4.20}$$

where $s$ is the distance in redshift space, $\mu$ is the angle between $\mathbf{s}$ and the plane of the sky and $\mathcal{P}_l$ is the legendre polynomial of order $l$ ($l = 0$ for monopole and $l = 2$ for quadrupole)

We use COSMOMC (Lewis & Bridle, 2002a) to run a Markov Chain Monte Carlo (MCMC) fit for the multipole moments of the correlation function. The correlation function is fit for 3 parameters $(f, \sigma_{FOG}, \nu)$, where $f$ is the growth rate, $\sigma_{FOG}$ is an additional velocity dispersion to model the Finger of God effect (Alam et al., 2015b) and $\nu$ is the overdensity which determines the first and second order bias through the peak-background split (White, 2014):

$$F' = \frac{1}{\delta_c}\left[a\nu^2 - 1 + \frac{2p}{1+(a\nu^2)^p}\right] \tag{4.21}$$

$$F'' = \frac{1}{\delta_c^2}\left[a^2\nu^4 - 3a\nu^2 + \frac{2p(2a\nu^2+2p-1)}{1+(a\nu^2)^p}\right] \tag{4.22}$$

where a = 0.707, p = 0.3 gives the Sheth-Tormen mass function (Sheth & Tormen, 1999), and $\delta_c = 1.686$ is the critical density for collapse. $F'$ and $F''$ are the first and second order Lagrangian bias. The linear galaxy bias is given by $b_g = 1 + F'$. The scale used in the RSD fits cover 28 to 70 $h^{-1}$Mpc. The smaller scale is chosen to only use the scales where model is shown to be consistent with survey mocks and N-body simulations (Alam et al., 2015b) whereas the large scale is determined by the size of the jackknife regions.

### 4.2.4 $\Upsilon$ Estimator

In galaxy-galaxy lensing, we measure $\Delta\Sigma = \bar{\Sigma}(< r_p) - \Sigma(r_p)$. $\bar{\Sigma}(< r_p)$ and hence $\Delta\Sigma$ is affected by all scales $< r_p$. This induces considerable theoretical uncertainty in modeling the $\Delta\Sigma$ observable due to the lack of a good model for the small scales, for example, non-linear clustering and galaxy-dark matter cross-correlations (usually quantified through $r_{\rm cc}$ parameter) at scales comparable to or within the virial radii of dark matter halos. To mitigate the impact of small-scale theoretical uncertainty on large-scale structure measurements, Baldauf et al. (2010) suggested a new estimator

$$\Upsilon_{gm}(r_p, r_0) = \Delta\Sigma_{gm}(r_p) - \frac{r_0^2}{r_p^2}\Delta\Sigma(r_0) \tag{4.23}$$

Expanding $\Delta\Sigma$ in terms of the correlation function, it can be shown that $\Upsilon$ is independent of information from small scales, $r < r_0$:

$$\Upsilon_{gm}(r_p, r_0) = \bar{\rho}_M\left[\frac{2}{r_p^2}\int_{r_0}^{r_p} \mathrm{d}r'\, w_{gm}(r') - w_{gm}(r_p) + \frac{r_0^2}{r_p^2}w_{gm}(r_0)\right]. \tag{4.24}$$

Analogous to this lensing observable, we can define $\Upsilon_{gg}$, starting by defining

$$\Sigma_{gg}(r_p) = \rho_{\rm crit}\int \mathrm{d}\Pi\, \xi_{gg}(r_p, \Pi), \tag{4.25}$$

where we have ignored the effects of the mean density and $\rho_{\rm crit}$ is used to get $\Sigma_{gg}$ in units of density. The choice of $\rho_{\rm crit}$ is also useful for the definition of the $E_G$ parameter in Section 4.2.5. The definition of $\Upsilon_{gg}$ follows from definition of $\Delta\Sigma$ and eq. (4.23).

### 4.2.5 $E_G$ Statistic

The perturbed metric in the conformal Newtonian gauge is written as

$$ds^2 = a(\tau)^2\{(1-2\psi)d\tau^2 + (1-2\phi)(dr^2 + r^2 d\Omega^2)\} \tag{4.26}$$

As stated in section 4.2.1, $\psi = \phi$ in $\Lambda$CDM, in the absence of anisotropic stresses. However, several modified gravity theories lead to modifications in either or both the potentials and in general violate the equality of two potentials (Jain & Khoury, 2010).

Zhang et al. (2007) proposed an estimator, $E_G$, to test the equality of two metric potentials

$$E_G(k,z) = \frac{P_{g\kappa}(k,z)}{P_{g\theta}(k,z)} \tag{4.27}$$

The real-space analog of $E_G$ was defined by Reyes et al. (2010) (see also Leonard et al. 2015) as

$$E_G(r_p) = \frac{\Upsilon_{gm}(r_p)}{\beta\Upsilon_{gg}(r_p)} \tag{4.28}$$

Equivalently, we can define $E_G$ in terms of projected surface mass density

$$E_G(r_p) = \frac{\Sigma_{gm}(r_p)}{\beta\Sigma_{gg}(r_p)} \tag{4.29}$$

In the case of galaxy-galaxy lensing, we will measure $E_G$ using $\Upsilon$ estimators, while for CMB lensing we will present results using both $\Upsilon$ and $\Sigma$. $\Sigma$ provides better measurement at small scales as there is no mixing of scales or removal of information as in $\Delta\Sigma$ or $\Upsilon$. However, $\Sigma$ has lower signal to noise (S/N) at large scales due to the low amplitude and higher contributions from cosmic variance, which introduces noise bias in $E_G$, in addition to the measurement being noisy. $\Upsilon$ on the other hand has lower S/N at small scales but is better at large scales due to higher S/N.

$E_G$ is sensitive to the difference between $\psi, \phi$ as $P_{g\kappa}$ or $\Upsilon_{gm}$, measured from lensing, are sensitive to the Weyl potential ($\Psi = \psi + \phi$), while the linear growth rate of matter perturbations depends on the Newtonian potential $\psi$. In $\Lambda$CDM, with $\psi = \phi$

$$\langle E_G(r_p, z)\rangle = \frac{\Omega_m(z=0)}{f(z)} \tag{4.30}$$

In this paper we focus on measurements of $E_G$ and their implications for the $\Lambda$CDM model. We do not test for any specific models of modified gravity, and refer the reader to Pullen et al. (2015b) for expectation values of $E_G$ in some alternate theories of gravity.

### 4.2.6 Possible Systematics

Theoretically, in the linear regime, $E_G$ is a clean null test for the equality of the two metric potentials and hence a test of GR. However, it can be subject to several observational and theoretical systematics that we discuss here.

*Non-linear clustering* at small scales can lead to significant deviations from the expected value of $E_G$, especially since it affects lensing and clustering measurements differently. The clustering ($\Upsilon_{gg}$) amplitude goes as $b_g^2$, which includes difficult-to-model non-linear galaxy bias. Lensing ($\Upsilon_{gm}$) is sensitive to $b_g r_{cc}$, where $b_g$ includes non-linear bias while $r_{\rm cc}$ is the galaxy-matter cross correlation coefficient. $r_{\rm cc} \sim 1$ at large scales but deviates strongly at small scales and depends on details of how

galaxies populate halos (see e.g. Baldauf et al., 2010, for more detailed discussion on $r_{cc}$). As discussed in Baldauf et al. (2010), use of $\Upsilon$ rather than $\Delta\Sigma$ partially mitigates the effects of non-linear clustering by removing small-scale information, though the efficiency of this small-scale removal depends on the choice of $r_0$. The choice of $r_0$ depends on the balance between theoretical uncertainties, which are large at small scales, and statistical errors, which increase with $r_0$ as more signal is removed. For our main analysis, we choose $r_0 = 2h^{-1}\mathrm{Mpc} \gtrsim 2r_{\rm vir}$, as suggested by Baldauf et al. (2010). This still leaves some residual effects from non-linear scales and we attempt to mitigate these effects by estimating them using simulations. For this we compute the correction factor $C_{nl}$ to correct for the effects of scale dependent bias and $r_{cc}$.

$$C_{nl}(r_p) = \frac{\Upsilon^{\rm sim}_{gg}(r_p)}{b_{\rm linear}\Upsilon^{\rm sim}_{gm}(r_p)} \tag{4.31}$$

$b_{\rm linear}$ is the linear galaxy bias.

The finite size of the top-hat window function in clustering measurement leaves some residual effects of *linear redshift space distortions* in the projected correlation function (see also Section 4.2.2) and hence in the $E_G$ measurement. The lensing measurement is not affected by RSD to first order and also the line-of-sight integration in lensing is much longer ($\Pi^{\rm lens}_{\rm max} \gg 100h^{-1}\mathrm{Mpc}$). We mitigate the effect of the limited window in projected clustering using the corrections computed with the linear theory and Kaiser formula (Kaiser, 1987).

$$C_{\rm rsd+win}(r_p) = \frac{w_{gg}(r_p|\Pi_{\rm max} = 100h^{-1}\mathrm{Mpc})}{w_{gg}(r_p|\Pi_{\rm max} = \infty)} \tag{4.32}$$

where $w_{gg}$ is defined in Eq. (4.15).

In addition to the above, the window function for lensing (depends on the broad lensing kernel and systematics weights) also varies, due to which effective weights to different $\Pi$ bins are different for lensing and clustering. We correct for this effect using simulations by computing $\Upsilon_{gm}$ with top hat and lensing weights which also includes the systematics weights defined in Section 4.2.1.

$$C_{\rm lens\ win}(r_p) = \frac{\int d\Pi\, \xi_{gm}(r_p, \Pi)}{\int d\Pi\, \mathcal{W}_{gm}(\Pi)\, \xi_{gm}(r_p, \Pi)} \tag{4.33}$$

Here $\mathcal{W}_{gm}$ is the lensing window function derived in Appendix 4.A.

Due to different redshift weights for lensing and clustering, the two measurements are also at different effective redshifts. Since the lensing amplitude scales as $b_g D(z)^2$ ($D(z)$ is the linear growth function) and clustering amplitude scales as $b_g^2 D(z)^2$, we correct for the different effective redshifts as

$$C_z(r_p) = \frac{D(z = z^{gg}_{\rm eff})}{D(z = z^{gm}_{\rm eff})} \sqrt{\frac{\int dz \mathcal{W}_{gg}(z) w_{gg}(z)}{\int dz \mathcal{W}_{gm}(z) w_{gg}(z)}}. \tag{4.34}$$

The ratio is directly computed from the data by measuring clustering with lensing weights ($\mathcal{W}_{gm}$) assigned to the galaxies.

There are also additional number density fluctuations due to the lensing effects of the intervening large scale structure between the observer and galaxies (Moradinezhad Dizgah & Durrer, 2016). We estimate the impact of these effects in appendix 4.C and apply correction for them $C_{\rm lens}$.

The final correction applied to the measured $E_G$ signal is the product of all corrections defined above.

$$C_{\rm tot}(r_p) = C_{\rm lens\ win}(r_p) \times C_{nl}(r_p) \times C_{\rm rsd+win}(r_p) \times C_z(r_p) \times C_{\rm lens} \tag{4.35}$$

## 4.3 Data

We use the same datasets as used in Singh et al. (2016a). In this section, we briefly describe these datasets for completeness.

### 4.3.1 SDSS

The SDSS (York et al., 2000) imaged roughly $\pi$ steradians of the sky, and the SDSS-I and II surveys followed up approximately one million of the detected objects spectroscopically (Eisenstein et al., 2001; Richards et al., 2002; Strauss et al., 2002). The imaging was carried out by drift-scanning the sky in photometric conditions (Hogg et al., 2001; Ivezić et al., 2004), in five bands ($ugriz$) (Fukugita et al., 1996; Smith et al., 2002) using a specially-designed wide-field camera (Gunn et al., 1998) on the SDSS Telescope (Gunn et al., 2006). These imaging data were used to create the catalogues of shear estimates that we use in this paper. All of the data were processed by completely automated pipelines that detect and measure photometric properties of objects, and astrometrically calibrate the data (Lupton et al., 2001a; Pier et al., 2003; Tucker et al., 2006). The SDSS-I/II imaging surveys were completed with a seventh data release (Abazajian et al., 2009), though this work will rely as well on an improved data reduction pipeline that was part of the eighth data release, from SDSS-III (Aihara et al., 2011); and an improved photometric calibration ('ubercalibration', Padmanabhan et al., 2008).

### 4.3.2 SDSS-III BOSS

Based on the photometric catalog, galaxies are selected for spectroscopic observation (Dawson et al., 2013), and the BOSS spectroscopic survey was performed (Ahn et al., 2012) using the BOSS spectrographs (Smee et al., 2013). Targets are assigned to tiles of diameter 3° using an adaptive tiling algorithm (Blanton et al., 2003), and the data were processed by an automated spectral classification, redshift determination, and parameter measurement pipeline (Bolton et al., 2012). We use SDSS-III BOSS data release 12 (Alam et al., 2015a; Reid et al., 2016) LOWZ galaxies, in the redshift range $0.16 < z < 0.36$ and CMASS galaxies in range $0.45 < z < 0.7$.

The LOWZ sample consists of Luminous Red Galaxies (LRGs) at $z < 0.4$, selected from the SDSS DR8 imaging data and observed spectroscopically in the BOSS survey. The sample is approximately volume-limited in the redshift range $0.16 < z < 0.36$, with a number density of $\bar{n} \sim 3 \times 10^{-4}\ h^3\mathrm{Mpc}^{-3}$ (Manera et al., 2015). To test for the redshift evolution of $E_G$, we also split the sample into two redshift bins, Z1: $0.16 < z < 0.26$ and Z2: $0.26 < z < 0.36$. Further, we also use a sample of field galaxies (Singh et al., 2016a) defined using counts-in-cylinders (CiC) technique (Reid & Spergel, 2009). Field galaxies are mainly central galaxies and have somewhat lower bias and hence are relatively less affected by non-linear effects.

The BOSS CMASS sample consists of higher redshift galaxies ($0.45 < z < 0.7$) targeted using color and magnitude cuts intended to select a uniform sample of massive galaxies (Reid et al., 2016).

We also use the weights defined by BOSS collaboration (Ross et al., 2012) for systematics ($w_{\rm sys}$), redshift failures ($w_{\rm no\text{-}z}$) and fiber collisions ($w_{\rm cp}$). The effects of these weights for both clustering and lensing measurements were studied in detail in Singh et al. (2016a) (see also Ross et al., 2012; Anderson et al., 2014). The final weights are defined as

$$w = w_{\rm sys}(w_{\rm no\text{-}z} + w_{\rm cp} - 1) \tag{4.36}$$

### 4.3.3 Shape sample

The catalogue of galaxies with measured shapes used in this paper (described in Reyes et al. 2012 and further characterized in Mandelbaum et al. 2013) was generated using the re-Gaussianization method (Hirata & Seljak, 2003) of correcting for the effects of the point-spread function (PSF) on the observed galaxy shapes. The catalogue production procedure was described in detail in previous work, so we describe it only briefly here. Galaxies were selected in a 9243 deg$^2$ region, with an average number density of 1.2 arcmin$^{-2}$. The selection was based on cuts on the imaging quality, data reduction quality, galactic extinction $A_r < 0.2$ defined using the dust maps from Schlegel et al. (1998) and the extinction-to-reddening ratios from Stoughton et al. (2002), apparent magnitude (extinction-corrected $r < 21.8$), and galaxy size compared to the PSF. The apparent magnitude cut used model magnitudes[2].

[2] http://www.sdss3.org/dr8/algorithms/magnitudes.php#mag_model

For comparing the galaxy size to that of the PSF, we use the resolution factor $R_2$ which is defined using the trace of the moment matrix of the PSF $T_{\rm P}$ and of the observed (PSF-convolved) galaxy image $T_{\rm I}$ as

$$R_2 = 1 - \frac{T_{\rm P}}{T_{\rm I}}. \tag{4.37}$$

We require $R_2 > 1/3$ in both $r$ and $i$ bands.

The software pipeline used to create this catalogue obtains galaxy images in the $r$ and $i$ filters from the SDSS 'atlas images' (Stoughton et al., 2002). The basic principle of shear measurement using these images is to fit a Gaussian profile with elliptical isophotes to the image, and define the components of the ellipticity

$$(e_+, e_\times) = \frac{1-(b/a)^2}{1+(b/a)^2}(\cos 2\phi, \sin 2\phi), \tag{4.38}$$

where $b/a$ is the axis ratio and $\phi$ is the position angle of the major axis. The ellipticity is then an estimator for the shear,

$$(\gamma_+, \gamma_\times) = \frac{1}{2\mathcal{R}}\langle(e_+, e_\times)\rangle, \tag{4.39}$$

where $\mathcal{R} \approx 0.87$ is called the 'shear responsivity' and represents the response of the distortion to a small shear (Kaiser et al., 1995; Bernstein & Jarvis, 2002); $\mathcal{R} \approx 1 - e^2_{\rm rms}$. In the course of the re-Gaussianization PSF-correction method, corrections are applied to account for non-Gaussianity of both the PSF and the galaxy surface brightness profiles (Hirata & Seljak, 2003).

For this work, we do not use the entire source catalogue, only the portion overlapping the LOWZ sample.

When computing the weak lensing signals around the LOWZ galaxies, we need estimates of the redshifts for the fainter source galaxies. For this purpose, we use the maximum-likelihood estimates of photometric redshifts (photo-$z$) based on the five-band photometry from the Zurich Extragalactic Bayesian Redshift Analyzer (ZEBRA, Feldmann et al., 2006), which were characterized by Nakajima et al. (2012) and Reyes et al. (2012). In this work, we used a fair calibration sample of source galaxies with spectroscopic redshifts as defined in Nakajima et al. (2012) to calculate biases in weak lensing signals due to bias and scatter in the photo-$z$, and applied corrections that were of order 10 per cent ($\pm 2$ per cent) to the weak lensing signals. We will further test the accuracy of these corrections using the clustering-z method (e.g. Ménard et al., 2013) in appendix 4.E.

### 4.3.4 Planck CMB lensing maps

For CMB lensing, we use the publicly available lensing maps from Planck collaboration (Planck Collaboration et al., 2015b). We convert the provided convergence values in Fourier space $\kappa_{\ell,m}$ into real space $\kappa_{\theta,\phi}$ using HEALPY (Górski et al., 2005) with $n_{\rm side} = 1024$. We use the full $\ell$ range ($8 < \ell < 2048$) when constructing the map. Planck Collaboration et al. (2015b) found some evidence of systematics in lensing maps and use $40 < \ell < 400$ for the main cosmological analysis. However, in Singh et al. (2016a) no evidence of systematics was found when cross-correlating the full convergence map with the BOSS galaxies and hence we will use the full $\ell$ range in this work as well. We refer the reader to Singh et al. (2016a) for more details about choice of pixel size in the maps and various systematic tests that were carried out.

### 4.3.5 Simulations

To test our $E_G$ pipeline and compute the corrections described in section 4.2.6, we use the 'Med-Res' simulations described in Reid et al. (2014), with snapshots at $z = 0.25$ and $z = 0.40$ for the LOWZ sample and $z = 0.60$ for the CMASS sample. The sample of galaxies is generated using the HOD model from Zheng et al. (2005) to fit the clustering of galaxies for full LOWZ or CMASS samples, assuming a fixed abundance of halos (see Reid et al., 2014, for more details). We will use same

| Sample | Lensing map | $f\frac{\sigma_8}{\sigma_{8,fid}}$ | $b_g\frac{\sigma_8}{\sigma_{8,fid}}$ | $A_l\frac{\sigma_8}{\sigma_{8,fid}}$ | $E_G$ |
|---|---|---|---|---|---|
| Z1 | SDSS | $0.74 \pm 0.11$ [0.64] | $1.73 \pm 0.03$ | $0.87 \pm 0.07$ | $0.38 \pm 0.06$ [0.49] |
| Z1 | Planck | $0.74 \pm 0.11$ [0.64] | $1.74 \pm 0.03$ | $1.4 \pm 0.23$ | $0.61 \pm 0.13$ [0.49] |
| LOWZ | SDSS | $0.63 \pm 0.06$ [0.67] | $1.85 \pm 0.02$ | $0.76 \pm 0.05$ | $0.405 \pm 0.047$ [0.47] |
| LOWZ | Planck | $0.63 \pm 0.06$ [0.67] | $1.86 \pm 0.02$ | $0.92 \pm 0.14$ | $0.46 \pm 0.07$ [0.47] |
| Field | SDSS | - | $1.52 \pm 0.02$ | $0.76 \pm 0.05$ | $0.40 \pm 0.04$ [0.47] |
| Field | Planck | - | $1.52 \pm 0.02$ | $0.98 \pm 0.15$ | $0.48 \pm 0.08$ [0.47] |
| Z2 | SDSS | $0.54 \pm 0.07$ [0.69] | $1.96 \pm 0.04$ | $0.69 \pm 0.08$ | $0.44 \pm 0.07$ [0.45] |
| Z2 | Planck | $0.54 \pm 0.07$ [0.69] | $1.96 \pm 0.03$ | $0.64 \pm 0.16$ | $0.36 \pm 0.1$ [0.45] |
| CMASS | Planck | $0.7 \pm 0.04$ [0.78] | $2.05 \pm 0.01$ | $0.87 \pm 0.08$ | $0.39 \pm 0.04$ [0.4] |

Table 4.1: Values of the growth function, $f$, linear galaxy bias, $b_g$ (from fits to projected correlation function $w_{gg}$), lensing amplitude, $A_l$ and $E_G$ for different samples and using SDSS and Planck CMB lensing maps. The values in square brackets, [], are the predicted values from the Planck ΛCDM model. The errorbars in this table are statistical only. There is additional systematic uncertainty in $E_G$ of order 2% from the applied corrections. Note that for the Field sample, we use $f$ measured from the full Lowz sample. Due to the selection effects from CiC method, $f$ from field sample is somewhat biased ($f = 0.74 \pm 0.06$).

simulation catalogs to compute corrections for the subsamples as well. To compute the galaxy-matter cross-correlations, we cross-correlate the galaxies with the randomly subsampled matter particles. Comparison of galaxy-galaxy and galaxy-matter correlations are shown in appendix 4.B.

## 4.4 Results

In this section we presents the results, starting with growth rate measurements in data, then $E_G$ measurements in simulations and tests for systematics. Finally we present the $E_G$ measurements in data. The clustering and lensing measurements in data (from Singh et al., 2016a) are presented in appendix 4.B.

### 4.4.1 Growth rate measurement

We performed the RSD analysis as described in Section 4.2.3. For each of the subsamples used in this paper, we divide the survey area into 100 jackknife regions. We then obtain galaxy clustering monopole and quadruple moments for each of the jackknife realizations for each sample. These monopole and quadruple moments are fit independently using our perturbation theory model. Our measurements and model fits are shown in figure 4.1 and the values of parameters are presented in table 4.1. Note that we only fit monopole and quadrupole in the range $28 < r_p < 70h^{-1}$Mpc, though we also show the hexadecapole measurements in figure 4.1, which are not very constraining and were ignored in the analysis. The growth rate measurements are consistent with ΛCDM within $2\sigma$. Also, our RSD analysis is performed with a fixed cosmology because the constraints on cosmological parameters from Planck are so tight that using Planck priors is equivalent to fixing the cosmology. We do not marginalize over the Alcock-Paczynski parameter and hence obtain smaller error than when doing the full shape RSD analysis (see for example Alam et al. 2015b). Sánchez et al. (2014) reported $f\sigma_8(z = 0.32) = 0.48 \pm 0.10$ and $f\sigma_8(z = 0.57) = 0.42 \pm 0.045$ using the SDSS DR11 sample, which is consistent with our LOWZ ($f\sigma_8(z = 0.32) = 0.44$)) and CMASS ($f\sigma_8(z = 0.57) = 0.43$). Gil-Marín et al. (2016b) also measured $f\sigma_8$ consistent with ours for both LOWZ and CMASS sample. Our LOWZ measurement of $f\sigma_8$ is $0.5\sigma$ lower than Gil-Marín et al. (2016b) with fixed Alcock-Paczynski parameter but higher than what they measured after marginalizing over Alcock-Paczynski parameters. The lower $f\sigma_8$ after AP marginalization is dominated by the information in the position of BAO peak. Since we do not use BAO scale in our RSD analysis and hence we do not allow extra freedom on top

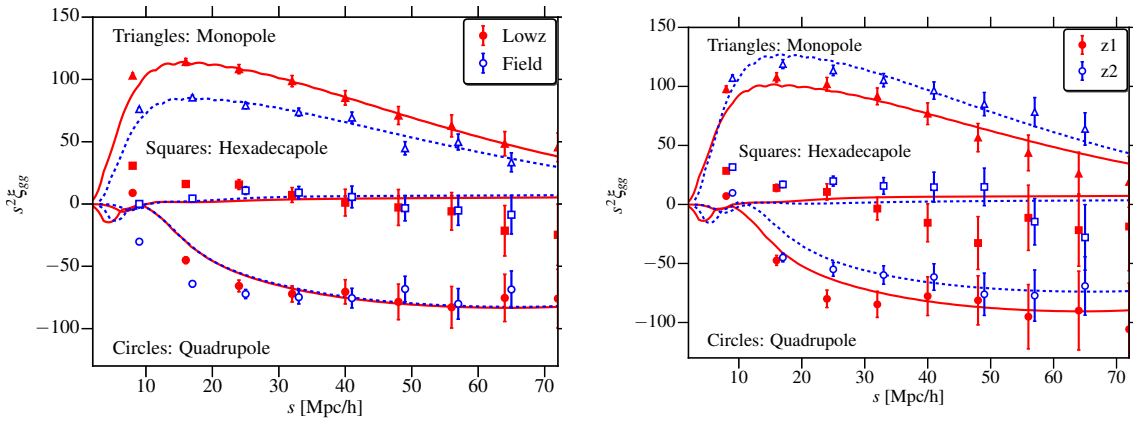



Figure 4.1: For different subsamples of the LOWZ sample in the two panels, we show the multipole moments – the monopole ($l = 0$), quadrupole ($l = 2$) and hexadecapole ($l = 4$) – of the galaxy correlation function, along with the best-fitting model.

of ΛCDM background we are doing a fair and consistent treatment of $f\sigma_8$. Our measurements of the growth rate are consistent with previous measurements using the SDSS sample (Alam et al., 2015b; Sánchez et al., 2014; Samushia et al., 2014; Satpathy et al., 2016; Beutler et al., 2016; Grieb et al., 2017; Sanchez et al., 2016; Alam et al., 2016a).

We note that the measurement of $\beta$ is obtained using scales above 25 $h^{-1}$Mpc whereas our final $E_G$ measurements extends to scales below 25 $h^{-1}$Mpc. Ideally one would want to perform measurement of $\beta$ using the same scales. Unfortunately the modeling technique used in the current measurement of $\beta$ is not accurate enough to extend it to smaller scales. The $\beta$ consists of two quantities growth rate and bias. We do account for the fact that bias will be scale dependent and different at smaller scales compared to the large scale measurement through a correction factor $C_{nl}$ (see Section 4.2.6 for details). But we have an inherent assumption that the growth rate measured using larger scales is constant and applicable for smaller scales. This makes the current measurements of $E_G$ slightly weaker than its full potential which should be improved upon in the future measurements with better RSD modeling to smaller scales.

### 4.4.2 $E_G$ in simulations and corrections

In Fig. 4.2 we show the $E_G$ measurements in the simulation box with snapshots at $z = 0.25$ and $z = 0.4$, in both real space and redshift space. Without applying any corrections, $E_G$ is biased low with considerable scale dependence. The average $E_G$ over all scales is low by $\sim 5\%$ compared to the ΛCDM predictions both in real space and redshift space. RSD effects are important at large scales, but due to relatively lower $S/N$ at those scales these do not contribute much when computing the mean $E_G$Still we do correct for these effects as described in Section 4.2.6. The biggest contribution to the bias in $E_G$ is from the combined effects of non-linear galaxy bias and the galaxy matter cross-correlation coefficient. We compute this bias directly from simulations and then correct for it.

The corrections computed using different simulations give consistent (to within 1%) mean $E_G$ measurements. Thus for our main results in the data, we will use the corrections from $z = 0.25$ snapshot for both full LOWZ and LOWZ-Z1 sample and $z = 0.40$ snapshot for LOWZ-Z2 sample. The mean $E_G$, computed after applying the corrections is within $\sim 1\%$ of the predicted value from the ΛCDM model (compared to 5% bias in the uncorrected $\langle E_G \rangle$). In appendix 4.D, we provide more tests using another set of mocks with varying HODs and show that the corrections used for our main results ($z = 0.25$ snapshot) recover $\langle E_G \rangle$ to within $\lesssim 2\%$. Thus we will add 2% systematic uncertainty in our error budget to account for the variations in corrections computed, especially for non-linear

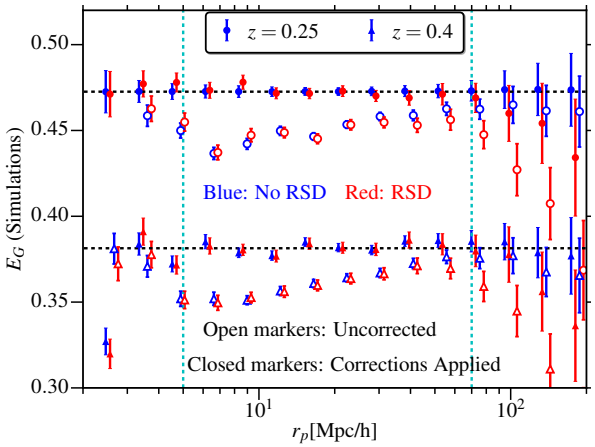



Figure 4.2: $E_G$ measurements using simulations. Points with errorbars show the measurements, while dashed black lines show ΛCDM predictions at two redshifts. Vertical cyan lines mark the scales $5 < r_p < 70h^{-1}$Mpc within which we measure the $\langle E_G \rangle$. Results for $z = 0.4$ have been shifted below by 0.05. Blue points show the measurements done using real space positions for galaxies while red points show measurements done in redshift space. Open markers show measurements without any corrections applied while closed markers show results obtained after applying all corrections (but note that the RSD correction is applied in all cases.). We find that the most important correction is for the combined effects of non-linear bias and $r_{\rm cc}$ $(\frac{\Upsilon_{gg}}{b_g^{\rm lin}\Upsilon_{gm}})$. Various corrections terms for LOWZ are shown in Fig. 4.5.

scales.

### 4.4.3 $E_G$ measurement

In Fig. 4.3, we show the $E_G$ measurement using the BOSS LOWZ sample along with both SDSS and CMB lensing maps. In the case of $E_G$ measured using galaxy-galaxy lensing, the errors are dominated by uncertainties in the RSD measurements ($\sim 10\%$), with lensing uncertainties being subdominant. $E_G$ is $\sim 2\sigma$ low compared to Planck ΛCDM predictions, which is primarily driven by the low amplitude of lensing measurements as discussed in Section 4.B. In case of CMB lensing, the error contributions from lensing are greater and thus lensing dominates the uncertainties in $E_G$The mean $E_G$ from CMB lensing is also consistent with Planck ΛCDM predictions, though the measurement is noisier when compared to galaxy lensing and is consistent with the galaxy lensing measurement.

$E_G$ measured from field sample is consistent with the results from full LOWZ sample (using same corrections for both). Since field sample does not contain groups, the effects of non-linear bias and satellite contamination are expected to be smaller. However, with CiC cylinder size, $r_p \lesssim 0.8h^{-1}$Mpc we do not observe any significant effects of removing groups at scales $r_p > 5h^{-1}$Mpc, other than reduced effective bias of the sample. CiC selection effects still bias the RSD measurements and hence we do not use the $f = 0.74 \pm 0.06$ measured from field sample.

In fig. 4.4, we show the $E_G$ measurement for the CMASS sample using CMB lensing measurement and two estimators, $\Sigma$ and $\Upsilon$, for clustering and lensing measurements. Both measurements are consistent with Planck ΛCDM predictions, though the mean $E_G$ measured from $\Sigma$ is noisier compared to $\Upsilon$ due to different signal to noise with scales. At small scales, $\Sigma$ gives better signal-to-noise ratio $(S/N)$ while at larger scales the $\Sigma$ measurement gets noisier. $\Upsilon$ on the other hand has lower $S/N$ at small scales since we are subtracting out some signal while at large scales it has better $S/N$ because of a reduced contribution from cosmic variance. Note that corrections for CMASS sample are computed using the $z = 0.6$ snapshot.

In Figure 4.5, we show the magnitude of several corrections applied to the LOWZ $E_G$ mea-

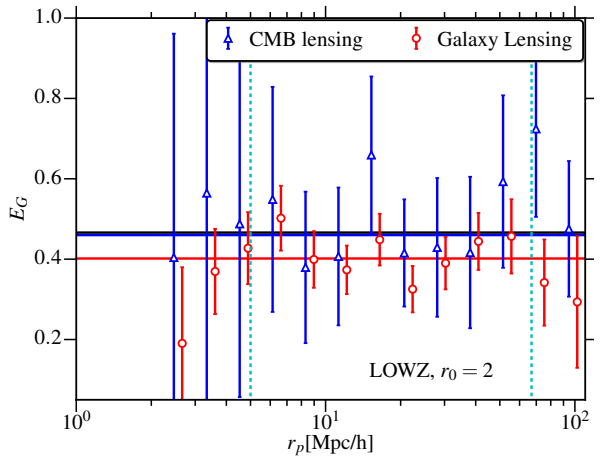



Figure 4.3: $E_G$ as a function of scale for LOWZ galaxies. The solid black line shows the ΛCDM prediction at the effective redshift of $z = 0.27$, while the solid red and blue lines show measured $\langle EG \rangle$ over scales $5 < r_p < 65h^{-1}$Mpc. The bands show the $1\sigma$ errors on $\langle E_G \rangle$.

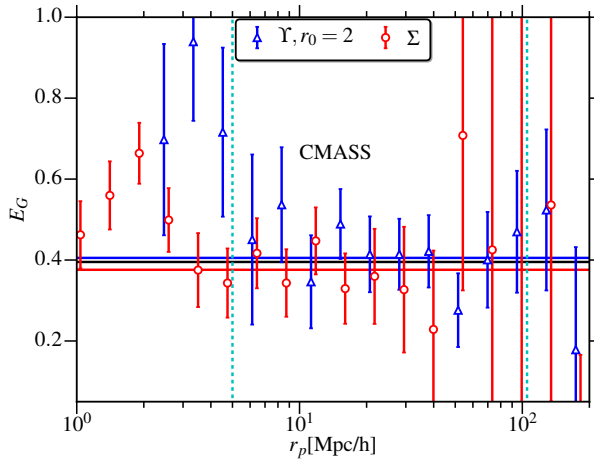



Figure 4.4: $E_G$ measurement for the CMASS sample with CMB lensing. We show the measurement using the $\Upsilon$ ($r_0 = 2h^{-1}$Mpc) estimator as well as using $\Sigma$, for both CMB lensing and clustering. $\Sigma$ provides a better $S/N$ at small scales, while $\Upsilon$ has better $S/N$ at large scales due to reduced cosmic variance (Baldauf et al., 2010).

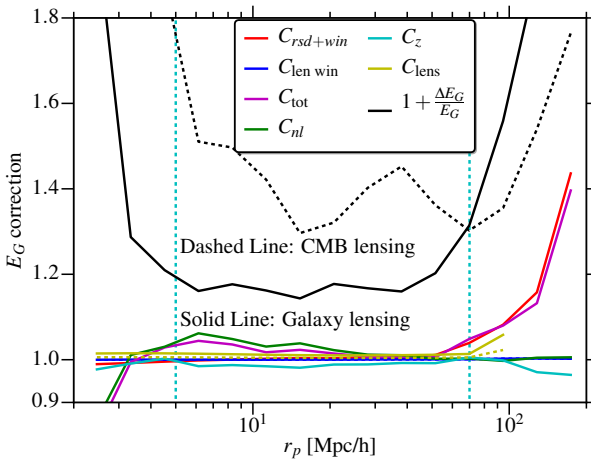



Figure 4.5: Various correction terms applied to the $E_G$ measurements in Fig. 4.3 (LOWZ sample). The dashed black line also shows the relative uncertainty in $E_G$ as a function of scale.

surements along with the relative (statistical) errors in the $E_G$ measurement for comparison. The $C_{\rm rsd+win}$ correction is computed from a theoretical model (linear theory+Kaiser correction). We also computed this correction using simulations (computing correlations with and without RSD) and find good agreement between theoretical estimates and simulations. Simulations are noisier, hence we use the theoretical estimates in the final results.

All the scale-dependent corrections we applied are subdominant when compared to the statistical uncertainties in the $E_G$ at that $r_p$, and the combined effect of corrections is to change $E_G$ by $\sim 1-2\%$ ($\lesssim 0.2\sigma$) for different samples. While this is not a significant change in our results, this suggests that future work with much better measurements will have to model these corrections to higher precisions.

We also note that lensing measurement from galaxies is susceptible to systematics from uncertainties in shear estimation and photometric redshifts. As mentioned in section 4.3.3, Reyes et al. (2012); Nakajima et al. (2012); Mandelbaum et al. (2013) did extensive testing of the shape sample for these systematics and we use the calibration factors derived for the shape sample derived in those papers. We further test the accuracy of the calibration factors for photo-z in appendix 4.E using a method with independent assumptions from the method used in Nakajima et al. (2012). There, the existence of a representative spectroscopic sample was assumed; here, we use the clustering redshift method, which assumes the existence of a non-representative spectroscopic sample that can be used to derive the ensemble redshift distribution by modeling the clustering signals. In appendix 4.E we show that while the calibration factors from the clustering redshift method depend on the priors on the galaxy bias, they agree with the factors derived by Nakajima et al. (2012) to within $\sim 5\%$. Using physically-motivated priors on the galaxy bias further improves the agreement.

There is also a possibility of contamination from intrinsic alignments (IA) of source galaxies. Blazek et al. (2012) did estimations for the IA contamination in the source sample used in this work and found no evidence for contamination with the conservative upper limit of contamination being 5%. We thus ignore the possible contamination from IA in this work.

Finally, in Fig. 4.6, we show the redshift dependence of $E_G$ along with $\Lambda$CDM predictions using different values of $\Omega_m$. In general our measurements for various samples are consistent with the Planck $\Lambda$CDM prediction at $2\sigma$ level or less, though measurements are not constraining enough right now to rule lower values of $\Omega_m$, especially with galaxy-galaxy lensing measurements preferring somewhat lower values. For the full Lowz sample and CMASS sample, our measurements are consistent with those of Blake et al. (2016), who measured $E_G = 0.48 \pm 0.10$ for Lowz ($0.15 < z < 0.43$) and $E_G = 0.30 \pm 0.07$ for the CMASS sample and also with measurement of (Alam et al., 2017), $E_G = 0.420 \pm 0.056$, for CMASS sample.

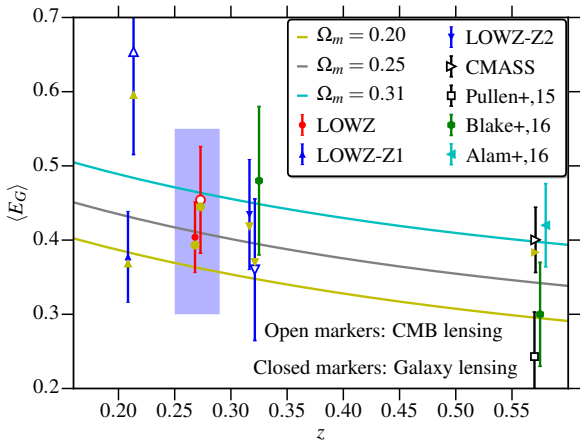



Figure 4.6: $E_G$ measurement between scales $5 < r_p < 70\ h^{-1}$Mpc for different samples, along with ΛCDM predictions with different $\Omega_m$ values. Measurements are consistent with the Planck ΛCDM model ($\Omega_m = 0.31$), though given the size of the statistical uncertainties, it is not possible to rule out other models. The blue band around the red points is drawn to indicate that the LOWZ measurement is correlated with the measurements from the Z1 and Z2 samples. Solid yellow points show the uncorrected value of $\langle E_G \rangle$.

The measured amplitude of $E_G$ is primarily affected by the (relative) lensing amplitude $A_l$ and the growth rate, $f$, measurement from RSD. In the case of galaxy-galaxy lensing, the lensing amplitude for Z1 is lower than 1 ($\sim 2\sigma$) and $f$ is higher than the Planck ΛCDM predictions (though consistent at the $1\sigma$ level), both of which lower the $E_G$ value. In the case of the Z2 sample, the lensing amplitude is significantly lower than 1 but $f$ is also lower, which partially cancels the effects of the lower $A_l$. The final $E_G$, is below but consistent with Planck predictions at the $1\sigma$ level. For the full LOWZ sample, $f$ is close to the ΛCDM prediction, but the lensing amplitude is much lower, and the latter effect dominates the final $E_G$ measurement. We again note that the lensing amplitude is degenerate with galaxy bias, and lower lensing amplitude does not necessarily mean systematics in lensing alone.

In the case of CMB lensing, the $E_G$ measurements are largely consistent with the predictions from the Planck ΛCDM model. One notable discrepancy is between our measurement for CMASS sample and that of Pullen et al. (2016). A significant part of the discrepancy in measurement of Pullen et al. (2016) is from large scales ($r_p \gtrsim 80h^{-1}$Mpc), where the galaxy-CMB lensing amplitude was observed to be low in their measurements. Our results on the other hand are dominated by the measurements at small scales, with larger scales being noisier and not contributing much to the mean $E_G$ measurement. Even when considering the full scale-dependent $E_G$ measurement, we do not observe any significant deviations at large scales, though the uncertainties in measurement at those scales are rather large. Given that the two studies use very different estimators (Pullen et al. 2016 measurements are in Fourier space), it is hard to reconcile the measurements in the two works without doing additional tests that are outside the scope of this work.

## 4.5 Conclusions

We have presented the measurements of $E_G$ and its redshift dependence using BOSS galaxies and lensing measurements from SDSS galaxy lensing and Planck CMB lensing maps. Measurements from CMB lensing and galaxy lensing are consistent within noise for the LOWZ sample. With the higher redshift CMASS sample, the $E_G$ measurement is of comparable significance to the galaxy lensing measurements. This highlights the potential of CMB lensing to provide complimentary observations and strong consistency checks when combined with the galaxy lensing measurements(see also Singh

et al., 2016a; Schaan et al., 2016).

We also highlighted several theoretical uncertainties in computing $E_G$ and computed corrections for them. Our results showed that after applying corrections in simulations we can recover $E_G$ to about 2% accuracy. With $\gtrsim 10\%$ error in our measurement, these correction are good enough for this work. However, in future several surveys will be able to measure $E_G$ (or perform similar tests on GR) to sub-percent accuracy (Leonard et al., 2015; Pourtsidou, 2016). In case of these surveys, it will be important to compute the corrections for systematics to even higher accuracy, using better theoretical modeling with analytical models and/or simulations.

While our $E_G$ measurements are largely consistent with the predictions from Planck ΛCDM predictions, there are some deviations at $\lesssim 2\sigma$ level, especially when using galaxy lensing measurements. Though statistically not very significant in $E_G$, these deviations are primarily driven by the low amplitude of lensing measurements, which can possibly be due to systematics in either clustering and/or lensing measurements, since clustering bias and lensing amplitude are degenerate. Lensing amplitude is also degenerate with the $\sim \sigma_8\Omega_m^{1/2}$ value and low amplitude could also mean that data prefers lower $\Omega_m$ or $\sigma_8$ (or both) values compared to the Planck ΛCDM model assumed. We also see similar redshift-dependent deviations in the growth rate measurements (and measurements with CMB lensing), though uncertainties are too large to make a definitive statement. Since growth rate is also degenerate with galaxy bias, this redshift dependent deviation does suggest that it is possible that problem is from the some residual sample selection effects which affect the bias measurements.

Low amplitude lensing measurements (compared to predictions from Planck ΛCDM model) has also been observed by other lensing studies (see for ex. Hildebrandt et al., 2016; Leauthaud et al., 2017; Joudaki et al., 2017) (see also van Uitert et al., 2017, for results consistent with Planck cosmology). Leauthaud et al. (2017) performed tests for several systematics in galaxy-galaxy lensing (e.g. effects of Baryons, neutrinos, modified gravity, assembly bias, sample selection) and showed that effects of systematics can be significant, especially with several systematics being combined together. Joudaki et al. (2017) performed similar analysis for cosmic shear measurements and also showed that marginalizing over some models for systematics can relieve some of the tensions between Planck ΛCDM model and lensing measurements. These studies, along with our tests suggest that more work is required to study and model the impact of systematics in both lensing and clustering measurements.

In the near future, data from the eBOSS, DES, KiDS and HSC surveys will help to extend the growth rate and $E_G$ measurements to higher redshifts with measurement uncertainties likely to be around $\gtrsim 5\%$ level within those bins. With the advent of LSST, DESI, WFIRST, SKA and CMB Stage-IV surveys, the statistical uncertainties on $E_G$ will decrease considerably, providing percent level or better measurements. However, to make $E_G$ a strong test of gravity and ΛCDM it is imperative to improve the modeling to mitigate observational systematics as well as theoretical uncertainties to much higher precisions than done in this work.

## Acknowledgements

We thank Martin White and Beth Reid for providing us halo catalog from simulations. We also thank the SDSS-I/II/III and Planck collaboration for their efforts in providing the datasets used in this work.

RM acknowledges the support of the Department of Energy Early Career Award program. SS acknowledges support from John Peoples Jr. Presidential Fellowship from Carnegie Mellon University. SA is supported by the European Research Council through the COSFORM Research Grant (#670193)

Some of the results in this paper have been derived using the HEALPix package (Górski et al., 2005).

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.SDSS3.org/.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/ JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## 4.A Lensing window function

We define window function as the relative weight assigned at a given line-of-sight distance from lens galaxy. For a given lens-source pair, with lens at $z_l$ and source at $z_s$ with photometric redshift $z_p$, the (unnormalized) window function is given by

$$w(\Pi) = \Sigma_{\rm crit}^{-2}(z_p, z_l) \int_0^{z_s} dz \Sigma_{\rm crit}^{-1}(z_s, z) \Sigma_{\rm crit}(z_p, z_l) \delta_D \left( z - z_l - \frac{cH(z_l)}{\Pi} \right) \tag{4.40}$$

where $\Pi$ is the line-of-sight separation from the lens. $\delta_D$ is the dirac delta function and ensures that we only integrate at right distance from the lens. $\Sigma_{\rm crit}^{-2}(z_p, z_l)$ factor arises from the weights we used in $\Delta\Sigma$ estimator. $\Sigma_{\rm crit}^{-1}(z_s, z)$ is the true critical density for the at $z$, while $\Sigma_{\rm crit}(z_p, z_l)$ is the critical density that we use to convert shear to matter density when measuring the signal.

Integrating over the lens and source samples we get

$$w(\Pi) = \frac{1}{\mathcal{N}} \int dz_l P(z_l) \int_{z_l}^{\infty} dz_p P(z_p) \Sigma_{\rm crit}^{-2}(z_p, z_l) \int_0^2 dz_s P(z_s|z_p) \int_0^{z_s} dz \Sigma_{\rm crit}^{-1}(z_s, z) \Sigma_{\rm crit}(z_p, z_l) \times \delta_D \left( z - z_l - \frac{cH(z_l)}{\Pi} \right) \tag{4.41}$$

where the normalization factor $\mathcal{N}$ is defined such that $w(\Pi = 0) = 1$. Also note that the effects of bias and scatter in photometric redshifts are included in the calibration biases derived in Nakajima et al. (2012) and hence we do not include them in our calculation. We set $P(z_s|z_p) = \delta_D(z_s - z_p)$.

## 4.B Galaxy Clustering and Lensing

In this section we present the results from galaxy clustering and galaxy-lensing cross correlations with galaxy lensing and CMB lensing. These measurements were first presented in Singh et al. (2016a) and we refer the reader to that paper for more details and several systematic tests, especially for the CMB lensing measurements.

Fig. 4.7 shows the clustering and lensing measurements, $w_{gg}$ and $\Delta\Sigma$, as well as $\Upsilon_{gg}$ and $\Upsilon_{gm}$ for various samples along with the Planck $\Lambda$CDM model with best fit values for bias and lensing amplitude, $A_l$. To get the values of linear bias and $A_l$, we jointly fit the clustering and lensing measurements with fixed cosmology, where due to very different S/N in two measurements, the bias is determined by clustering and $A_l$ is determined by the lensing. The values of bias and lensing amplitude are presented in Table 4.1. We fit the model only in the range $10 < r_p < 65h^{-1}$Mpc. Below $10h^{-1}$Mpc, non-linear effects not included in the model become important, while $65h^{-1}$Mpc is approximately the size of jackknife regions at the minimum redshift of the LOWZ sample, $z = 0.16$. For the CMASS sample, we extend the upper limit in $r_p$ to $100h^{-1}$Mpc.

The values of galaxy bias that we obtain are consistent with other works (Singh et al., 2015; Rodríguez-Torres et al., 2016) and follow the expected trends, with higher redshift subsamples having

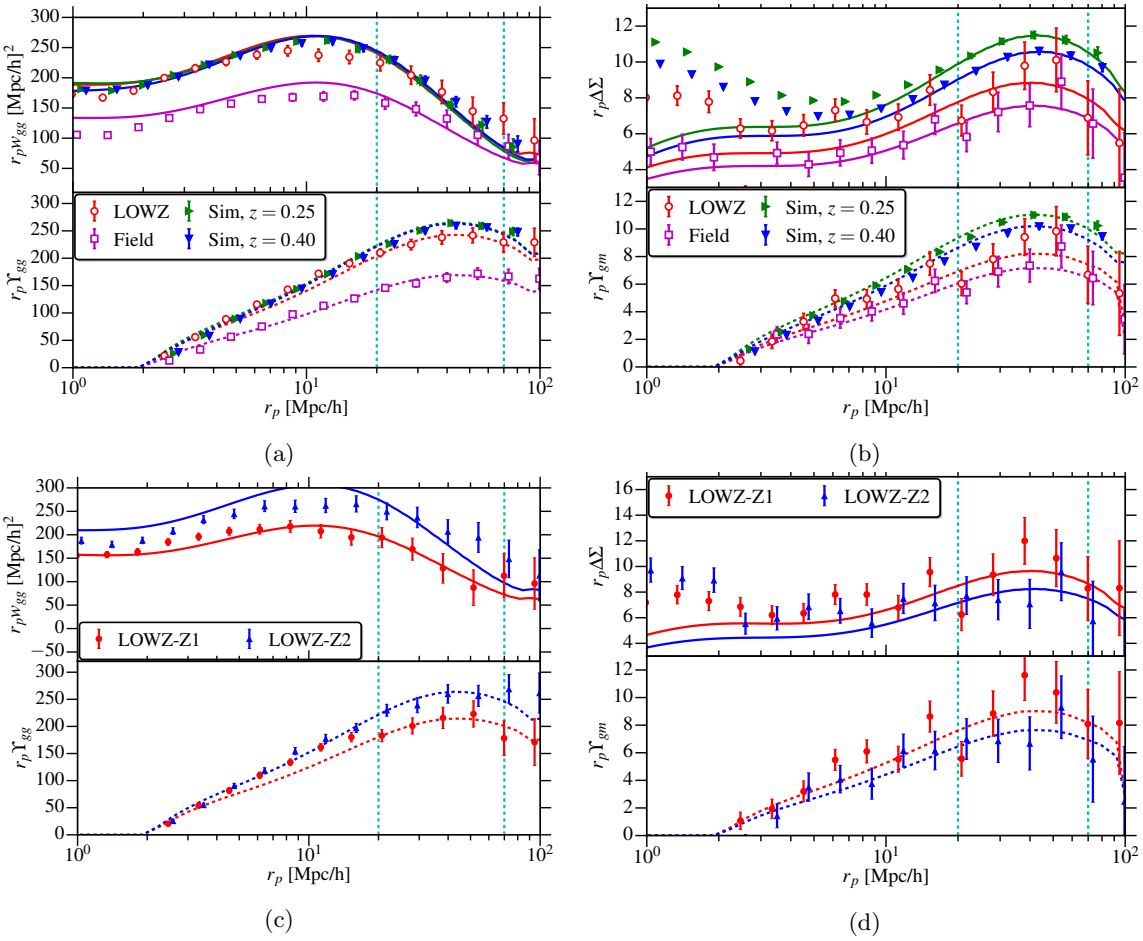



Figure 4.7: *Left Column*: Galaxy-galaxy correlation functions, $w_{gg}$ and $\Upsilon_{gg}$, for the LOWZ sample and various sub-samples. *Right Column*: The galaxy-matter cross-correlation functions from weak lensing, $\Delta\Sigma$ and $\Upsilon_{gm}$, for the LOWZ sample and various sub-samples. The solid (dashed) lines are the best fit models (Planck ΛCDM model with best fit bias) to $\Delta\Sigma$ and $w_{gg}$ (or $\Upsilon_{gm}$ and $\Upsilon_{gg}$). In all cases we fit the model jointly to lensing and clustering ($w_{gg}$ and $\Delta\Sigma$ or $\Upsilon_{gg}$ and $\Upsilon_{gm}$) in the range $10 < r_p < 65h^{-1}$Mpc, marked by vertical dashed lines. In the top row, we also show the measurements from the simulations. The galaxy clustering measurements match well by construction (HOD in simulations match the clustering with data), while the lensing measurements from data have lower amplitude compared to simulations, consistent with lower than 1 lensing amplitude measured from model fits.

higher bias (expected for galaxies with similar halo mass as they tend to form in more biased regions at higher redshifts and also from selection effects as mean luminosity and halo mass also tend to increase with redshift). However, the values of lensing amplitude, $A_l$, are smaller than the expected value of 1, especially in the case of galaxy-galaxy lensing; this discrepancy increases with redshift. This suggests there is a possibility of systematics in the lensing source sample, especially in photometric redshifts, though Nakajima et al. (2012) did not find any evidence of systematics (over and above the systematics corrections we have applied) in the extensive analysis of the source sample we use. Other possible explanations come from the fact that the $A_l$ value is degenerate with the galaxy bias and the $\Omega_m$ value assumed in the model. The clustering measurement is relatively more sensitive to the (mis-estimation of) sample selection effects; since the lensing amplitude is degenerate with the galaxy bias, mis-estimation of the latter can also introduce some redshift dependent bias in the lensing amplitude.

Another possibility is that there is some discrepancy in the fiducial cosmological model we use, with the lensing data preferring lower $\Omega_m$ (compared to Planck $\Lambda$CDM), consistent with some other lensing measurements (Massey et al., 2007; Mandelbaum et al., 2013; Kilbinger et al., 2013; Heymans et al., 2013a). The lensing amplitude is degenerate with $\Omega_m$, and a lower $\Omega_m$ value will push the $A_l$ higher. In the case of CMB lensing (not shown, see Singh et al. (2016a) for details), the lensing amplitude is consistent with the expectation of 1, though redshift trends are similar to galaxy lensing. CMB lensing measurements are noisier at low redshifts due to the rapid drop in the CMB lensing kernel and are consistent with the galaxy-galaxy lensing measurements within the errors. It is also worth noting from Table 4.1 that growth rate measurements (described in Section 4.4.1), though consistent with $\Lambda$CDM predictions within the noise ($\sim 2\sigma$ deviation for the $Z2$ sample), also follow trends similar to lensing, with values being lower at higher redshifts. Since the $f$ measurement is also degenerate with the bias, this suggests the presence of some low level systematics in the clustering measurement and/or that data prefers lower $\Omega_m$ value.

## 4.C Estimating lensing magnification bias to $E_G$

### 4.C.1 Bias in Clustering

Due to the effect of lensing by the foreground structure, the apparent galaxy over density is modified from the true over density. This modification arises from two effects, the volume perturbations due to magnification, and the modification to galaxy selection function in a flux limited sample.

To estimate the impact of lensing contaminations, we begin by assuming that the true number density of galaxies, $n_0$, for a given flux limit, $f$, follows a simple powerlaw relation

$$n_0(>f) = Af^{-\alpha} \tag{4.42}$$

Where slope $\alpha$ is measured as

$$\alpha = -\frac{d\log\left(n_0(>f)\right)}{d\log\left(f\right)} = 2.5\frac{d\log\left(n_0(<m)\right)}{dm} \tag{4.43}$$

The observed number density of galaxies, $n$, changes due to the magnification $\mu$ as

$$n(>f) = \frac{1}{\mu}n_0(>\frac{f}{\mu}) \tag{4.44}$$

$$= \frac{1}{\mu}\mu^{\alpha}n_0(>f) \tag{4.45}$$

$$\approx (1+2(\alpha-1)\kappa)n_0(>f) \tag{4.46}$$

Where we used the relation between magnification and convergence, $\kappa$, $\mu \approx 1/(1-\kappa)^2$.

The observed galaxy overdensity, $\widehat{\delta}_g$, is then related to the true overdensity , $\delta_g$, upto first order as

$$\widehat{\delta}_g \approx \frac{(1+\kappa')n_g}{\overline{n}_g} - 1 \approx \delta_g + \kappa' \tag{4.47}$$

Where $\kappa' = 2(\alpha-1)\kappa$.

The measured correlation function is then

$$\widehat{\xi}_{gg}(r_p,\Pi) = \xi_{gg}(r_p,\Pi) + \int dz p_g(z) p_g(z+\Delta z_\Pi)\kappa'(z)\kappa'(z+\Delta z_\Pi)\left(\theta = \frac{r_p}{f_k(\chi_z)}\right) + \int dz p_g(z) p_g(z+\Delta z_\Pi)\delta_g(z)\kappa'(z+\Delta z_\Pi)(r_p) \tag{4.48}$$

$$\widehat{w}_{gg}(r_p) = w_{gg}(r_p) + \int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} d\Pi \int dz p_g(z) p_g(z+\Delta z_\Pi)\kappa'(z)\kappa'(z+\Delta z_\Pi)\left(\theta = \frac{r_p}{f_k(\chi_z)}\right) + \int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} d\Pi \int dz p_g(z) p_g(z+\Delta z_\Pi)\delta_g(z)\kappa'(z+\Delta z_\Pi)(r_p) \tag{4.49}$$

$$= w_{gg}(r_p) + [2(\alpha-1)]^2 \int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} d\Pi \int dz W(z) \int_0^{z+\Delta z} d\chi_m \frac{\overline{\rho}_m^2 w_{mm}\left(r'_p = \frac{r_p f_k(\chi_{z_m})}{f_k(\chi_z)}, z_m\right)}{\Sigma_c(z_m,z)\Sigma_c(z_m,z+\Delta z)} + 2(\alpha-1)\int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} d\Pi \int dz W(z) \frac{\overline{\rho}_m w_{gm}(r_p,z)}{\Sigma_c(z,z+\Delta z)}\Theta(z,z+\Delta z) \tag{4.50}$$

where $\Delta z_\Pi = \frac{\Pi H(z)}{c}$. The last term can be approximated as

$$\int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} d\Pi \int dz W(z) \frac{\overline{\rho}_m w_{gm}(r_p,z)}{\Sigma_c(z,z+\Delta z)}\Theta(z,z+\Delta z) = \int dz W(z)\overline{\rho}_m w_{gm}(r_p,z) \times \frac{4\pi G(1+z)}{c^2}\int_0^{\Pi_{\rm max}} d\Pi \frac{\Pi}{1+\Pi/f_k(\chi_l)} \tag{4.51}$$

$$\approx \frac{3H_0^2\Omega_m}{2c^2}\frac{\Pi_{max}^2}{2}\int dz W(z)(1+z) w_{gm}(r_p,z) \tag{4.52}$$

$$\approx 2.5\times 10^{-4}\int dz W(z)(1+z) w_{gm}(r_p,z) \tag{4.53}$$

$$\approx 2.5\times 10^{-4}(1+\overline{z}) w_{gm}(r_p) \tag{4.54}$$

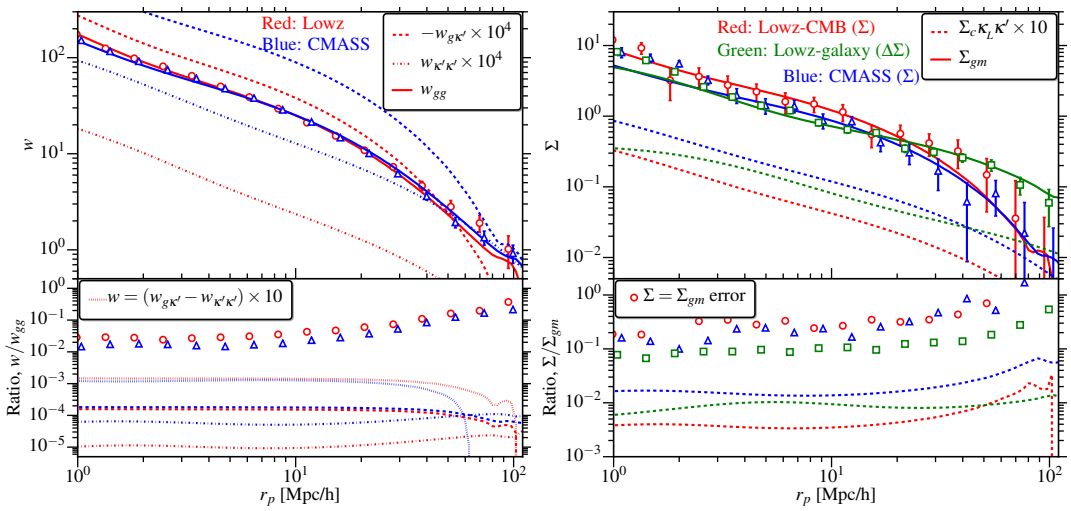
In fig. 4.8a we show the estimated contamination from lensing to the clustering measurement for the case of both LOWZ and CMASS samples. With estimated $\alpha = 0.4$ for LOWZ and $\alpha = 0.5$ for CMASS, the dominant contribution is from $w_{g\kappa'}$, which biases the clustering low by factor of $\sim 10^{-4}$, much less than the signal to noise in the measurements.

### 4.C.2 Bias in lensing

For the case of lensing, we do the stacking analysis where magnification bias is not expected to contaminate the measurement. However, we do apply the boost factor, $b_F = \frac{\sum_{ls} w_{ls}}{\sum_{Rs} w_{Rs}}$, to correct for the effect of sources associated with the lens sample. The boost factor also includes the effects from lensing magnification, $b_F \sim 1/\mu$, and thus biases the lensing measurement.

We can model the magnification bias in a similar way to clustering, as the magnification bias is changing the observed density contrast of the lens sample. For the case of galaxy-galaxy lensing

$$\widehat{\Delta\Sigma} = \langle \Sigma_c \widehat{\delta}_g \gamma_t \rangle = \Delta\Sigma + \langle \Sigma_c \gamma_t \kappa' \rangle \tag{4.55}$$




(a) (b)

Figure 4.8: a) Lensing contamination to the clustering measurements. Upper panel show the correlation function, $w_{gg}$ and the correlations due to lensing $w_{g\kappa'}$ and $w_{\kappa'\kappa'}$. Lower panel shows the ratio of predicted contamination with the best fit model. Also, the dotted red line shows the signal to noise ratio in the measurements. Note that in the upper panel (but not in lower panel, unless stated in lower panel legend), we multiplied lensing predictions with a constant factor to make them same order of magnitude as the clustering signal. b) Same as a), but showing the lensing contamination to $\Sigma$ and $\Delta\Sigma$ measurement.

And for the case of galaxy-cmb lensing

$$\widehat{\Sigma} = \langle \Sigma_c \widehat{\delta}_g \kappa_{\rm cmb} \rangle = \Sigma + \langle \Sigma_c \kappa_{\rm cmb} \kappa' \rangle \tag{4.56}$$

In fig. 4.8b, we show the estimated contamination in lensing measurements, which introduces a bias of order $\sim 1\%$, lower than the measurement noise levels of $\sim 10\%$.

### 4.C.3 Bias in $E_G$

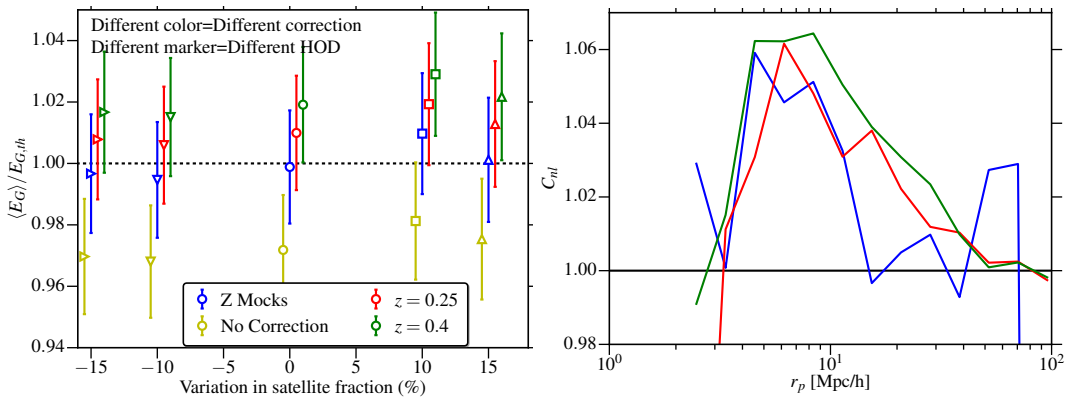
Ignoring the bias to the growth rate measurement from lensing (note that lensing bias varies with the line of sight separation and hence will bias the quadrupole in clustering as well), the bias in $E_G$ can be estimated as

$$\frac{\Delta E_G}{E_G} = \frac{\Delta\Sigma'}{\Delta\Sigma}\frac{w_{gg}}{w'_{gg}} - 1 \approx \frac{\langle \Sigma_c \gamma_t \kappa' \rangle}{\Delta\Sigma} + \frac{w_{g\kappa'}}{w_{gg}} - \frac{w_{\kappa'\kappa'}}{w_{gg}} \approx 10^{-2} \tag{4.57}$$

Final corrections applied to $E_G$ measurement are shown in fig.4.5, for the case of Lowz sample.

## 4.D Testing the $E_G$ corrections using different mocks

In this section we test the corrections, especially $C_{nl}$, on mocks with different HOD models. For this purpose, we construct light-cones from the BigMultiDark Planck simulation (BigMDPL Klypin et al., 2016) using all the available snapshots in the range $0.16 < z < 0.36$. This is a simulation of a flat $\Lambda$CDM model, with Planck1 cosmological parameters (Planck Collaboration et al., 2014a), $\Omega_m = 0.307$, $\Omega_\Lambda = 0.693$, $h = 0.678$ and $\sigma_8 = 0.829$. The BigMDPL has $3840^3$ particles with a mass resolution of $2.4{\times}10^{10}\ h^{-1}M_\odot$ and a box size of 2.5 $h^{-1}$Mpc. These features allow to produce light-cones with a volume comparable to BOSS and resolve dark matter halos in the mass range predicted




(a) (b)

Figure 4.9: a) $E_G$ measured using mocks with different HODs (variable satellite fraction) and different sets of corrections applied. b) $C_{nl}$ correction computed from different mocks.

for the LOWZ and CMASS galaxies. Dark matter halos are defined using the Robust Overdensity Calculation using K-Space Topologically Adaptive Refinement halo finder (ROCKSTAR Tinker et al., 2013).

The resulting galaxy mocks will reproduce the clustering at different redshift and the radial selection function of the observed sample. In order to populate the dark matter halos of the simulation we implement the halo abundance matching used by Rodríguez-Torres et al. (2016). The scatter is included using the maximum circular velocity over the whole history of the halo, $V_{peak}$, using the relation

$$V_{peak}^{scat} = (1 + \mathcal{G}(0, \sigma_{\rm HAM}))V_{peak}, \tag{4.58}$$

where $\mathcal{G}$ is a random gaussian number from a distribution with mean 0 and standard deviation $\sigma_{\rm HAM}$. We sort all halos using $V_{peak}^{scat}$ and then, select objects from the largest velocity and continuing down until reaching the number density of the observed sample. The scatter between halos and galaxies is fixed using the projected correlation function of the LOWZ data between $2 < r_p < 30\ h^{-1}$ Mpc for 3 different redshift bins in the range $0.16 < z < 0.36$.

We do not distinguish between host halos and subhalos in the first light-cone. This produces a LOWZ mock with a fraction of satellites equal to 12.5%. In order to produce different HOD, we include an additional parameter which models the fraction of satellites, $f_{sat}$. It will force the light-cone to have a desired fraction of subhalos. Including the new parameter, we produce 4 different light-cones, first, increasing the fraction of satellites by 10 and 15 percent and then decreasing by the same factors. In all cases, we fix the scatter to reproduce the observed clustering.

In addition to $C_{nl}$ measured using simulations presented in section 4.3.5, we also compute $C_{nl}$ from one of BigMDPL mocks. We then apply the three sets of $C_{nl}$ to mocks with different HODs to check if we recover $E_G$ accurately. In fig.4.9a, we show the relative bias in the corrected $\langle E_G \rangle$ relative to the theoretical prediction. Applying no corrections, $\langle E_G \rangle$ is biased low by $\sim 3\%$. Applying the correction from same MDPL mocks, we recover the $\langle E_G \rangle$ at level $\lesssim 0.5\%$ independent of the HOD. However, applying correction from different sims over correct and $\langle E_G \rangle$ is biased high. As we show in fig. 4.9b, this is primarily driven by differences in the non-linear physics, which leads to different $C_{nl}$, across different sims. Corrections used for our main results, $z = 0.25$ snapshot, recovers $\langle E_G \rangle$ to $\sim 2\%$ for all HODs.

## 4.E Photo-z bias estimates using clustering redshifts

In this section, we present the estimates of the true redshift distribution $N(z)$ for our photometric redshifts, closely following the formalism developed by Ménard et al. (2013); van Daalen & White (2017). We use a random subsample of our source sample containing $\sim 2 \times 10^6$ galaxies and split it into 40 photo-$z$ bins containing approximately equal numbers of galaxies. These samples are then cross-correlated with the BOSS spectroscopic sample (LOWZ+CMASS), which is split into bins of size $\delta z = 0.05$ in the range $z \in [0.1, 0.7]$ and two additional bins in the ranges $z \in [0, 0.1]$ and $z \in [0.7, 1]$.

To compute auto- and cross-correlations between the spectroscopic samples, we use the the standard Landy-Szalay estimator, but for cross-correlations involving the source sample we use the sub-optimal estimator as we do not have a good estimate for the on-sky selection function of the source catalog and hence we lack random catalogs for that sample:

$$w_{s_i-s_j}(\theta) = \frac{\sum_\theta (D_{s_i} - R_{s_i})(D_{s_j} - R_{s_j})}{\sum_\theta R_{s_i} R_{s_j}} = \frac{\sum_\theta (D_{s_i} D_{s_j} - R_{s_i} D_{s_j} - D_{s_i} R_{s_j} + R_{s_i} R_{s_j})}{\sum_\theta R_{s_i} R_{s_j}} \tag{4.59}$$

$$w_{p_i-s_j}(\theta) = \frac{\sum_\theta D_{p_i} (D_{s_j} - R_{s_j})}{\sum_\theta D_{p_i} R_{s_j}} = \frac{\sum_\theta (D_{p_i} D_{s_j} - D_{p_i} R_{s_j})}{\sum_\theta D_{p_i} R_{s_j}} \tag{4.60}$$

Here the subscript $p_i$ refers to the $i^{th}$ photometric samples, $s_i$ refers to the $i^{th}$ spectroscopic sample and $\sum_\theta$ indicates that summation is carried over pairs with separation within the given bin's limits. Then we integrate the correlation functions to obtain

$$W_{a-b} = \int_{\theta_{\rm min}}^{\theta_{\rm max}} w_{a-b}(\theta) \mathcal{W}(\theta) \theta d\theta \tag{4.61}$$

The weight function $\mathcal{W}(\theta) = 1/\theta$, as suggested from Ménard et al. (2013). We use $\theta_{\rm min} = 0.03$ degree and $\theta_{\rm max} = 0.5$ degree. These choices are primarily dictated by the noise and systematic bias in the measured cross-correlation functions, primarily dictated by the SDSS number density at small $\theta$ and our use of the sub-optimal estimator at large $\theta$, which increases the variance (Singh et al., 2016b, there is also some evidence of biases due to selection function effects in some bins).

The cross correlation between a photometric sample, $p$, and spectroscopic sample $s_i$ can be modeled as

$$W_{p-s_i} = \sum_{j=1}^{N_s} b_{R,p-s_j} f_{p,s_j} W_{s_i-s_j}. \tag{4.62}$$

$N_s$ is the number of spectroscopic subsamples, $f_{p,s_j} \approx p(z_s)\Delta z_s$ is the fraction of the photometric sample within the redshift range of spectroscopic bin $j$ and the bias ratio $b_{R,p-s_j}$ is defined as

$$b_{R,p-s_j} = \begin{cases} \frac{b_{p,s_j}}{b_{s_j}} & z_{s_j} \leq z_{s_i} \\ 1 & \text{otherwise.} \end{cases}$$

$b_{p,s_j}$ is the (integrated) bias of photometric galaxies that lie in redshift bin $z_{s_j}$ and $b_{s_j}$ is the bias of spectroscopic galaxies (measured by fitting the linear+halofit matter correlation function to $w_{s_j-s_j}$) in the same bin. If $z_{s_j} \leq z_{s_i}$, galaxies at $z_{s_j}$ are either correlated with galaxies at $z_{s_i}$ ($z_{s_j} = z_{s_i}$) or they are at lower redshift and are lensing the galaxies at higher redshift $z_{s,i}$. When $z_{s_j} > z_{s_i}$, galaxies at $z_{s_i}$ act as lens and hence the correlations depend only on the bias of the spectroscopic galaxies in $z_{s_i}$ and the ratio is 1.

In our fitting procedure, we leave $b_{p,s_j}$ as a free parameter, though we do put hard lower and upper limits on the values it can take and also assume that it follows some simple redshift evolution. For hard limits, we use $b_l < b_{p,s_j} < 3 \,\forall\, p, s_j$, where the upper limit is fixed to 3 and for the lower limit we try 3 different values $b_l = [0, 0.4, 0.7]$. $b_l \approx 0.7$ is motivated by the lower limit of halo bias in simulations (e.g., Tinker et al., 2010). However, to accommodate any effects of non-linear physics

(non-linear bias and non-unity cross-correlation coefficient between galaxy sample), we also do the analysis with different lower limits of galaxy bias. To fit for the redshift evolution, we try two different functions, linear and power-law, with two free parameters, $b_0$ and $b_1$, to be fitted,

$$b_p(z) = b_0 \left(\frac{1+z}{1+\overline{z}_p}\right)^{b_1} \tag{4.63}$$

$$b_p(z) = b_0 \left[1 + b_1(z - \overline{z}_p)\right] \tag{4.64}$$

where $\overline{z}_p$ is the mean photometric redshift of the given photometric subsample.

For each photometric redshift bin, we adopt a Gaussian likelihood as

$$L_W = \frac{1}{\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\widehat{W}_{p-s_i} - W_{p-s_i})^T \Sigma^{-1} (\widehat{W}_{p-s_i} - W_{p-s_i})\right) \tag{4.65}$$

We fit for $f_{p,s_j}, b_0$, and $b_1$. To fit for the fractional distribution of galaxies, $f_{p,s_j}$, we use two separate methods. In the first method, we assume the $p(z_s)$ is Gaussian and integrate the $p(z_s)$ within a bin to obtain $f_p$. In this case, the fitting parameters are the mean and variance of the Gaussian $p(z_s)$. However, using a Gaussian function for the $p(z_s)$ is a simplification that may not be justified in reality. So we also use a non-parametric fitting procedure where each of the $f_{p,s_j}$ are free parameters. In this case, we require that $-f_l < f_{p,s_j} < 1 + f_l$. We test the values $f_l = [0, 0.05, 0.1]$, where values $> 0$ are necessary to accommodate shot noise. While the second method is in principle the most general form for $f_p$, the results have a higher variance and can be more susceptible to the effects of bias-$f_p$ degeneracy. When computing the photo-z bias in lensing, any bins $j$ with best-fitting $f_{p,s_j} < 0$ are set to zero (since bins with no galaxies due to shot noise will not contribute anything to lensing).

In addition, we also require that the sum of $f_{p,s_j}$ ($\sum_j f_{p,s_j}$) should be close to 1. For this we add an additional term in the likelihood as

$$\mathcal{L} = L_W \frac{1}{\sqrt{N_{p,f}}} \exp\left(\frac{-1}{2} \frac{N_{p,f}^2 (\sum_j f_{p,s_j} - 1)^2}{N_{p,f}}\right) \tag{4.66}$$

$$N_{p,f} = \frac{500}{\overline{z}_p} \tag{4.67}$$

Where $N_{p,f}$ should ideally be the number of galaxies in the given photometric sample. However, since the redshift extent of the spectroscopic sample is less than that of the photometric sample, we loosen the likelihood for samples at higher photometric redshift, hence the dependence on $z_p$. 500 was chosen so that $500/min(z_p)$ is of order 50,000, which is the number of galaxies in each photometric bin.

Once we have the spectroscopic redshift distribution for given photometric samples, we can estimate the bias, $B_L$ in measured $\widehat{\Delta\Sigma}$ as (Nakajima et al., 2012)

$$B_L = \frac{\widehat{\Delta\Sigma}}{\Delta\Sigma} = \frac{\int dz_l p(z_l) D(z_l)^2 \int dz_p p(z_p) w_{ls}(z_l, z_p) \int dz_s p(z_s|z_p) \frac{\Sigma_c(z_l,z_p)}{\Sigma_c(z_l,z_s)}}{\int dz_l p(z_l) D(z_l)^2 \int dz_p p(z_p) w_{ls}(z_l, z_p) \int dz_s p(z_s|z_p)} \tag{4.68}$$

$z_l$ is the lens redshift, $z_p$ is the photometric redshift and $z_s$ is the true source redshift. $D(z_l)$ is the growth function at lens redshift, $w_{ls}$ is the weight as described in eq. (4.9).

In Fig. 4.10, we show the fractional distribution obtained assuming a Gaussian model for $p(z_s)$ (blue points) and when using the non-parametric $f_p$ model. We also show the $f_p$ obtained using a representative calibration sample with spectroscopic redshifts. $f_p$ obtained using the Gaussian model is less noisy, but for any given photometric sample, the bins in true redshift are strongly correlated. In the case where the $f_p$ are treated as independent parameters, the correlations across bins are not significant and the distributions appear to capture more features, though to some extent this can be caused by noise in clustering measurements and the degeneracies with the galaxy bias.

In Fig. 4.11 we show the bias in lensing due to the impact of photometric redshift bias and scatter, calculated using the true redshift distributions estimated from the clustering analysis. We show the

estimates using different bias and $f_p$ models along with different priors on the galaxy bias and $f_p$. Our results are not very sensitive to the bias evolution models, likely due to the relatively narrow redshift bins. However, in the case where the $f_p$ values are treated as free parameters, are results are sensitive to the priors on $f_p$ and the bias, especially the lower limit on bias. This is primarily caused by the degeneracy between the bias and $f_p$, along with the noise in the correlation function measurements that makes it difficult to infer so many free parameters. We attempt to overcome this difficulty using the gaussian $p(z_s)$. The results in this case are less sensitive to priors on galaxy bias. However, the signal to noise in the lensing bias is not improved by much, which suggests we are not gaining much more information by constraining $p(z_s)$ (the $f(z_s)$ obtained in this case are correlated across bins). As a result we do not attempt to use more complicated functions than gaussian which may capture more features of $p(z_s)$.

For the physically-motivated choice of priors, $b_l \in [0.4, 0.7]$ and $f_l \in [0.05, 0.1]$, our results are consistent (within $\lesssim 5\%$) with the bias in the lensing signal estimated using the calibration sample. This value has been used in this and previous works to correct the lensing measurements.

Our results suggest that more work is required to carefully choose the priors and functional forms for $f_p$ to further develop the clustering redshift method (as used in this work) for lensing calibration. However, our results demonstrate that the bias in our galaxy-galaxy lensing signals due to photometric redshift bias and scatter is well constrained at the $\lesssim 5\%$ level and hence cannot explain the $\sim 20\%$ discrepancy observed between the $E_G$ measurements and the prediction from the Planck $\Lambda$CDM model.

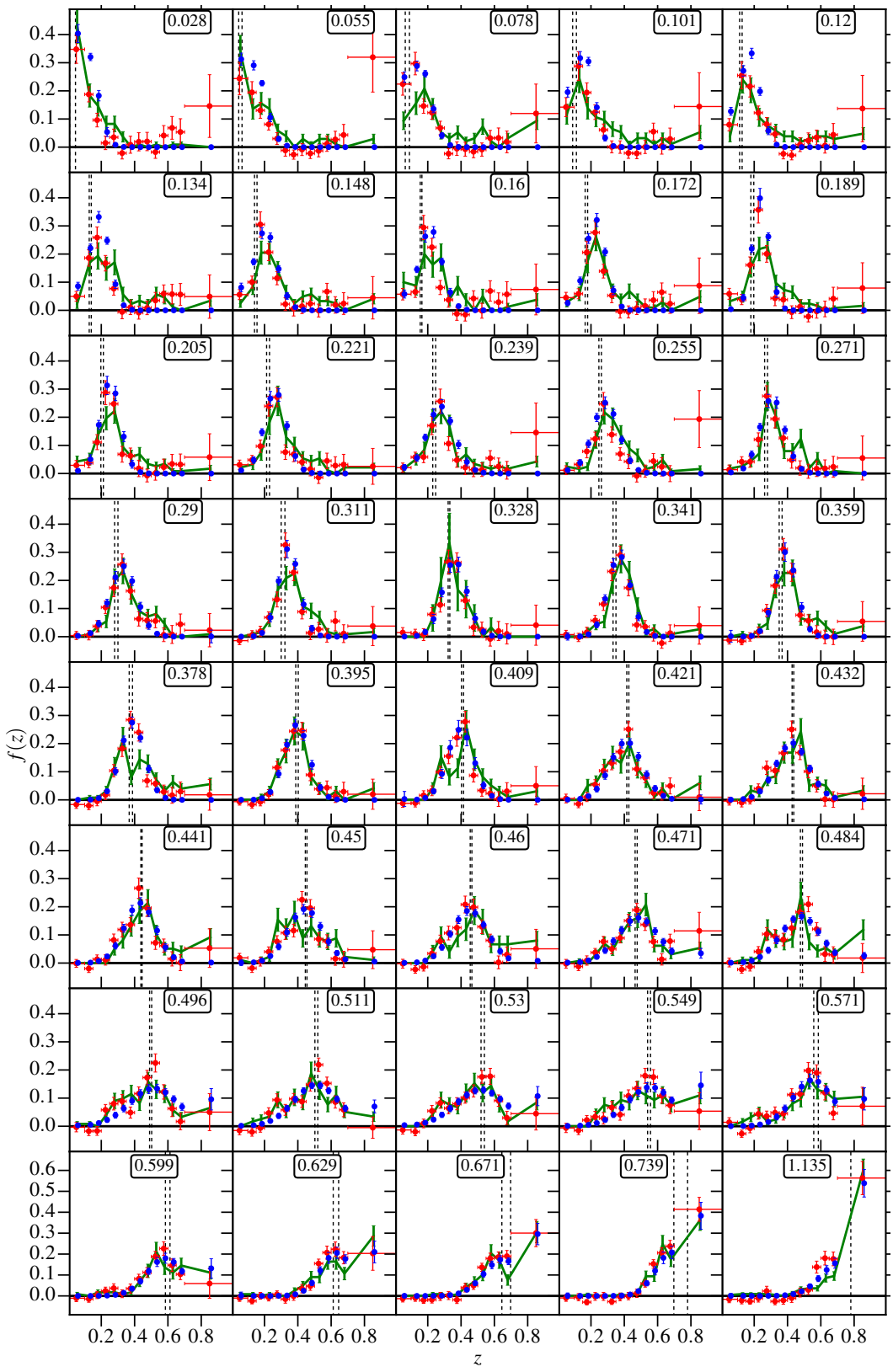
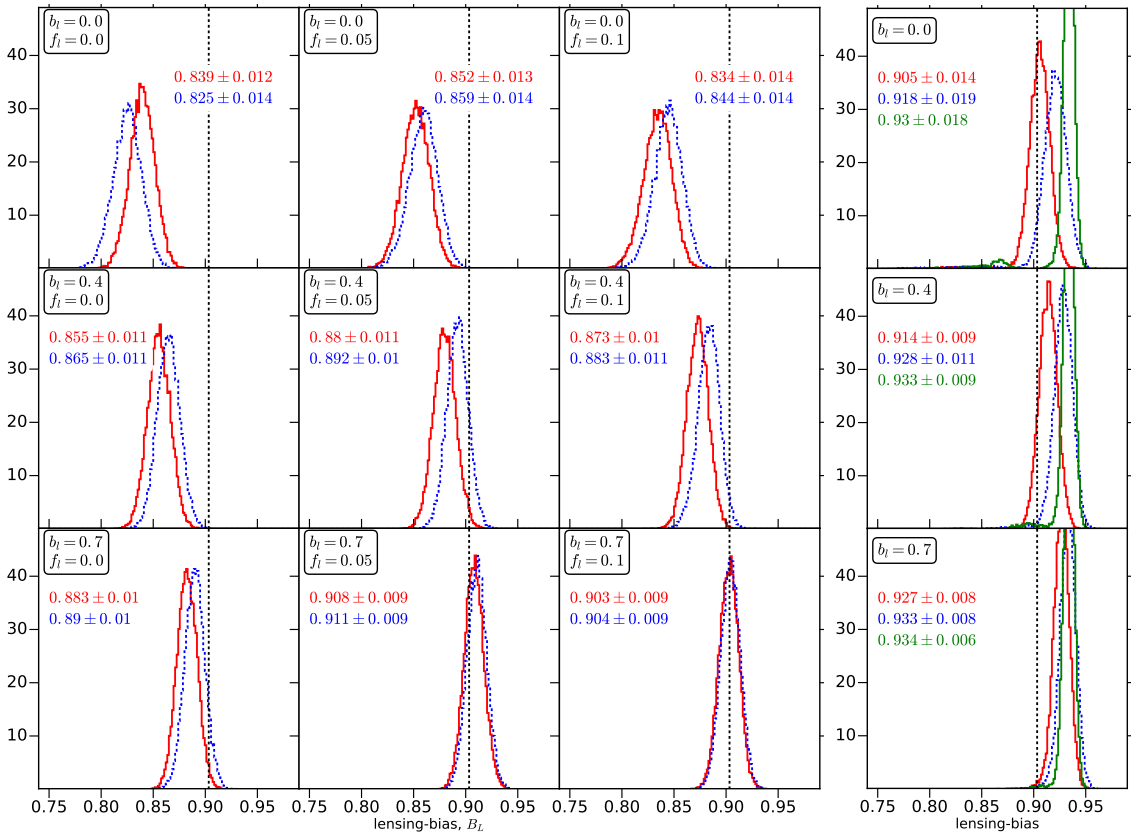



Figure 4.10: Redshift distribution estimates $f(z)$ for different photometric redshift bins, where each panel represents a given photo-z selected sample. Red and blue points are for the fits with $f(z)$ in each bin as a free parameter and assuming a Gaussian redshift distribution, respectively. For both red and blue curves in this plot we use $b_l = 0.7$; for the red points, $f_l = 0.05$. The green curves show the $f(z)$ obtained using spectroscopic redshifts from the representative calibration sample.




(a) (b)

Figure 4.11: a) Probability distribution of bias in the lensing signal, $B_L$, due to photometric redshifts, for different choices of lower limit on bias and redshift fraction. The vertical dashed black lines show the estimate using spectroscopic redshifts in the representative calibration sample. The solid red and dotted blue lines use the power-law and linear models for the bias-redshift relation in the photometric sample, respectively. There is strong dependence of the inferred bias on the lensing signal on the priors chosen, though the choice of model for the bias-redshift relation does not affect it too strongly. In all cases, the difference between the lensing bias from the clustering redshift method and that derived using spectroscopic redshifts ($\sim 0.91$) is $\lesssim 10\%$. b) Same as a), using gaussian model for $p(z)$. Green curves show the estimate when we assume galaxy bias to be constant with redshift ($b_1 = 0$ in eq.(4.63)).

# 5 Intrinsic alignments of SDSS-III BOSS LOWZ sample galaxies

Sukhdeep Singh[1], Rachel Mandelbaum[1], Surhud More[2]

[1]McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[2]Kavli Institute for the Physics and Mathematics of the Universe (WPI), TODIAS, The University of Tokyo, Chiba, 277-8583, Japan

**Abstract**

Intrinsic alignments (IA) of galaxies, i.e. correlations of galaxy shapes with each other (II) or with the density field (gI), are a major astrophysical source of contamination for weak lensing surveys. We present the results of IA measurements of galaxies on 0.1–200$h^{-1}$Mpc scales using the SDSS-III BOSS LOWZ sample, in the redshift range $0.16 < z < 0.36$. We extend the existing intrinsic alignments measurements for spectroscopic LRGs to lower luminosities, and show that the luminosity dependence of large-scale intrinsic alignments can be well-described by a power law. Within the limited redshift and color range of our sample, we observe no significant redshift or color dependence of intrinsic alignments. We measure the halo mass of LOWZ galaxies using galaxy-galaxy lensing, and show that the mass dependence of large-scale intrinsic alignments is also well described by a power law. We detect variation in the scale dependence of intrinsic alignments with mass and luminosity, which underscores the need to use flexible templates in order to remove the IA signal. We also study the environment dependence of intrinsic alignments by splitting the sample into field and group galaxies, which are further split into satellite and central galaxies. We show that group central galaxies are aligned with their halos at small scales and also are aligned with the tidal fields out to large scales. We also detect the radial alignments of satellite galaxies within groups, which results in a null detection of large-scale intrinsic alignments for satellites. These results can be used to construct better intrinsic alignment models for removal of this contaminant to the weak lensing signal.

## 5.1 Introduction

The deflection of light rays due to the gravitational effects of matter, gravitational lensing, changes both the observed shape and size of distant galaxies. Weak lensing, the statistical study of such tiny shape distortions, has emerged as an important tool in cosmology to study and map the distribution of dark matter as well as to study the effects of dark energy. Due to its sensitivity to the total matter content, weak lensing is theoretically a clean probe for dark matter and dark energy, but it is affected by a number of potential systematic sources (see, e.g., Weinberg et al., 2013). Much effort has been put into controlling the observational systematics arising from the effect of the point-spread function (PSF) on the estimated galaxy shape distortions. Future surveys like LSST[1] (LSST Science Collaboration et al., 2009), Euclid[2] (Laureijs et al., 2011) and WFIRST-AFTA [3] (Spergel et al., 2013) aim to bring these systematics below the percent level, with comparable improvement being needed on astrophysical sources of contamination to the weak lensing measurements.

Intrinsic alignments of galaxies is perhaps the most important astrophysical systematic for weak lensing. It is the coherent alignments of the shapes of physically nearby galaxies which can mimic the

[1]http://www.lsst.org/
[2]http://sci.esa.int/euclid/, http://www.euclid-ec.org
[3]http://wfirst.gsfc.nasa.gov/

weak lensing signal. One likely explanation for intrinsic alignments is that they are an environmental effect, whereby the local or large-scale tidal fields can shear and align the galaxy shapes, producing shape correlations. These correlations violate the assumption in weak lensing studies that observed correlations in galaxy shapes are only caused by gravitational lensing. Hirata & Seljak (2004) developed the linear alignment model for the intrinsic alignments, assuming that intrinsic alignments are set by tidal fields at the time of galaxy formation (Catelan et al., 2001). This model was later extended by Bridle & King (2007) to non-linear scales using the full non-linear matter power spectrum. To extend this model inside of dark matter halos, Schneider & Bridle (2010) developed the halo model of intrinsic alignments, which assumes that satellites are aligned radially within the halo, with central galaxies (BCGs or BGGs) being preferentially aligned with the shape of the halo. Constraining these models observationally is important for all techniques for mitigating intrinsic alignments in future weak lensing surveys that involve marginalizing over models with some free parameters (e.g., Joachimi & Bridle, 2010); these techniques require valid models, and priors on the model parameters.

Intrinsic alignments are also interesting from the galaxy formation perspective, as the alignments of galaxies with their dark matter halos and the alignments of dark matter halos themselves are intricately related to the processes involved in galaxy formation and evolution. Several studies have concentrated on studying intrinsic alignments of dark matter halos in cosmological simulations (for example, Croft & Metzler, 2000; Heymans et al., 2006; Faltenbacher et al., 2008; Kang et al., 2007; Kuhlen et al., 2007). These studies have found that dark matter halo shapes tend to align with each other, while subhalos tend to align radially within their host halos. Joachimi et al. (2013a,b) studied shapes and intrinsic alignments of COSMOS galaxies by populating dark matter halos in the Millennium simulation (Springel et al., 2005) with galaxies using semi-analytical methods, finding a strong dependence of intrinsic alignments on scale, redshift and galaxy type. Tenneti et al. (2014b) also studied the orientations of stellar components of galaxies within their halos using the MassiveBlack-II (MB-II) hydrodynamical simulation (Khandai et al., 2014), and found that galaxy shapes are preferentially aligned with their dark matter halos with average misalignments angle of $\sim 10 - 30$ degrees, with the degree of misalignment being a strong function of mass.

Observationally, several studies have detected the intrinsic alignments for Luminous Red Galaxies (LRGs) to tens of Mpc scales (Mandelbaum et al., 2006; Hirata et al., 2007; Okumura et al., 2009; Joachimi et al., 2011), finding the radial dependence of the signal to be consistent with the linear alignment model. Hirata et al. (2007) and Joachimi et al. (2011) also observed luminosity dependence of intrinsic alignments, with more luminous objects having stronger alignments. Many studies have also tried to measure intrinsic alignments within halos using large cluster samples (Sifón et al., 2014; Schneider et al., 2013; Hung & Ebeling, 2012; Hao et al., 2011), but so far, satellite alignments within groups and clusters have not been conclusively detected. Zhang et al. (2013) used the Sloan Digital Sky Survey (SDSS) DR7 data to reconstruct the cosmic tidal field, and observed that the major axis of galaxies tends to be aligned with the direction of the filament or the plane of the sheet in which they live. They also observed color, environment and luminosity dependence of this alignment, with central, red and brighter galaxies showing stronger alignments. Blazek et al. (2012) and Chisari et al. (2014) also developed methods to measure intrinsic alignments using more general photometric samples from SDSS. When using photometric redshifts, intrinsic alignments measurements suffer from contamination from weak lensing. After accounting for possible weak lensing contamination, Blazek et al. (2012) and Chisari et al. (2014) found the intrinsic alignments signals for typical galaxies to be consistent with null detection.

In this work, we will extend the study of intrinsic alignments in LRGs to more faint objects using the low redshift (LOWZ) spectroscopic sample of LRGs from the SDSS-III BOSS survey. The LOWZ sample covers the same redshift range as the LRG sample used by Hirata et al. (2007), but it extends to lower luminosities, with comoving number density $\sim 3 \times 10^{-4} h^{-3}\mathrm{Mpc}^3$, an increase by a factor of 3. We test the validity of the non-linear alignment model in both the linear and quasi-linear regimes using shape-density cross correlations for various samples defined using different galaxy properties. We study the environmental dependence of intrinsic alignments by identifying galaxies in groups and study the intrinsic alignments for satellites, BGGs (brightest group galaxies), and field galaxies. We

also model the small-scale signal using a halo model fitting function from Schneider & Bridle (2010), and study the variations in small-scale intrinsic alignment amplitude with different galaxy properties.

This paper is organized as follows: in section 5.2, we describe the linear alignment model of intrinsic alignments along with the halo model and the two point correlation functions we use to measure intrinsic alignments. Section 5.3 describes the SDSS-III BOSS LOWZ sample and selection criteria for the different sub-samples that we use. We present our results in section 5.4 and conclude in section 5.5.

Throughout we use standard cosmology with $h = 0.7$, $\Omega_b = 0.046$, $\Omega_{DM} = 0.236$, $\Omega_\Lambda = 0.718$, $n_s = 0.9646$, $\sigma_8 = 0.817$ (WMAP9, Hinshaw et al., 2013). All the distances are in comoving $h^{-1}$Mpc, though $h = 0.7$ as mentioned above was used to calculate absolute magnitudes and to generate predictions for the matter power spectrum.

## 5.2 Formalism and Methodology

We model galaxy shapes as some purely random shape, plus two small but coherent shape distortions ("shears"), one due to intrinsic alignments (e.g., from some large-scale tidal field) and the other due to gravitational lensing. The coherent part of the shears can be written as $\gamma = \gamma^G + \gamma^I$, where $\gamma^G$ is the lensing shear and $\gamma^I$ is the intrinsic shear. The two point correlation function for shear thus has the following contributions:

$$\langle\gamma\gamma\rangle = \langle\gamma^G\gamma^G + \gamma^G\gamma^I + \gamma^I\gamma^I\rangle = \xi_{GG} + \xi_{GI} + \xi_{II} \tag{5.1}$$

$\xi_{GG}$ is the desired lensing signal, while the $\xi_{II}$ contribution comes from two nearby galaxies that are affected by the same tidal field and $\xi_{GI}$ contribution comes from the pair where one galaxy shape is affected by the tidal field of the gravitational potential it lives in while a background galaxy is lensed by the same potential, see Hirata & Seljak (2004). In Sec. 5.2.1 we describe the linear alignment model used to model intrinsic alignments at large scales, followed by the halo model prescription in Sec. 5.2.2, which describes the intrinsic alignments at small scales.

### 5.2.1 Linear alignment model

One popular formalism to study intrinsic alignments at large scales is the linear alignment model (Catelan et al., 2001; Hirata & Seljak, 2004; Chisari & Dvorkin, 2013). In this section, we briefly describe the main features of the model.

The linear alignment model is based on the assumption that the intrinsic shear of galaxies is determined by the tidal field at the time of formation of the galaxy (assumed to be during matter domination; Catelan et al., 2001). Thus, we can write the intrinsic shear in terms of the primordial potential as

$$\gamma^I = (\gamma^I_+, \gamma^I_\times) = -\frac{C_1}{4\pi G}(\partial_x^2 - \partial_y^2, \partial_x\partial_y)\phi_p \tag{5.2}$$

Using the form of $\gamma^I$ described in Eq. (5.2), Hirata & Seljak (2004) derived the power spectrum for the two-point matter-intrinsic alignments correlation functions, relating them to the linear matter power spectrum, $P_\delta^{\rm lin}$.

$$P_{g+}(\mathbf{k}, z) = A_I b \frac{C_1\rho_{\rm crit}\Omega_m}{D(z)}\frac{k_x^2 - k_y^2}{k^2}P_\delta^{\rm lin}(\mathbf{k}, z) \tag{5.3}$$

$$P_{++}(\mathbf{k}, z) = \left(A_I \frac{C_1\rho_{\rm crit}\Omega_m}{D(z)}\frac{k_x^2 - k_y^2}{k^2}\right)^2 P_\delta^{\rm lin}(\mathbf{k}, z) \tag{5.4}$$

$$P_{g\times}(\mathbf{k}, z) = A_I b \frac{C_1\rho_{\rm crit}\Omega_m}{D(z)}\frac{k_x k_y}{k^2}P_\delta^{\rm lin}(\mathbf{k}, z) \tag{5.5}$$

Here, $P_{g+}$ ($P_{g\times}$) is the cross-power spectrum between the galaxy density field and the galaxy intrinsic shear along (at 45° from) the line connecting a pair of galaxies; $P_{++}$ is the shape- shape power spectrum for alignment along the line joining each pair of galaxies in the two-point correlations; and $b$ is the linear galaxy bias. Following Joachimi et al. (2011), we fix $C_1\rho_{\rm crit} = 0.0134$, and use an arbitrary constant $A_I$ to measure the amplitude of intrinsic alignments for different samples. $D(z)$ is the linear growth factor, normalized to unity at $z = 0$.

To measure the two-point correlation functions, we cross-correlate two samples of galaxies, with one sample used as a biased tracer of the matter density ("density sample", denoted $D$) and the other sample with shapes to trace the intrinsic shear ("shape sample", denoted $S$). See Sec. 5.2.3 for more details. The galaxy bias entering Eq. (5.3) is the bias of the density sample, while $A_I$ is measured for shape sample.

Bridle & King (2007) suggested using the full non-linear matter power spectrum, $P_\delta^{\rm nl}$, in Eqs. (5.3)–(5.5) to extend the linear alignment model to quasi-linear scales. This model is called the non-linear linear alignment (NLA) model. While there are a number of reasons why this substitution of the non-linear matter power spectrum is unlikely to be valid to very small scales (see discussion in Bridle & King, 2007), it should allow at least some extension to smaller scales, and use of this model will permit us to compare our results more easily with many other works that have also adopted this model. In this work, we will use the non-linear matter power spectrum based on halofit model (Smith et al., 2003), generated with software CAMB (Lewis & Bridle, 2002b), with fixed WMAP9 cosmology from Hinshaw et al. (2013). The expressions in Eqs. (5.3)–(5.5) can be Fourier transformed to get the 3d correlation functions:

$$\xi_{AB}(r_p,\Pi,z) = \int \frac{{\rm d}^2k_\perp {\rm d}k_z}{(2\pi)^3} P_{AB}(\mathbf{k},z)\left(1+\beta_A\mu^2\right) \left(1+\beta_B\mu^2\right) e^{i(r_p.k_\perp+\Pi k_z)} \tag{5.6}$$

Here $\mu = k_z/k$, and $\beta$ is the linear redshift distortion parameter[4] with $(1+\beta\mu^2)^2$ accounting for the effects of redshift-space distortions (RSD, see Appendix 5.A). For quantities that include a galaxy overdensity, the relevant $\beta$ is defined as $\beta(z) = f(z)/b$, where $f(z)$ is the logarithmic growth rate at redshift $z$ and $b$ is the galaxy bias of the sample. For the case of the shear, there is no RSD effect at leading order ($\beta_+ = 0$, see Appendix 5.A). In ΛCDM, for a wide range of redshifts, $f(z) \sim \Omega_m(z)^{0.55}$. From data, we most easily measure $w(r_p)$, the 2d (projected) correlation function, which can be obtained by integrating the 3d correlation function over the line of sight separation Π

$$w_{AB}(r_p) = \int {\rm d}z\, W(z) \int {\rm d}\Pi\, \xi_{AB}(r_p,\Pi,z) \tag{5.7}$$

Here $r_p$ is the 2d projected separation and $W(z)$ is the redshift window function (Mandelbaum et al., 2006).

$$W(z) = \frac{p_A(z)p_B(z)}{\chi^2(z){\rm d}\chi/{\rm d}z} \left[\int \frac{p_A(z)p_B(z)}{\chi^2(z){\rm d}\chi/{\rm d}z}{\rm d}z\right]^{-1} \tag{5.8}$$

$p_A(z)$ and $p_B(z)$ are the redshift probability distributions for shape and density sample respectively and $\chi(z)$ is the comoving distance to redshift $z$.

Doing the integral in Eq. (5.7), we get the 2d correlation functions:

$$w_{g+}(r_p) = \frac{A_I b_D C_1\rho_{\rm crit}\Omega_m}{\pi^2} \int {\rm d}z \frac{W(z)}{D(z)} \int_0^\infty {\rm d}k_z \int_0^\infty {\rm d}k_\perp \frac{k_\perp^3}{(k_\perp^2+k_z^2)k_z} P(\mathbf{k},z)\sin(k_z\Pi_{\rm max})J_2(k_\perp r_p) \left(1+\beta_D\mu^2\right) \tag{5.9}$$

[4] We do not explicitly account for the Fingers of God effect due to the virial motions of galaxies within halos. However, these are expected to be negligible in the projected correlations which are obtained by integrating along the line-of-sight between $\pm 100h^{-1}$Mpc.

where $b_D$ is the galaxy bias for density sample, while $\beta_D = f/b_D$. $w_{g\times}$ is expected to be zero by symmetry. The projected intrinsic alignments auto- correlation function is defined as

$$w_{++}(r_p) = \frac{(A_I C_1 \rho_{\rm crit} \Omega_m)^2}{2\pi^2} \int {\rm d}z \frac{W(z)}{D(z)^2} \int_0^\infty {\rm d}k_z \int_0^\infty {\rm d}k_\perp \frac{k_\perp^5}{(k_\perp^2 + k_z^2)^2 k_z} P(\mathbf{k}, z) \sin(k_z \Pi_{\rm max}) \times [J_0(k_\perp r_p) + J_4(k_\perp r_p)] \tag{5.10}$$

As discussed earlier, there is no RSD correction in $\gamma_I$ at leading order. Thus, the RSD correction[5] in $w_{g+}$ is only $(1 + \beta_D \mu^2)$ (from the galaxy overdensity $\delta_D$) and there is no RSD correction in $w_{++}$ at leading order (see Appendix 5.A for derivation).

To measure the bias of the density sample and of the shape sample, $b_S$, we measure the two-point galaxy-galaxy cross correlation function

$$w_{gg}(r_p) = \frac{b_S b_D}{\pi^2} \int {\rm d}z \frac{W(z)}{D(z)} \int_0^\infty {\rm d}k_z \int_0^\infty {\rm d}k_\perp \frac{k_\perp}{k_z} P(\mathbf{k}, z) \sin(k_z \Pi_{\rm max}) J_0(k_\perp r_p) \left(1 + \beta_S \mu^2\right) \left(1 + \beta_D \mu^2\right) \tag{5.11}$$

Intrinsic alignments measurements that do not use spectroscopic redshifts suffer from gravitational lensing as a potential contaminant (Joachimi et al., 2011; Blazek et al., 2012; Chisari et al., 2014). Since we use spectroscopic redshifts to select galaxy pairs, we expect negligible lensing contamination in our measurements from galaxy-galaxy lensing. Hence we do not include any contamination signals in our model.

### 5.2.2 Halo model

To extend this model of intrinsic alignments into scales comparable to or smaller than the sizes of dark matter halos, Schneider & Bridle (2010) proposed the halo model for intrinsic alignments. The halo model assumes that BCGs or BGGs are at the center of dark matter halos, with their shapes aligned with the host halos. Satellite galaxies tend to align radially with the major axis pointing towards the center of the halo. In practice there can be substantial contamination in small scale intrinsic alignments signal due to off-centering of BCGs, as well as misalignments between dark matter and baryons.

Schneider & Bridle (2010) calculated the 1-halo intrinsic alignments power spectrum using Monte Carlo simulations with the aforementioned complications, and they provide a fitting function for the halo model, given by

$$P^{1h}_{\delta,\gamma_I} = a_h \frac{(k/p_1)^2}{1 + (k/p_2)^{p_3}} \tag{5.12}$$

The parameters $p_i$ are given by

$$p_i = q_{i1} \exp\left(q_{i2} z^{q_{i3}}\right) \tag{5.13}$$

$z$ is the redshift and $q_{ij}$ are the parameters that we keep fixed to values described in Table 5.1. Our choice of $q_{i2}$ and $q_{i3}$ values are the same as in Schneider & Bridle (2010), while the $q_{i1}$ are different, chosen to fit the shape of $w_{g+}$ for LOWZ sample. As in Sec. 5.2.1, we do the Fourier transform and line-of-sight integration to get

$$w^{1h}_{g+} = \int {\rm d}z\, W(z) \int {\rm d}\Pi \int \frac{{\rm d}^2 k_\perp {\rm d}k_z}{(2\pi)^3} P^{1h}_{\delta,\gamma_I}(\mathbf{k}, z) e^{i(r_p . k_\perp + \Pi k_z)} \tag{5.14}$$

[5] This is different from equation C.1 in Blazek et al. (2011) and equation 12 in Chisari & Dvorkin (2013), which have an incorrect RSD factor in $w_{g+}$.

| Parameter Index | $q_{i1}$ | $q_{i2}$ | $q_{i3}$ |
|---|---|---|---|
| 1 | 0.005 | 5.909 | 0.3798 |
| 2 | 0.6 | 1.087 | 0.6655 |
| 3 | 3.1 | 0.1912 | 0.4368 |

Table 5.1: Halo model parameters used in Eq. (5.12).

When computing the correlation functions from data, we use $\Pi \in [-100, 100]h^{-1}$Mpc, which is much larger than a typical halo size. The $d\Pi$ integral can then be approximated as a delta function, and Eq. (5.14) can be written as

$$w_{g+}^{1h} = \int \mathrm{d}z W(z) \int \frac{\mathrm{d}^2 k_\perp}{(2\pi)^2} P_{\delta,\gamma_I}^{1h}(k_\perp, z) e^{i(r_p . k_\perp)} \tag{5.15}$$

We emphasize that in order to fully explain and interpret the small-scale intrinsic alignments, a full halo model description would require us to describe the halo occupation statistics of each sample including the Fingers of God effect, along with a model for radial dependence of intrinsic shear and misalignment angles. Some of the assumptions made by Schneider & Bridle (2010) when computing the power spectrum might not be valid in the context of some (or all) of our measurements. However, here we have a much simpler goal: we want to provide a simple fitting formula that approximately describes the observed intrinsic alignments of LOWZ galaxies on small scales. For this reason, we use the above formulae and fit for the amplitude parameter $a_h$ as a simple way to encapsulate the 2-point correlation functions below $\sim 1.5h^{-1}$Mpc in terms of one number. Thus, $a_h$ cannot be simply interpreted as a single quantity such as the amplitude of a satellite galaxy radial shear.

### 5.2.3 Correlation function estimators

In order to compute the intrinsic alignments signals, we cross-correlate the shapes of galaxies in the shape sample (S) with the positions of galaxies in the density sample (D). We use a generalized Landy-Szalay correlation function estimator (Landy & Szalay, 1993) to calculate the cross-correlations

$$\xi_{gg} = \frac{(S - R_S)(D - R_D)}{R_S R_D} = \frac{SD - R_S D - S R_D + R_S R_D}{R_S R_D}, \tag{5.16}$$

$$\xi_{g+} = \frac{S_+ D - S_+ R_D}{R_S R_D}, \tag{5.17}$$

$$\xi_{++} = \frac{S_+ S_+}{R_S R_S}. \tag{5.18}$$

$R_S$ and $R_D$ are sets of random points corresponding to the shape sample and the density sample, respectively. Here the terms involving shears for the galaxies are

$$S_+ X = \sum_{i \in S, j \in X} \gamma_+^{(i)} \langle j|i \rangle \tag{5.19}$$

and

$$S_+ S_+ = \sum_{i \in S, j \in S} \gamma_+^{(i)} \gamma_+^{(j)} \langle j|i \rangle. \tag{5.20}$$

$\gamma_{+,\times}$ measure the components of the shear along the line joining the pair of galaxies and at 45 degrees from that line, respectively. In our sign convention, positive $\gamma_+$ implies radial alignments, while negative $\gamma_+$ implies tangential alignments.

Finally, the projection uses summation over bins in $\Pi$,

$$w_{ab} = \int_{-\Pi_{\max}}^{\Pi_{\max}} \xi_{ab}(r_p, \Pi) \mathrm{d}\Pi. \tag{5.21}$$

We use $\Pi_{\rm max} = 100h^{-1}$Mpc, with $d\Pi = 10h^{-1}$Mpc. Our choice is different from that used in Mandelbaum et al. (2006), but our results and conclusions are not significantly affected by this difference, and when modeling the results we self-consistently include this choice in the model predictions. $w_{g\times}$, $w_{\times\times}$, and $w_{+\times}$ are defined in a similar fashion. Note that our correlation function estimator is different from those used by Joachimi et al. (2011). Compared to the estimator $DD/DR$, the Landy-Szalay estimator is less affected by geometrical effects from survey boundaries and holes, and it also minimizes the covariance matrix.

To get the error bars, we divide the sample (described in Sec. 5.3) into 100 jackknife regions of approximately equal area on the sky, and compute the cross-correlation functions by excluding 1 jackknife region at a time.

When the shape and density sample are the same, we fix the cosmology and then simultaneously fit $w_{gg}$ and $w_{g+}$ for the galaxy bias, $b$, and intrinsic alignments amplitude, $A_I$, for each jackknife sample. The final values for both parameters are the jackknife mean, with errors calculated from the jackknife variance. If the shape and density sample are different, the bias for the density sample ($b_D$) is measured separately from its auto-correlation function, and then is held fixed during the fitting procedure that uses the cross-correlation functions $w_{gg}$ and $w_{g+}$ to get the bias of the shape sample $b_S$ and its intrinsic alignments amplitude $A_I$. This procedure leads to slightly underestimated errors on our bias ($b_S$) measurements, but has a negligible effect on the intrinsic alignments amplitude, since the errors on the intrinsic alignments amplitude are dominated by shape noise. We also note that the size of our jackknife regions on the sky is approximately 8 degrees on a side, which corresponds to $\sim 65h^{-1}$Mpc at redshift of 0.16. We measure (and show) the correlation function out $\sim 200h^{-1}$Mpc, but we fit the models only for $r_p < 65h^{-1}$Mpc to avoid any issues resulting from the size of jackknife regions (most notably the error bars for $r_p > 65h^{-1}$Mpc will be underestimated). We test the effect of limiting our model to $r_p < 65h^{-1}$Mpc, by fitting out to $r_p \sim 200h^{-1}$Mpc and found that our results and conclusions remain unchanged.

### 5.2.4 Weak lensing

To study the possible halo mass dependence of intrinsic alignments, we estimate the halo mass of galaxies in our sample using galaxy-galaxy weak lensing, with our LOWZ galaxies as the lens sample. As described earlier, the observed shear of a galaxy has two contributions, $\gamma = \gamma^G + \gamma^I$, where $\gamma^G$ is the tangential shear from lensing and $\gamma^I$ is the intrinsic shear of the galaxy. Blazek et al. (2012) used the same SDSS source sample as used in this work to estimate the intrinsic alignments contamination in galaxy-galaxy lensing (with SDSS LRG sample as lenses) and found that the intrinsic alignments contamination is well below the statistical error for this source sample. Thus, for our lensing measurements we will assume that the contribution of the $\gamma^I$ term is well below the shape noise limit, and will ignore its contribution when modeling the lensing signal. In the weak lensing regime, the observable is then tangential shear $\gamma_t = \gamma_t^G$, which relates to the surface density contrast $\Delta\Sigma(r_p) = \bar{\Sigma}(< r_p) - \Sigma(r_p)$ as

$$\gamma_t(r_p) = \frac{\Delta\Sigma(r_p)}{\Sigma_c} \tag{5.22}$$

$\Sigma_c$ is the critical surface density, given in comoving coordinates as

$$\Sigma_c = \frac{c^2}{4\pi G}\frac{D_s}{(1+z_l)^2 D_l D_{ls}} \tag{5.23}$$

$z_l$ is the lens redshift; $D_s$ and $D_l$ are the angular diameter distances to the source galaxy and lens respectively while $D_{ls}$ is the angular diameter distance between the lens and source.

To measure the lensing signal, we compute the surface density contrast around a lens galaxy using the statistic

$$\Delta\Sigma(r_p) = \frac{\sum_{ls} w_{ls}\gamma_t^{(ls)}\Sigma_c^{(ls)}}{\sum_{rs} w_{rs}} \tag{5.24}$$

Here $w_{ls}$ is the weight for the lens-source pair,

$$w_{ls} = \frac{\Sigma_c^{-2}}{\sigma_\gamma^2 + \sigma_{SN}^2} \tag{5.25}$$

where the denominator includes the shape noise and measurement error for the galaxy shear estimates added in quadrature. $w_{rs}$ is similar to $w_{ls}$, but with random points acting as the lens sample. Thus dividing by $w_{rs}$ accounts for any galaxies that are physically-associated with the lens but are accidentally included in the source sample, which would lower the lensing signal if left uncorrected (Sheldon et al., 2004; Mandelbaum et al., 2005).

Since we are primarily interested in measuring the average mass of galaxies within our sample, we only model the lensing using an NFW profile (Navarro et al., 1996), for $r_p < 1h^{-1}$Mpc. The NFW profile is given by

$$\rho = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2} \tag{5.26}$$

with projected surface density given by

$$\Sigma(r_p) = \int_0^{r_{\rm vir}} \rho(r = \sqrt{r_p^2 + \chi^2}) d\chi. \tag{5.27}$$

For NFW profile, $\alpha = -1$ and $\beta = 3$ are fixed. We define the concentration, $c_{180b} = r_{180b}/r_s$, and mass, $M_{180b}$, using a spherical overdensity of 180 times the mean density:

$$M_{180b} = \frac{4\pi}{3} r_{180b}^3 (180\bar{\rho}_m) \tag{5.28}$$

Using the mass-concentration relation from Mandelbaum et al. (2008) in Eq. (5.29), we fit the signal for each jackknife sample in the range $0.05h^{-1}\text{Mpc} < r_p < 0.3h^{-1}\text{Mpc}$. We limit the model to $r_p < 0.3h^{-1}$Mpc to avoid contamination from the 1-halo satellite term and the halo-halo terms, which are important at larger $r_p$ values for galaxies in this mass range (see, e.g., figure 3 of Velander et al. 2014, which shows the contributions of these terms for a lower mass sample; all contributions tend to shift to the right for higher mass). The final quoted values for the mass are jackknife mean and standard error. When presenting mass results we will discuss sensitivity to this choice of mass-concentration relation:

$$c = 5 \left( \frac{M_{180b}}{10^{14} h^{-1} M_\odot} \right)^{-0.1} \tag{5.29}$$

## 5.3 Data

The SDSS (York et al., 2000) imaged roughly $\pi$ steradians of the sky, and followed up approximately one million of the detected objects spectroscopically (Eisenstein et al., 2001; Richards et al., 2002; Strauss et al., 2002). The imaging was carried out by drift-scanning the sky in photometric conditions (Hogg et al., 2001; Ivezić et al., 2004), in five bands (*ugriz*) (Fukugita et al., 1996; Smith et al., 2002) using a specially-designed wide-field camera (Gunn et al., 1998) on the SDSS Telescope (Gunn et al., 2006). These imaging data were used to create the catalogues of shear estimates that we use in this paper. All of the data were processed by completely automated pipelines that detect and measure photometric properties of objects, and astrometrically calibrate the data (Lupton et al., 2001b; Pier et al., 2003; Tucker et al., 2006). The SDSS-I/II imaging surveys were completed with a seventh data release (Abazajian et al., 2009), though this work will rely as well on an improved data reduction pipeline that was part of the eighth data release, from SDSS-III (Aihara et al., 2011); and an improved photometric calibration ('ubercalibration', Padmanabhan et al., 2008).

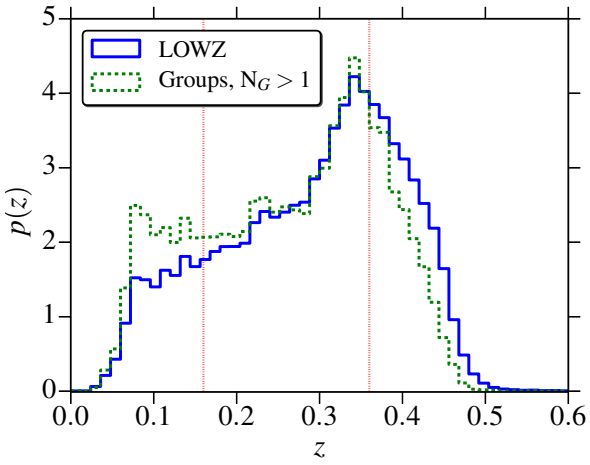



Figure 5.1: Redshift distribution of LOWZ galaxies and our group sub-sample. Vertical lines mark the boundary of the redshift range that we use in this paper, $z = [0.16, 0.36]$.

### 5.3.1 Redshifts

Based on the photometric catalog, galaxies are selected for spectroscopic observation (Dawson et al., 2013), and the BOSS spectroscopic survey was performed (Ahn et al., 2012) using the BOSS spectrographs (Smee et al., 2013). Targets are assigned to tiles of diameter 3° using an adaptive tiling algorithm (Blanton et al., 2003), and the data was processed by an automated spectral classification, redshift determination, and parameter measurement pipeline (Bolton et al., 2012).

We use SDSS-III BOSS DR11 LOWZ galaxies, in the redshift range $0.16 < z < 0.36$. The LOWZ sample consists of Luminous Red Galaxies (LRGs) at $z < 0.4$, selected from the SDSS DR8 imaging data and observed spectroscopically in BOSS survey. Fig. 5.1 shows the normalized redshift distribution of the LOWZ galaxies. The sample is approximately volume limited in the redshift range $0.16 < z < 0.36$, with a number density of $\bar{n} \sim 3 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$ (Manera et al., 2014). We combine the spectroscopic redshifts from BOSS with galaxy shape measurements from Reyes et al. (2012). BOSS DR11 has 225334 LOWZ galaxies within our redshift range. However, due to higher Galactic extinction in some regions of the sky, Reyes et al. (2012) did not have shape measurements in those regions. After cutting out those regions, we are left with 173855 galaxies for our LOWZ density sample, of which there are good shape measurements for 159621 galaxies, which constitutes our LOWZ shape sample.

To compute absolute magnitudes, we use distances in the units of $h^{-1}$Mpc for distance modulus computation and then apply $k + e$ corrections (to $z = 0$) as described in Wake et al. (2006). Fig. 5.2 shows the average $k + e$ corrected magnitudes as function of redshift. Throughout, we will use the $k + e$ corrected magnitudes.

To study the evolution of intrinsic alignments with redshift and different galaxy properties, we cut the LOWZ sample into different sub-samples. Table 5.2 describes the selection criteria and various properties of different sub-samples. Since our $k+e$ corrections are not perfect, we observe some redshift evolution in luminosity and color. To ensure that we select a fair samples based on luminosity and color cuts, those cuts are applied within redshift bins where each bin has 10 per cent of the sample (equal sized bins in redshift percentile space). We also observe evolution of color with luminosity but we avoid selecting color samples from luminosity bins, as that changes the redshift distribution ($\mathrm{d}n/\mathrm{d}z$) for different color samples, which complicates our correlation function calculations. For simplicity, we stick to redshift-based selection and keep the luminosity evolution in mind when interpreting the results for different color samples. Color and luminosity cuts are also based on percentile values. For example, the $L_1$ sample contains the galaxies in the $0 - 20^{th}$ percentile in luminosity within each $z$ bin. $L_1 - L_4$ samples go from brightest to faintest, with $L_1$–$L_3$ each containing 20% of the galaxies,

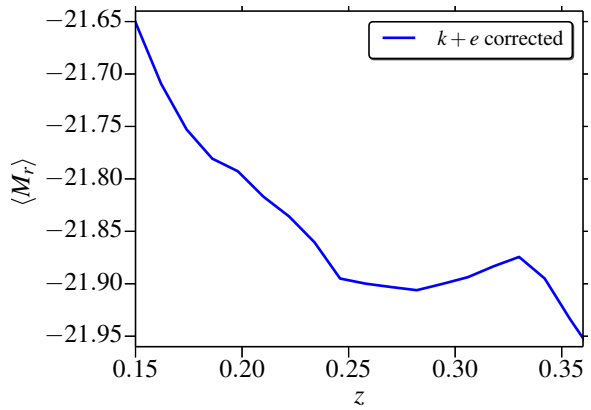



Figure 5.2: Average $k + e$ corrected $r$-band absolute magnitude, $\langle M_r \rangle$, for the LOWZ sample as function of redshift.

and $L_4$ having the faintest 40 per cent of the galaxies ($L_4$ has more galaxies to get sufficient S/N). Similarly each color sample has 20 per cent of the galaxies and go from blue to red end, with $C_1$ being the bluest and $C_5$ being the reddest.

Due to the finite size of the optical fibers required to target galaxies for spectroscopy, the SDSS is not able to simultaneously take spectra for galaxy pairs separated by $\theta < 62''$ on the sky. In regions of the sky which are observed multiple times (due to overlapping tiles), a number of such cases can be resolved. Nevertheless a significant fraction of targets remain unassigned. This effect is known as fiber collisions. Since fiber collisions occur preferentially in denser environments, the resulting incompleteness introduces a bias in correlation functions at all scales, with small scales being more strongly affected. One way to correct for this, is to upweight the nearest neighbour of a fiber collided galaxy. This is based on the assumption that two very nearby red sequence galaxies are likely to be within the same group or cluster. The nearest neighbour upweighting scheme has been shown to correct the bias in galaxy-galaxy correlation function on large scales ($\gtrsim$ fiber collision scale), but can introduce some bias at small scales since the redshift separation between the collided pair of galaxies has been artificially set to zero, which is not always correct (Reid et al., 2014). Also while measuring the clustering of galaxies in subsamples, the assumption that the fiber collided galaxy would also have been part of the same subsample as its nearest neighbour would be incorrect (see e.g., More et al., 2014, in the context of SDSS-III CMASS galaxies). Therefore, instead of upweighting the neighbours, we include the fiber collided galaxies in our sample, but assign them redshifts of their nearest neighbours. Nevertheless, bearing in mind sample incompleteness and fiber collision correction as possible source of systematics at small scales, we do not use the points below the fiber collision scale when fitting our models. Throughout, the fiber collision scale will be marked with solid black lines at $r_p \sim 300$ kpc/h (fiber collision scale at $z = 0.36$). In addition to the fiber collided galaxies, we also assign the nearest neighbour redshifts to galaxies for which the spectroscopic pipeline failed to obtain a redshift estimate. We calculate the absolute magnitudes along with $k + e$ corrections for these galaxies based on the nearest neighbour redshift estimates.

We use the sets of random points provided by the BOSS collaboration. However, some of our subsamples have different redshift distributions from the full LOWZ sample, in which case we generate new redshifts for the randoms from the subsample redshift distribution using the acceptance-rejection method. The acceptance-rejection method generates a set of random variates by uniformly sampling the area of the probability distribution of the random variable.

| Dataset | Cuts | $N_D$ | $N_S$ | $\langle z \rangle$ | $\langle L_r/L_r^p \rangle$ | $\langle M_g\text{-}M_i \rangle$ | $\log\left(\frac{M_{180b}}{h^{-1}M_\odot}\right)$ | $A_I$ | $b_S$ | $a_h$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Lowz | | 173855 | 159621 | 0.28 | 0.95 | 1.18 | 13.18±0.05 | 4.6±0.5 | 1.77±0.04 | 0.08±0.01 |
| BGG | $N_g > 1$, BGG=True | 19319 | 17916 | 0.27 | 1.31 | 1.16 | 13.5±0.1 | 8.1±1.5 | 2.54±0.13 | 0.41±0.05 |
| Group | $N_g > 1$ | 44780 | 40928 | 0.27 | 1.01 | 1.19 | 13.3±0.1 | 3.6±1.0 | 2.67±0.14 | 0.26±0.03 |
| Satellite | $N_g > 1$, BGG=False | 25461 | 23012 | 0.27 | 0.78 | 1.22 | 13.1±0.1 | 0±1 | 2.77±0.15 | 0.16±0.04 |
| Field | $N_g = 1$ | 129075 | 118693 | 0.28 | 0.93 | 1.18 | 13.13±0.06 | 5.0±0.6 | 1.46±0.07 | 0.011±0.007 |
| Z1 | $0.16 < z < 0.26$ | 67880 | 63180 | 0.21 | 0.91 | 1.17 | 13.17±0.06 | 4.1±0.8 | 1.66±0.07 | 0.05±0.01 |
| Z2 | $0.26 < z < 0.36$ | 105975 | 96441 | 0.32 | 0.98 | 1.19 | 13.18±0.07 | 5.1±0.8 | 1.88±0.05 | 0.16±0.02 |
| $L_1$ | $0\% < M_r < 20\%$ | 34760 | 31910 | 0.28 | 1.55 | 1.14 | 13.49±0.07 | 8.5±0.9 | 2.0±0.1 | 0.20±0.02 |
| $L_2$ | $20\% < M_r < 40\%$ | 34768 | 31910 | 0.28 | 1.04 | 1.16 | 13.25±0.09 | 5±1 | 1.74±0.08 | 0.09±0.02 |
| $L_3$ | $40\% < M_r < 60\%$ | 34768 | 31910 | 0.28 | 0.87 | 1.17 | 13.1±0.1 | 4.7±1.0 | 1.67±0.09 | 0.06±0.02 |
| $L_4$ | $60\% < M_r < 100\%$ | 69530 | 63830 | 0.28 | 0.65 | 1.23 | 12.96±0.09 | 2.2±0.9 | 1.70±0.08 | 0.03±0.02 |
| $C_1$ | $0\% < M_g - M_i < 20\%$ | 34760 | 31910 | 0.28 | 1.16 | 1.08 | 13.0±0.1 | 4.6±1.1 | 1.52±0.09 | 0.07±0.02 |
| $C_2$ | $20\% < M_g - M_i < 40\%$ | 34768 | 31910 | 0.28 | 0.99 | 1.16 | 13.24±0.10 | 5.0±1.0 | 1.72±0.09 | 0.06±0.02 |
| $C_3$ | $40\% < M_g - M_i < 60\%$ | 34768 | 31910 | 0.28 | 0.96 | 1.19 | 13.22±0.09 | 5.4±0.9 | 1.84±0.10 | 0.10±0.02 |
| $C_4$ | $60\% < M_g - M_i < 80\%$ | 34768 | 31910 | 0.28 | 0.92 | 1.21 | 13.36±0.09 | 5.8±1.0 | 1.95±0.10 | 0.10±0.02 |
| $C_5$ | $80\% < M_g - M_i < 100\%$ | 34760 | 31910 | 0.28 | 0.73 | 1.29 | 13.0±0.1 | 2±1 | 1.8±0.1 | 0.079±0.025 |

Table 5.2: Table describing various sub-samples in our analysis. 'Cuts' describes the cuts implemented in our pipeline to select the sample, where $N_g$ is the group multiplicity. $N_D$ is the total number of galaxies in the sample and $N_S$ is the number of galaxies in that sample with good shape measurements. $\langle z \rangle$ and $\langle L_r/L_r^p \rangle$ are the average redshift and $r$-band absolute luminosity for the sample, with $L_r^p$ being the pivot luminosity corresponding to r-band absolute magnitude $M_r^p = -22.0$. $M_{180b}$ is the average halo mass for the sample, measured from weak lensing. % denotes the percentile values for the sample. $b_S$ is the linear galaxy bias, and $A_I$ and $a_h$ are the intrinsic alignments model amplitude for the sample, calculated by cross-correlating the given shape sample with the full LOWZ sample as density sample, with its bias $b_D = 1.77$ measured from LOWZ-LOWZ correlations and then held fixed. For joint fitting $w_{gg}$ and $w_{g+}$, we get reduced $\chi^2 \in [0.8, 2]$ for all samples presented here with a probability to

### 5.3.2 Shapes

The catalogue of galaxies with measured shapes used in this paper (described in Reyes et al. 2012 and further characterized in Mandelbaum et al. 2013) was generated using the re-Gaussianization method (Hirata & Seljak, 2003) of correcting for the effects of the point-spread function (PSF) on the observed galaxy shapes. The catalogue production procedure was described in detail in previous work, so we describe it only briefly here. Galaxies were selected in a 9243 deg$^2$ region, with an average number density of 1.2 arcmin$^{-2}$. The selection was based on cuts on the imaging quality, data reduction quality, galactic extinction $A_r < 0.2$ defined using the dust maps from Schlegel et al. (1998) and the extinction-to-reddening ratios from Stoughton et al. (2002), apparent magnitude (extinction-corrected $r < 21.8$), and galaxy size compared to the PSF. The apparent magnitude cut used model magnitudes[6]. For comparing the galaxy size to that of the PSF, we use the resolution factor $R_2$ which is defined using the trace of the moment matrix of the PSF $T_{\rm P}$ and of the observed (PSF-convolved) galaxy image $T_{\rm I}$ as

$$R_2 = 1 - \frac{T_{\rm P}}{T_{\rm I}}. \tag{5.30}$$

We require $R_2 > 1/3$ in both $r$ and $i$ bands.

The software pipeline used to create this catalogue obtains galaxy images in the $r$ and $i$ filters from the SDSS 'atlas images' (Stoughton et al., 2002). The basic principle of shear measurement using these images is to fit a Gaussian profile with elliptical isophotes to the image, and define the components of the ellipticity

$$(e_+, e_\times) = \frac{1-(b/a)^2}{1+(b/a)^2}(\cos 2\phi, \sin 2\phi), \tag{5.31}$$

where $b/a$ is the axis ratio and $\phi$ is the position angle of the major axis. The ellipticity is then an estimator for the shear,

$$(\gamma_+, \gamma_\times) = \frac{1}{2\mathcal{R}}\langle(e_+, e_\times)\rangle, \tag{5.32}$$

where $\mathcal{R} \approx 0.87$ is called the 'shear responsivity' and represents the response of the distortion to a small shear (Kaiser et al., 1995; Bernstein & Jarvis, 2002); $\mathcal{R} \approx 1 - e^2_{\rm rms}$. In the course of the re-Gaussianization PSF-correction method, corrections are applied to account for non-Gaussianity of both the PSF and the galaxy surface brightness profiles (Hirata & Seljak, 2003).

For this work, we do not use the entire source catalogue, only the portion overlapping the LOWZ sample.

When computing the intrinsic alignment correlation functions, we use the shear estimates from this catalog together with the redshifts from BOSS. However, when computing the weak lensing signals around the LOWZ galaxies, we need estimates of the redshifts for the fainter source galaxies. For this purpose, we use the maximum-likelihood estimates of photometric redshifts (photo-$z$) based on the five-band photometry from the Zurich Extragalactic Bayesian Redshift Analyzer (ZEBRA, Feldmann et al., 2006), which were characterised by Nakajima et al. (2012) and Reyes et al. (2012). In this work, we used a fair calibration sample of source galaxies with spectroscopic redshifts as defined in Nakajima et al. (2012) to calculate biases in weak lensing signals due to bias and scatter in the photo-$z$, and applied corrections that were of order 10 per cent ($\pm 2$ per cent) to the weak lensing signals.

### 5.3.3 Groups

We use the counts in cylinder (CiC) technique (Reid & Spergel, 2009) to find the galaxies in groups. The CiC technique assumes that for galaxies in the same halos, the dominant source of redshift separation between satellite and central galaxies are the line of sight relative velocities. Under such an assumption, a cylinder with size determined by the size and velocity dispersion of the halo should be

[6] http://www.sdss3.org/dr8/algorithms/magnitudes.php#mag_model

| $N_g$ | $N_G$ | $f$ | $N_g$ | $N_G$ | $f$ | $N_g$ | $N_G$ | $f$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 130414 | 0.75 | 6 | 84 | 0.0029 | 11 | 1 | $6.3\times10^{-5}$ |
| 2 | 14944 | 0.17 | 7 | 33 | 0.0013 | 12 | 1 | $6.9\times10^{-5}$ |
| 3 | 2784 | 0.048 | 8 | 17 | 0.00078 | 13 | 0 | 0 |
| 4 | 780 | 0.018 | 9 | 7 | 0.00036 | 14 | 1 | $8.1\times10^{-5}$ |
| 5 | 215 | 0.0062 | 10 | 2 | 0.00012 | 15 | 1 | $8.6\times10^{-5}$ |

Table 5.3: Group multiplicity function for the group sample in redshift range, $0.16 < z < 0.36$. $N_g$ is the number of galaxies within a group, $N_G$ is the total number of groups with $N_g$ galaxies, and $f$ is the fraction of the sample at a given group multiplicity $N_g$.

able to identify galaxy pairs within a single halo. Reid & Spergel (2009) calibrated the CiC technique for halos hosting LRGs using mock LRG catalogs and we use their dimensions for the cylinders, given by $R_\perp < 0.8h^{-1}$Mpc and $|\Pi_\parallel| < 20h^{-1}$Mpc in comoving coordinates. Once within-halo pairs are identified, we can use the friends-of-friends algorithm to find all the galaxies belonging to the same group.

We note that the CiC parameters were optimized by Reid & Spergel (2009) for the SDSS LRG sample, and the LOWZ sample has a number density that is three times higher, extending to lower mass halos. Thus, using the CiC parameters optimized for LRG sample can potentially lead to spurious group identification in our LOWZ group sample. To roughly quantify the level of contamination, we use the halo mass function of Tinker et al. (2008) and find that the lower limit of halo mass for a sample of halos with the LOWZ sample abundance is a factor of $\sim 2.5$ less than a sample with the abundance of the LRG sample. Since the cylinder parameters should scale roughly like $M^{1/3}$, this will give $R_\perp \lesssim 0.6h^{-1}$Mpc and $|\Pi_\parallel| \lesssim 15h^{-1}$Mpc for the LOWZ sample. Using these parameters, we find that the satellite fraction in LOWZ reduces to 11.7 per cent, versus 14.6 per cent using the CiC parameters from Reid & Spergel (2009). Most of this difference is coming from the reduction in the $R_\perp$ size, since most of the satellites are within $|\Pi_\parallel| \lesssim 10h^{-1}$Mpc as shown in Fig. 5.3. However, our calculation of CiC parameters for LOWZ is only approximate, as the CiC technique needs to be calibrated using mock catalogs, which is beyond the scope of this work. Hence for our analysis we will use the group catalog obtained using CiC parameters of Reid & Spergel (2009).

Table 5.3 shows the multiplicity function of our group sample and Fig. 5.3 shows the stacked distribution of satellites with respect to BGGs, which we assume as the centre of groups. BGGs are selected as the brightest galaxy (in $r$ band) within the group. The assumption that the brightest galaxy truly resides at the center of the halo fails for about 25%-40% of the groups within our mass range (Skibba et al., 2011; Hikage et al., 2013), in which case the brightest galaxy is in fact a satellite galaxy. This is a potential source of contamination that should be kept in mind when interpreting our results on the environmental dependence of intrinsic alignments.

## 5.4 Results

In this sections we present the results, beginning with correlation functions for full LOWZ sample and some tests for systematics. After that we primarily focus on evolution of intrinsic alignments amplitudes with different galaxy properties and end the section with some comparisons with other studies of intrinsic alignments, using different intrinsic alignments estimators where necessary. A summary of most of the results discussed in this section is presented in table 5.2 and Fig. 5.8.

### 5.4.1 LOWZ sample

We begin by showing our measurements for the full LOWZ sample. Fig. 5.4 shows our results for the galaxy-galaxy correlation function $w_{gg}$ and the density-shape correlation function, $w_{g+}$. Throughout, when referring to two point correlation measurements, our naming convention for different results

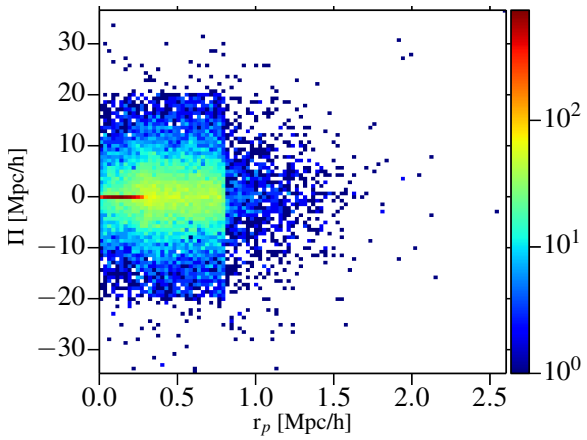



Figure 5.3: The stacked distribution of satellite galaxies in our group sample, with respect to BGGs in the groups. The apparent sharp boundary at $r_p = 0.8h^{-1}$Mpc is due to the size of our CiC cylinders and the fact that our sample is dominated by two-galaxy (single satellite) groups. The strong peak at $\Pi = 0$ and $r_p < 0.3h^{-1}$Mpc arises primarily from the fiber collision corrected pairs.

will be in the form of "Shape sample-Density sample". We fit these data to the following models: the galaxy-galaxy correlation function calculated using non-linear matter power spectrum and linear galaxy bias, and the NLA model at $r_p > 6h^{-1}$Mpc. For the full LOWZ sample, we find a bias $b = 1.77 \pm 0.04$ and intrinsic alignment amplitude $A_I = 4.6 \pm 0.5$. The NLA model is a good fit to the $w_{g+}$ measurement for $r_p > 6h^{-1}$Mpc, and can be extended to $r_p \sim 4h^{-1}$Mpc, but there are significant deviations below that scale, due to differing scale dependence.

Our bias measurement appears to be in tension with results from Parejko et al. (2013), who measured $b \sim 2$ for the BOSS DR9 LOWZ sample. However, there are a number of differences betweeen our study and theirs: Parejko et al. (2013) (i) use a different redshift range, $z \in [0.2, 0.4]$, (ii) use a smaller area DR9 sample, (iii) neglect redshift space distortions which can affect the large scale bias (see e.g., van den Bosch et al., 2013), and (iv) use a FKP weighting scheme which favourably weights the large bias high redshift galaxies. Using the same redshift range as theirs, we also get a higher value of bias $b = 1.85 \pm 0.04$. When RSD corrections in our model are neglected, we obtain an even higher value of $b = 1.90 \pm 0.04$. We expect the remaining discrepancy to be a result of the further differences in the weighting scheme and/or the sample size. Our values are closer to the $b = 1.85$ used by Tojeiro et al. (2014) in the reconstruction scheme for the BAO measurement (Tojeiro et al. 2014 use $\sigma_8 = 0.8$ and $n_s = 0.95$).

On non-linear scales, we use the halo model fitting formula for intrinsic alignments defined in Sec. 5.2.2, fitting to the data with $0.3 < r_p < 1.5h^{-1}$Mpc. We find a good agreement with the data using the halo model parameters as described in Table 5.1, with halo model amplitude $a_h = 0.084 \pm 0.010$. Keeping $a_h$ fixed, but using fitting formula parameters from Schneider & Bridle (2010), the model does not fit data well. This is not surprising as the halo model fitting function does not explicitly model effects such as scale dependence of non-linear bias and intrinsic alignments themselves, which can change both the amplitude and scale dependence of the signal. For this reason and those described in Sec. 5.2.2, we caution against interpreting our $a_h$ values directly as an average intrinsic alignments shear.

The halo model fitting formula combined with the LA model (linear alignment model with linear matter power spectrum) does not fit the data well in the intermediate regime $2 < r_p < 10h^{-1}$Mpc. The NLA model combined with the halo model fitting function can fit the data well at all scales, but the interpretation of such a model is not clear, since both the NLA model and the halo model fitting function assume different things in modeling the 1-halo term. For this reason, we do not show the combined fits for the NLA and halo models, but simply use the halo model fitting function as a

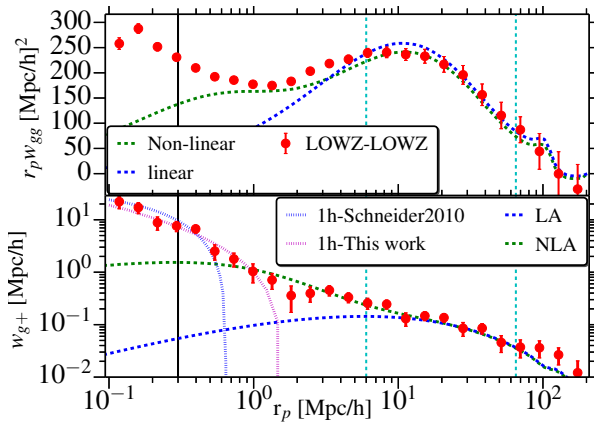



Figure 5.4: The galaxy-galaxy correlation, $w_{gg}$ (top, multiplied by $r_p$), and the density-shape correlation function, $w_{g+}$ (bottom), for the full LOWZ sample in the redshift range $0.16 < z < 0.36$. The red points are the measurements from the data, the dashed green lines are the linear model with the non-linear matter power spectrum, and the dashed blue lines are the linear model with the linear matter power spectrum. The NLA model is fitted only in the range $6h^{-1}$Mpc $< r_p < 65h^{-1}$Mpc (shown by a dashed vertical line), while the LA model is shown with the same parameters as the NLA model. The dotted red line in the bottom plot is the halo model fit to $w_{g+}$ at small scales. The black solid line shows the SDSS fiber collision scale at $z = 0.36$.

convenient way to encapsulate the amplitude of the small-scale signal in a single number.

Fig. 5.5 shows our results for the $w_{++}$ measurement. The NLA model with the best-fitting parameters from $w_{g+}$ predicts a small but non-zero $w_{++}$. Our signal is dominated by shape noise, and our results (at $6h^{-1}$Mpc $< r_p < 65h^{-1}$Mpc) marginally prefer the null model (best fit $A_I = 2 \pm 96$, $\chi^2 = 7.21$ with 8 bins, $\chi^2_0 = 7.22$,) over the NLA prediction using the amplitude from the fits to the $w_{g+}$ signal ($\chi^2 = 8.0$) with a $\Delta\chi^2 = 0.8$, which is not a statistically significant difference. Theoretically, from the NLA model with the best-fitting value of $C_1$ from the fits to $w_{g+}$, we expect $w_{++}/w_{g+} \sim 0.01$ (the ratio is scale dependent with a peak value of $\sim 0.01$ for $r_p > 6h^{-1}$Mpc). In the shape noise-dominated limit, in the case that the shape and density sample are the same, the standard deviations of $w_{g+}$ and $w_{++}$ are proportional to $\sigma_\gamma$ and $\sigma^2_\gamma$ (see, e.g., Schneider et al., 2002), respectively, where $\sigma_\gamma$ is the shape noise per component for our sample. For this sample, $\sigma_\gamma \sim 0.2$. Given a detection signal-to-noise ratio (SN) for $w_{g+}$ on large scales of 9.2 ($= A_I/\sigma_{A_I}$), the expected SN for $w_{++}$ on large scales is then $\mathrm{SN}(w_{++}) \sim 0.01\,\mathrm{SN}(w_{g+})/0.2 \sim 0.5$. Thus at $r_p > 6h^{-1}$Mpc we expect to have a null detection of $w_{++}$, consistent with our observations. If we use scales below $6h^{-1}$Mpc, the $\chi^2 = 16.7$ for a null signal[7], with 9 bins, giving a $p$-value$= 0.1$ ($< 2\sigma$).

Our $w_{++}$ results are inconsistent with those for SDSS-I/II LRGs from Blazek et al. (2011) ($\Delta\chi^2 \sim 11$, $p$-value$= 0.004$), as shown in Fig. 5.5. While calculating the projected correlation function from the 3d correlation functions in Okumura et al. (2009), Blazek et al. (2011) assumed isotropic $\xi_{++}$ along with Gaussian and independent errors. These assumptions are likely to break down at some level, and can cause a potential discrepancy with our results.

Other possible sources of the discrepancy are as follows. First, the LOWZ sample is fainter than the LRG sample in Blazek et al. (2011). Since the $w_{++}$ signal goes as $A_I^2$, our expected signal is lower by a factor of $\sim 4$ ($A_I^{\mathrm{LRG}} = 9.3$, $A_I^{\mathrm{LOWZ}} = 4.6$). Our $L_1$ sub-sample is closer to the LRG sample in term of luminosity and the number of galaxies. However, as shown in Fig. 5.5, the large scale $w_{++}$ signal for the $L_1$ sample is also consistent with a null signal (for $6h^{-1}$Mpc $< r_p < 65h^{-1}$Mpc, the best fit $A_I = 0 \pm 44$ with $\chi^2 = 6.3$ for 8 bins). The $\Delta\chi^2$ when we compare this with the $\chi^2$ value for

[7]All $p$-values in this paper calculated from either $\chi^2$ or $\Delta\chi^2$ use the simulation method described in Hirata et al. (2004) to account for the fact that the jackknife covariances are noisy, which modifies the expected $\chi^2$ distribution.

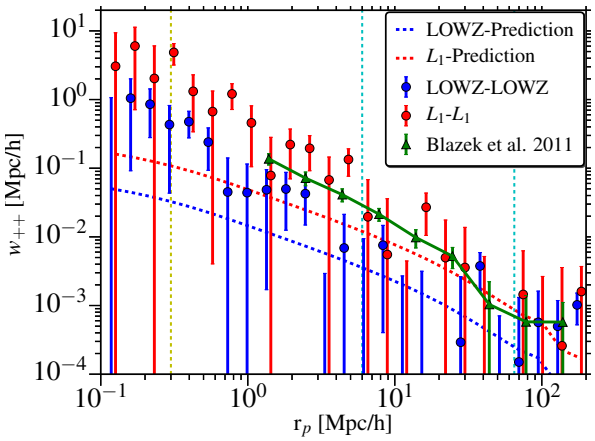



Figure 5.5: $w_{++}$ measurements for the full LOWZ sample (blue) and the $L_1$ sample (red). The green points show the $w_{++}$ measurement from Blazek et al. (2011). Our results are dominated by shape noise, and both the LOWZ and $L_1$ sample are consistent with a null detection at large scales. The dotted lines show NLA model predictions (using the best fit model to $w_{g+}$). We note that NLA fitting to $w_{g+}$ is only done for $6h^{-1}\text{Mpc} < r_p < 65h^{-1}\text{Mpc}$.

the value $A_I = 9.22$ from Blazek et al. (2011) is only $\sim 0.5$. Hence our $L_1$ sample results do not rule out the Blazek et al. (2011) results; however, what is interesting is the very large discrepancy in the significance of detection between the two studies (with Blazek et al. 2011 ruling out the null model at high significance).

A possible explanation for this discrepancy in detection significance comes from the use of different per-galaxy shear estimates. The results in Okumura et al. (2009) that were used to present $w_{++}$ in Blazek et al. (2011) used isophotal shape measurements from the SDSS pipeline which are defined at a low surface brightness level and are not corrected for the effects of the PSF. PSF contamination and other systematics can introduce spurious shape correlations which can mimic a $w_{++}$ signal. However, it is still possible that we are seeing a real physical effect. The SDSS isophotal shape measurements emphasize the outer parts of galaxies, while re-Gaussianization puts more weight on the inner regions of galaxies. If the outer regions of galaxy shapes are more responsive to tidal fields, this could result in legitimately different $w_{++}$ signals when using the two different shear estimates. Using the MB-II SPH simulation, Tenneti et al. (2014a) found that differences in shape measurement methods, weighted towards outer or inner regions, leads to a difference of about 10% in the $w_{g+}$ measurement[8], which would translate to a $\sim 20\%$ difference in $w_{++}$ (with the right sign to explain what is seen here: $w_{g+}$ is larger when weighting the outer regions more heavily). However, shape measurement methods in simulations may not correspond directly to shape measurement methods in observations and hence there may not be a direct correspondence between results from Tenneti et al. (2014a) and discrepancies observed between our results and that of Blazek et al. (2011). We also note that Blazek et al. (2011) detected $w_{\times\times}$ signal as well, but our $w_{\times\times}$ signal (not shown), like our $w_{++}$ signal, is consistent with zero. A definitive answer to this puzzle requires deeper analysis of shape measurement methods and their impact on IA measurement, which we defer to future work. Due to the null detection of $w_{++}$ in the LOWZ sample, we will not discuss $w_{++}$ hereafter.

[8]The trend is mass-dependent, and the difference is larger for lower mass. The number we quote here is for halo masses $M > 10^{13}h^{-1}M_\odot$, which are the most appropriate ones for a comparison with massive galaxies like those in the LOWZ sample.

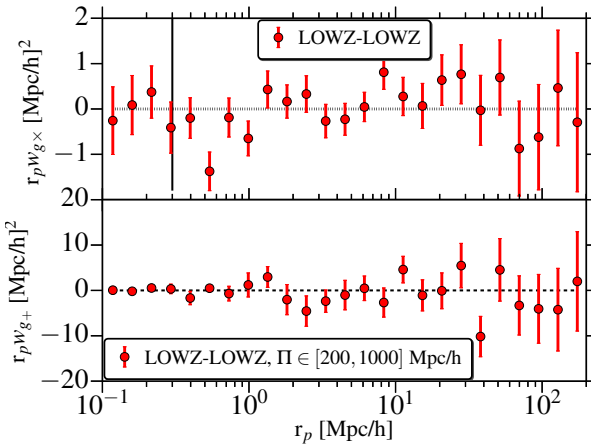



Figure 5.6: $w_{g\times}$ and $w_{g+}$ (at large $|\Pi|$) measurements for the full LOWZ sample. Both signal are consistent with zero. Note that both quantities are multiplied by $r_p$ in this plot. The increase in the statistical errors on large scales is driven by large-scale shear systematics; see Mandelbaum et al. (2013) for details.

### 5.4.2 Systematics tests

In this section, we present a $w_{g\times}$ measurement for the full LOWZ sample, and a $w_{g+}$ measurement when integrating only over large $\Pi \in [200, 1000]h^{-1}$Mpc values. Both these measurements are a good test for systematics in our shape measurements and correlation function calculations. From symmetry, $w_{g\times}$ is expected to be zero, while $w_{g+}$ is also expected to have null signal when only large $\Pi$ values are considered since the galaxy pairs are too far apart along the line-of-sight. As shown in Fig. 5.6, the $w_{g\times}$ signal is consistent with zero, with $\chi^2_{\rm red} = 1.48$ for 25 bins (across all scales shown on the plot) and $p$-value$= 0.34$. $w_{g+}$ at large $|\Pi|$ values is also consistent with zero, with $\chi^2_{\rm red} = 1.41$ and $p$-value$= 0.40$. We therefore conclude that these signals, which could reveal possible systematic errors, are fully consistent with zero. There are also no patterns evident in the residuals from zero, just a slight hint that errorbars might be underestimated.

### 5.4.3 Luminosity Dependence

Previous studies using LRGs have observed luminosity dependence of intrinsic alignments (Hirata et al., 2007; Joachimi et al., 2011), where brighter galaxies tend to have higher intrinsic alignments amplitude. LOWZ sample allows us to study luminosity dependence by going to fainter luminosities, with spectroscopic redshifts for all galaxies. To study luminosity dependence, we divide the LOWZ sample into four sub-samples, based on $r$-band absolute magnitude $M_r$. The sample are defined according $M_r$ percentiles within $z$ bins (each $z$ bin has 10% of the sample). $L_1$ has the brightest 20% of the galaxies, followed by $L_2$ and $L_3$ samples which have next 20% galaxies each based on $M_r$. $L_4$ has the faints 40% of the galaxies. Due to the decrease in signal to noise at fainter end, we do not cut $L_4$ sample into more sub-samples. Our density-shape correlation function measurements for the luminosity samples are shown in Fig. 5.7. All four samples are cross correlated with LOWZ as the density sample, and the LOWZ bias is fixed to $b_D = 1.77$ (Sec. 5.4.1) when modeling the results. Fits to $w_{gg}$ (not shown) reveal that the $L_1$ sample has the highest bias, followed by the $L_2$, $L_3$ and $L_4$ samples though differences in last three samples are statistically not very significant(see Table 5.2 for values of fit parameters like the bias). However, the overall trend in the evolution of bias with luminosity is consistent with the fact that brighter galaxies live in more massive and hence more biased halos.

We also see luminosity variation of the intrinsic alignments amplitude and therefore of the NLA

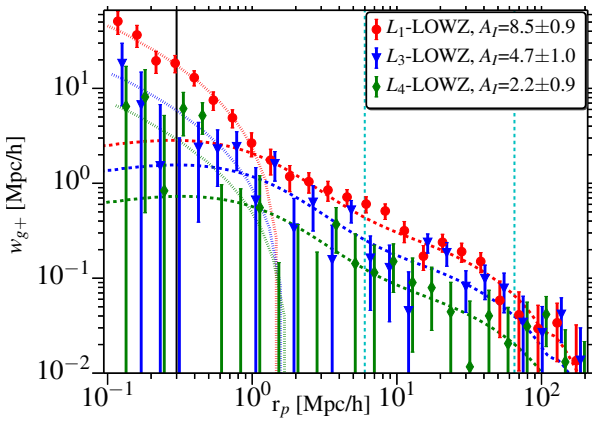



Figure 5.7: Density-shape correlation functions for samples defined using luminosity cuts, $L_1$, $L_3$, $L_4$ samples, as specified in Table 5.2. $L_1 - L_4$ are arranged in order of luminosity, with $L_1$ being brightest and $L_4$ being faintest. The $L_2$ sample is not shown in the figure for clarity. Brighter galaxies have higher intrinsic alignments amplitude, with luminosity trend of $A_I$ being well described by a power law (Eq. 5.33).

model best-fitting amplitude $A_I$, with brighter galaxies having higher amplitude. Since the density sample is the same in each case, the increasing $w_{g+}$ amplitude reflects the different intrinsic alignments for the different luminosity samples. Following Joachimi et al. (2011), we use a power law function to study the variation of $A_I$ with luminosity.

$$A_I(L_r) = \alpha \left( \frac{L_r}{L_r^p} \right)^{\beta} \tag{5.33}$$

$L_r$ and $M_r$ are the $r$-band luminosity and absolute magnitude, respectively, and $L_r^p(M_r^p)$ is a pivot luminosity (magnitude), chosen to be $M_r^p = -22.0$. Fitting to the samples defined by our luminosity cuts, we find $\alpha = 4.9 \pm 0.6$ and slope $\beta = 1.30 \pm 0.27$ (where these errors come from the jackknife). Our results are quite consistent with those of Joachimi et al. (2011), who found $\alpha = 5.76^{+0.60}_{-0.62}$ and $\beta = 1.13^{+0.25}_{-0.20}$ (using MegaZ-LRG + SDSS LRG + L4 + L3), with the same k-corrections and choice of $M_r^p$.

On small scales, we also see luminosity evolution of the halo model fitting function amplitude $a_h$, with brighter galaxies having higher amplitude. Fitting a power law similar to Eq. (5.33), we find a power-law amplitude $\alpha_h = 0.081 \pm 0.012$, and a power-law index $\beta_h = 2.1 \pm 0.4$ (subscript "h" denotes fits to halo model fitting function parameters). We note that scaling of $a_h$ with luminosity is somewhat steeper than that of $A_I$. This difference is likely due to that fact that more luminous galaxies live in denser regions, where tidal fields are stronger and hence leads to stronger intrinsic alignments. This will increase $w_{g+}$ at all scales, but the impact will be much stronger at smaller scales, leading to stronger scaling of $a_h$. This effect also contributes to the change in shape of $w_{g+}$ for $0.3 < r_p < 10h^{-1}$Mpc as the luminosity varies.

### 5.4.4 Redshift Dependence

To study the redshift dependence of intrinsic alignments, we divide our sample into two redshift bins, Z1: $0.16 < z < 0.26$, Z2: $0.26 < z < 0.36$. Due to the small redshift range of the LOWZ sample, we cannot cut our samples further, as the edge effects from the redshift boundaries can then introduce significant bias in the correlation functions.

Fig. 5.9 shows the density-shape correlation measurements for both of the redshift samples. From $w_{gg}$ measurement (not shown), we find higher galaxy bias for Z2 sample ($b = 1.88 \pm 0.05$) compared to

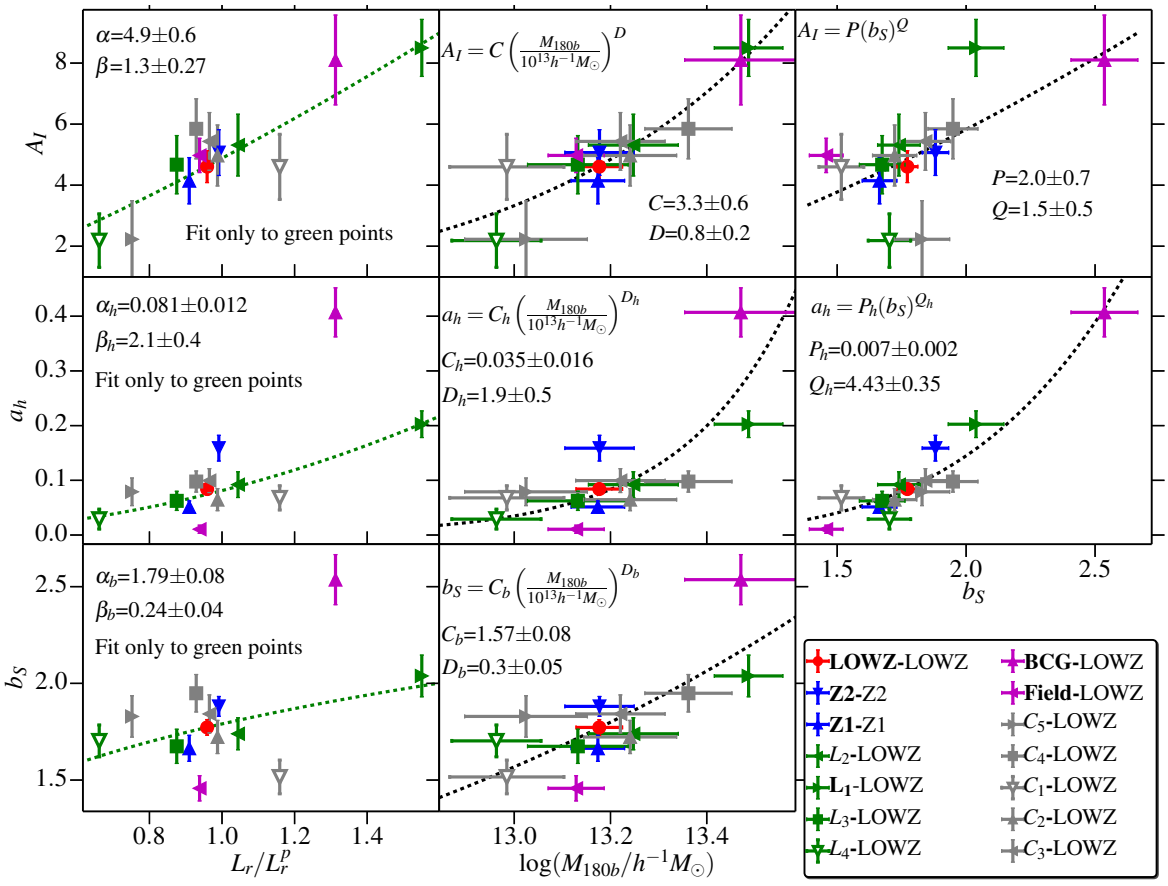



Figure 5.8: Intrinsic alignments amplitudes and bias for various shape samples, as a function of different galaxy properties of the shape sample. Note that the density sample is fixed to full LOWZ sample for all points, except for Z1 and Z2, for which density sample are Z1 and Z2 themselves, respectively. The top row shows $A_I$ as a function of different properties of shape sample. $A_I$ shows clear evolution with luminosity ($M_r$) as well as galaxy mass and bias, with negligible evolution in redshift. The black dotted line in $A_I$ vs. $M_r$ shows a power law fit to the luminosity samples (green points). Similarly in $A_I$ vs $\log(M_{180b}/h^{-1}M_\odot)$ ($M_{180b}$ is the halo mass from weak lensing), the black dotted line is the power law fit, using all the points. The middle row shows the halo model amplitude, $a_h$, as a function of different galaxy properties. For cases where the density sample is fixed to LOWZ, the effects of the non-linear bias of the density sample is the same. Therefore, the observed dependence of $a_h$ is likely due to the dependence of intrinsic alignments on the properties of the shape sample. We find tight correlation between large scale bias of shape sample and $a_h$, which could partially arise from the dependence of the halo model amplitude on the non-linear bias of the shape sample (see Sec. 5.4.7 for more detailed discussion). The black dotted line in $a_h$ vs $M_r$ is the power law fit to luminosity samples (green points), and the dotted line in $a_h$ vs $b_S$ is the powerlaw fit using all the points. Please see Sec. 5.4.6 for possible systematics in the $a_h$ vs. mass relationship. We emphasize that galaxy properties shown on the $x$-axis are correlated, for example, more luminous galaxies also have higher bias and live in more massive halos. Due to this correlation, we do not attempt to model intrinsic alignments amplitudes by simultaneously using more than one such property.

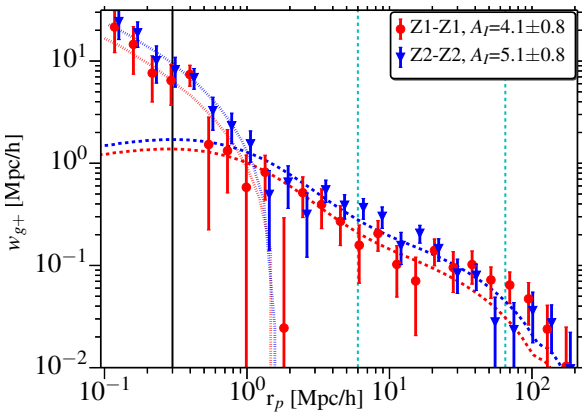



Figure 5.9: Density-shape correlation functions for samples defined with redshift cuts, Z1 ($0.16 < z < 0.26$) and Z2 ($0.26 < z < 0.36$). We do not find any statistically significant redshift evolution for intrinsic alignments.

Z1 sample ($b = 1.66 \pm 0.07$), which is consistent with the fact that among halos of similar mass, ones at higher redshift are more biased. We do not find any redshift evolution for $A_I$, which is consistent with the NLA paradigm, where galaxy alignments are set at the time of galaxy formation and hence we do not expect any significant redshift evolution for intrinsic alignments beyond what is implicitly included due to use of a redshift-dependent nonlinear matter power spectrum that determines the gravitational tidal field. However, given the short redshift baseline, our power to constrain evolution of intrinsic alignments with redshift is limited.

On the small scales where we fit the halo model fitting function, there are significant differences between the two redshift ranges. The Z2 sample has a higher halo model amplitude. However, the Z2 sample is also expected to have a higher non-linear bias which will contribute to the increase in $a_h$. Processes within groups and clusters, such as galactic mergers, stripping, peculiar motion and tidal torquing will also play an important role in determining the small scale signal and can contribute to the redshift evolution of $a_h$. Due to a degeneracy between intrinsic alignments amplitude and non-linear bias within the halo model, we cannot separate out the effects of these processes.

### 5.4.5 Color Dependence

Our sample consists of only luminous red galaxies, so we cannot divide it into subsamples to study difference in intrinsic alignments for red and blue galaxies. However, we do split our sample based on $M_g - M_i$ color to study intrinsic alignments dependence on color within the LRG sample. We divide our sample into five sub-samples based on color, with each sample having 20 per cent of the LOWZ galaxies. Color cuts were applied in $z$ bins to take out the $z$ evolution of color and make sure we select a fair sample. The five samples, $C_1 - C_5$ are arranged from the bluest to reddest. We observe significant luminosity evolution across the five samples, with redder samples getting progressively fainter. We do observe some intrinsic alignments evolution across the different samples, but find that the evolution can be well explained by the luminosity variations alone (see Fig. 5.10). We note that the $C_1$ sample is expected to have some contamination from late-type galaxies (Masters et al., 2011), but we do not observe a very significant deviation in intrinsic alignments signal from the expectation based on luminosity or mass scalings.

Our results suggests that luminosity and mass are more important properties in determining intrinsic alignments signal for red sequence galaxies, with color variations being absent or at least subdominant. This is an important test of the models that are commonly used to predict the intrinsic alignments contamination of future weak lensing surveys (e.g., Joachimi et al., 2011), which split galaxies into a red and a blue sample without permitting any variation in intrinsic alignments with

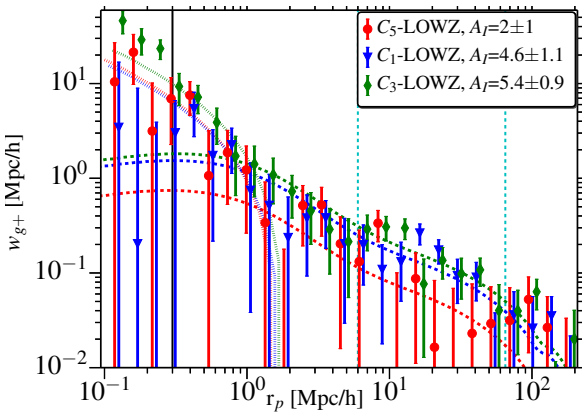



Figure 5.10: Density-shape correlation functions for samples defined with color cuts, with $C_1 - C_5$ arranged from bluest to reddest sample. We observe some intrinsic alignments variations between different samples, but these variations are explained by the luminosity difference between the different samples (see Fig. 5.8).

color within those two samples. Our results validate this choice for the red galaxies, which are currently more important since they are the ones for which there is a robust intrinsic alignments detection.

### 5.4.6 Weak Lensing

To study the halo mass dependence of intrinsic alignments, we compute the average halo mass of galaxies within different samples using galaxy-galaxy lensing. Figure 5.11 shows the weak lensing signal for LOWZ, BGG and field galaxies, with points being measurements from data and dashed lines are the NFW profile fits, with concentration fixed using concentration-mass relation defined in Eq. (5.29). The signal for BGGs deviates from an NFW profile for $r_p \gtrsim 0.5h^{-1}$Mpc. These deviations could be due to some satellite contamination in our BGG sample, which is expected at the level of tens of per cent. At small scales, BGG mis-centering effects (which we have not accounted for in our fits) can also lead to deviations from the NFW profile resulting in underestimated halo mass.

To check for the effect of using fixed concentration-mass relation defined in Eq. (5.29), we vary the amplitude in Eq. (5.29) by 20% and re-fit the signal from field sample, using $0.05h^{-1}\text{Mpc} < r_p < 1.0h^{-1}\text{Mpc}$ (different from range used for main results). We find that the final mass measured changes by $\sim -12\%$.

Mass measurements for different sample are given in table. 5.2. For full LOWZ sample, we get $\log(M/h^{-1}M_\odot) = 13.18\pm0.05$ and for our brightest sample, $L_1$, we get $\log(M/h^{-1}M_\odot) = 13.49\pm0.07$. Using the clustering analysis, Parejko et al. (2013) found $\log(M/h^{-1}M_\odot) \sim 13.72$ for the DR9 LOWZ sample. Their masses correspond to FOF halos with a linking length of 0.2 which are expected to have a higher overdensity (see More et al., 2011), worsening the discrepancy we see here. However, clustering mass estimates are obtained in a less direct way that can have some bias due to modeling assumptions. Clustering and galaxy-galaxy lensing also have different redshift window functions and Parejko et al. (2013) used a different redshift range, $z \in [0.2, 0.4]$, due to which it is hard to do a fair comparison between the two studies.

The third column in Fig. 5.8 shows the variations in intrinsic alignments amplitudes and galaxy bias as function of mass. As expected, more massive galaxies have higher bias, with the mass dependence of the bias being well-described by a power law, $b \propto M_{180b}^{0.30\pm0.05}$. Intrinsic alignments amplitudes are also strongly correlated with the mass, with more massive galaxies showing stronger intrinsic alignments. The relation between $A_I$ and mass can be well defined by a power law:

$$A_I = C\left(\frac{M_{180b}}{10^{13}h^{-1}M_\odot}\right)^D \tag{5.34}$$

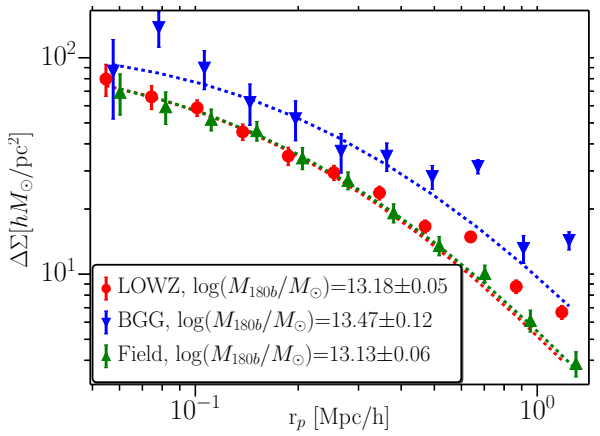



Figure 5.11: Weak lensing measurement for three different samples: LOWZ, BGGs, and field galaxies as indicated in the legend. The dashed lines are the NFW profile fits to the signals. Deviations from the NFW in case of BGGs and LOWZ are due to the presence of satellites in these samples (even the BGG sample, due to the fact that selecting a BGG purely based on luminosity is imperfect). The field galaxy sample is the least likely to suffer from satellite contamination, so its signal is well described by the NFW profile.

We find an amplitude $C = 3.3\pm0.6$ and index $D = 0.8\pm0.2$. Similarly for the small scale amplitude $a_h$, we get $C_h = 0.035\pm0.016$ and index $D_h = 1.9\pm0.5$. However, this fit is significantly effected by the BGG-LOWZ correlation point, where BGG mass measurement are affected by satellite contamination and BGG mis-centering effects. Excluding the BGG-LOWZ point, we find $C_h = 0.047 \pm 0.012$ and index $D_h = 1.2 \pm 0.3$. We also note that we can only test a small range in mass using our LOWZ sample, and this result is unlikely to be valid when extrapolated significantly outside this mass range.

### 5.4.7 Environment Dependence

One of the main goals of this work is to address the environmental dependence of intrinsic alignments. The halo model of intrinsic alignments (Schneider & Bridle, 2010) predicts that satellite galaxies tend to align radially within their host halos, while the BGGs tend to align with the halo shape, which itself tends to align with the large scale tidal gravitational field according to the linear alignment model. As described in Sec. 5.3.3, we identify galaxies in groups using the CiC technique and then separate them into satellites and BGGs. It is important to bear in mind that all results shown in this section for BGGs vs. satellites may be affected by the imperfect BGG selection, which will tend to dilute differences between the samples.

Figures 5.12 and 5.13 show our measurements for different combinations of galaxies within the groups and the field galaxies. In the plots, the density sample for both signals shown is the same, so the interpretation of large-scale amplitudes includes the same factor of density sample bias. At linear scales, we find that BGGs do tend to align with the tidal field and that BGG alignments are also stronger than that of field galaxies. This effect is consistent with the halo mass dependence of intrinsic alignments seen in Sec. 5.4.6, with BGGs living in more massive halos and thus showing stronger alignments. Satellite galaxies on the other hand, do not show any alignments at large scales, with $A_I = 0 \pm 1$. This is discrepant with the luminosity and mass relations of intrinsic alignments observed earlier, but consistent with the expectations from the halo model, where satellites are expected to be radially aligned within the halos and hence will have much weaker signal on large scales. Satellites are also the reason that the large-scale intrinsic alignments amplitude for group galaxies is lower than that of BGGs or even field galaxies.

At small scales, we detect a significant density-shape alignment signal for both satellite galaxies and BGGs. We note that in the BGG-LOWZ correlation, the small scale signal is dominated by BGG-

Group correlations (see also Fig. 5.16). In both cases what we are seeing is the fact that BGGs tend to point preferentially towards their own satellites or, equivalently, that the luminous red satellites of luminous red central galaxies are preferentially located along the major axes of their hosts. Also, in the satellite-LOWZ correlations, the primary contribution to the small scale signal is from within the groups as well, with satellite-BGG correlations (satellites pointing radially towards their hosts) being the dominant signal (see also Fig. 5.15). This confirms two of the assumptions made in the halo model: satellites do preferentially align radially within the halos, and BGGs preferentially align with the shape of their host halo. The results shown in Figures 5.12 and 5.13 are likely to be affected by the contamination in our group sample due to our choice of CiC parameters (see discussion in Sec. 5.3.3), but the results in Figures 5.15 and 5.16 are more robust since most of the satellites are within $|\Pi| \lesssim 10h^{-1}$Mpc of the BGG (see Fig. 5.3) and the signal for $r_p \lesssim 0.6h^{-1}$Mpc is not likely to be affected by the contamination in the group sample. We also note that even though our results do not agree with the null detection of satellite radial alignments in studies using cluster galaxies (e.g., Sifón et al., 2014; Schneider et al., 2013; Hao et al., 2011), our sample of LRGs is inherently different. The samples used in the group and cluster studies are dominated by fainter galaxies, which as shown in Sec. 5.4.3 have lower intrinsic alignments and hence lower the expected signal. Also, those samples are expected to have a small but non-negligible number of late type (blue/disk) galaxies for which intrinsic alignments may arise from angular momentum alignment, which is a second order effect (Hirata & Seljak, 2004). This will further weaken the likelihood of a detection of intrinsic alignments in the cluster studies. Hence, our detection of satellite radial alignments is not necessarily inconsistent with studies using cluster galaxies. Also see Sec. 5.4.10 for more discussion.

In Fig. 5.8, we also observe strong environmental dependence of the halo model fitting function amplitude $a_h$. This result could potentially be contaminated by the effects of non-linear bias of density sample which the halo model does not account for. We factor out this dependence by looking for variations in $a_h$ when the density sample of galaxies is fixed to LOWZ. As shown in Fig. 5.8, we fit a powerlaw relation between halo model amplitude $a_h$ and linear galaxy bias, $b_S$,

$$a_h = P_h(b_S)^{Q_h} \tag{5.35}$$

We find amplitude $P_h = 0.007 \pm 0.002$ and slope $Q_h = 4.43 \pm 0.35$, in the bias range $1.5 < b_S < 2.6$. dependence of small scale IA amplitude ($a_h$) on the large-scale linear bias ($b_S$) is not intuitively obvious. This relation implies that galaxies in more biased and hence more dense regions show higher intrinsic alignments at small scales, even though they are more likely to be effected by processes such galactic mergers and peculiar velocities which can potentially suppress the intrinsic alignments signal. This suggests that effects such as tidal torquing (which counteract mergers, boosting intrinsic alignments) do play an important role in determining intrinsic alignments at small scales.

Also, as described in Appendix 5.A, what we measure is the density-weighted intrinsic shear $\widetilde{\gamma_I} = \gamma_I(1 + \delta_S)$. On small scales where the $\delta_S$ term cannot be neglected, the non-linear bias of shape sample (which is correlated with the large-scale linear bias $b_S$) will also contribute to produce correlations between $a_h$ and $b_S$. These correlations will likely go as $a_h \propto b_S$ which is weaker than what we observe ($a_h \propto b_S^{4.4}$). Thus, the intrinsic alignments at small scales are still driving the correlations between $a_h$ and $b_S$. Given the small scatter in this relationship between $a_h$ and $b_S$ shown in Fig. 5.8, it is apparent that for a given sample of density tracers, the large-scale bias of the shape sample is the single best predictor of the small-scale intrinsic alignments amplitude considered in this work, better than luminosity or host halo mass. We note that the $a_h$ vs. halo mass relation shown in Fig. 5.8 is driven by the BGG-LOWZ data, where the BGG mass can be affected by satellite contamination and mis-centering effects as discussed in Sec. 5.4.6. However, that does not affect our conclusion that the large-scale bias $b_S$ is a more robust predictor of the small-scale amplitude $a_h$ than the galaxy luminosity. These results were all derived with a fixed density tracer sample; we relax this restriction in the following subsection.

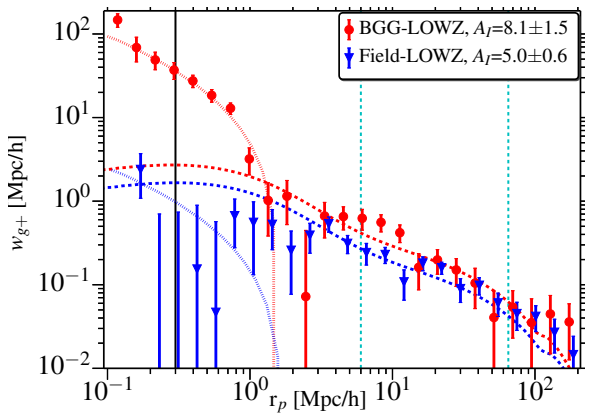



Figure 5.12: Density-shape correlation functions for BGGs and field galaxies. BGGs have a much higher intrinsic alignment amplitude (particularly on small scales), which is consistent with BGGs having their shape aligned with their halos and more massive halos having stronger shape alignments.

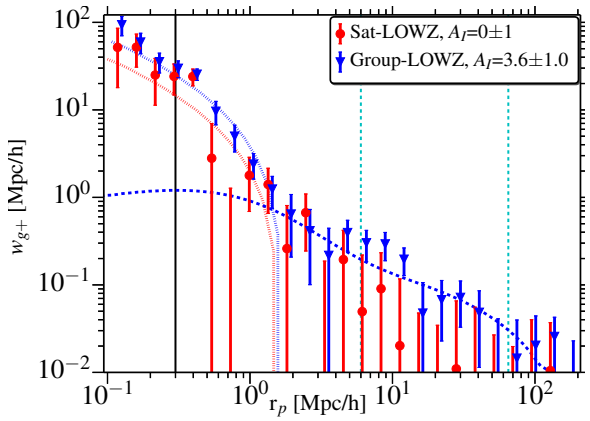



Figure 5.13: Density-shape correlation measurements for group galaxies and satellite galaxies correlated with the full LOWZ sample. On large scales, satellite galaxies show no statistically significant shape alignments. On small scales, satellites show some intrinsic alignments signal, which is primarily coming from radial satellite alignments (see also Fig. 5.15). Group galaxies (satellites + BGGs) exhibit shape alignments with the density field at both small and large scales, with the large scale signal coming primarily from BGGs, and satellites and BGGs both contributing to the small scale signal.

### 5.4.8 Variation of intrinsic alignments with density sample

In this section, we address the question of whether alignments of galaxy shapes also depend on the density sample, i.e., do galaxy shapes tend to be more aligned with more massive or biased objects beyond the simple linear dependence on the bias of the density tracer sample. At linear scales, galaxies are biased tracers of large scale tidal fields and thus within the NLA regime we expect this effect to be negligible. Any changes in $w_{g+}$ for the same shape sample and different density samples should be explained by variations in the bias of the density samples, and not by any changes in intrinsic alignments amplitude, $A_I$. It is unlikely that the same should be true on small scales, where different density samples could have quite different nonlinear bias and environmental effects that will modify the intrinsic alignments beyond what is expected from the large-scale linear bias of the density tracer.

The top panel in Fig. 5.14 shows the variations in $A_I$ when a shape sample is correlated with different density samples. For a given shape sample, we do not find any significant changes in $A_I$ with different density samples. The largest change observed is for the field shape sample, where the difference in $A_I$ for field-LOWZ and field-group is $\lesssim 2\sigma$, after accounting for the correlated errors (but see Fig. 5.19 for variations observed using a different estimator).

As described in Sec. 5.4.7, on small scales and especially within halos, more complex effects can be important in determining the intrinsic alignments signal for a given shape sample with different density samples. This effect is much harder to interpret using the 1-parameter halo model fitting function that we use in this work. As can be seen in Figs. 5.15 and 5.16, our fitting function does not fully capture the scale dependence of the small-scale signal once the density sample is different from the full LOWZ sample. This is not surprising, since we calibrated the shape of the fitting function to fit signals with LOWZ as the density sample. The scale dependence in the $w_{gg}$ signal suggests that non-linear bias should have significant contributions to the scale dependence of intrinsic alignments signal. This effect, along with some possible variation of intrinsic alignments itself, can be captured if we let more free parameters in the fitting function, but the physical interpretation in that case is not very clear and we defer the study of such effects to future work. The bottom panel in Fig. 5.14 also shows the variations in the halo model fit to $w_{g+}$ on small scales. There are significant deviations in $a_h$ observed with variations in the density sample, but as described earlier, it is hard to interpret these in terms of one parameter fitting function; a more detailed physical model is needed.

### 5.4.9 Putting it all together

In this section we briefly discuss and compare the scaling relations of intrinsic alignments amplitudes described earlier and shown in Fig. 5.8. For NLA amplitude, $A_I$, we find that $A_I$ scalings with halo mass and luminosity are well described by a power law and scaling with bias $b_S$ is also well described by a linear fit though with larger scatter. All of these scaling are consistent with the observation that more luminous galaxies, living in more massive halos have stronger alignments. It is interesting to note that scaling with luminosity is stronger than scaling with mass, though we caution that mass and luminosity are correlated. To check if this difference in scaling index has any physical meaning, we fit a power law of of the form $M \sim L^{\nu}$ to mass and luminosity of different samples. We get $\nu = 1.4 \pm 0.2$, which compares well with the expectation from simply comparing $A_I$ scalings, $\nu \sim 1.6 \pm 0.3$. From this we can conclude that these scaling represent same underlying physical effect and one can use either mass or luminosity to predict intrinsic alignments amplitude at large scale. We also note that when fitting for $A_I$ vs $L_r/L_r^p$ scaling, including all the points shown in Fig. 5.8, the parameters donot change significantly ($\alpha = 5.0 \pm 0.5$ and $\beta = 1.33 \pm 0.25$), which further strengthens our conclusion that either mass or luminosity of sample can be used to predict intrinsic alignments amplitude.

We also find similar mass and luminosity scaling for the halo model amplitude $a_h$ and we can also draw similar conclusions that both mass and luminosity scaling present same underlying physical effect. It is also interesting to note that mass and luminosity scaling indices for $a_h$ are higher than that for $A_I$, though scatter is also higher. The correlation with linear galaxy bias for $a_h$ is also much stronger, with linear galaxy bias being much better predictor of $a_h$ than mass or luminosity, within the narrow range of our sample. Since the scaling relations with mass/luminosity are steeper for $a_h$

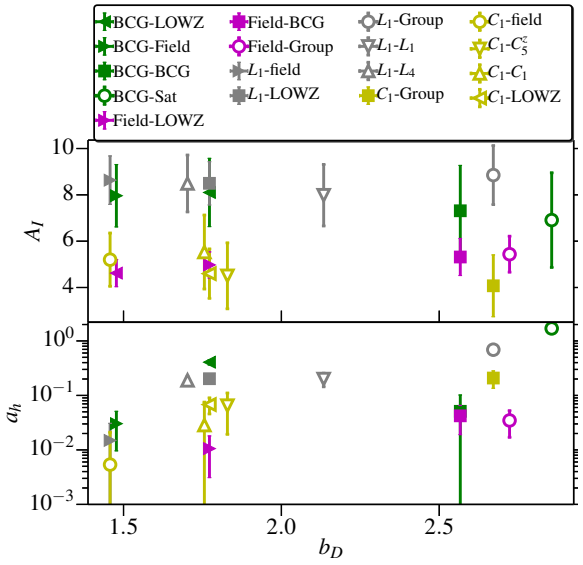



Figure 5.14: Variations of intrinsic alignments amplitude with the density samples. Points are color coded according to the shape samples. Consistent with expectations from the NLA model, we do not observe any significant changes in $A_I$ by varying density samples.

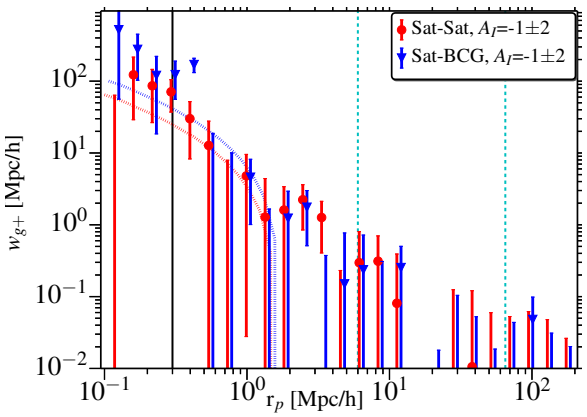



Figure 5.15: Variations in the density-shape correlation functions for satellite galaxies correlated with different density samples.

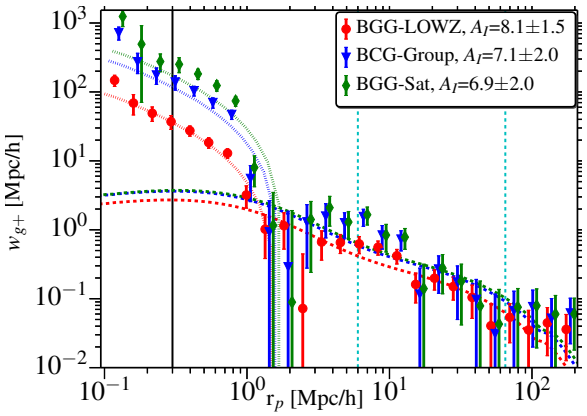



Figure 5.16: Variations in the density-shape correlation functions for BGGs correlated with different density samples. For BGG-Group (blue) and BGG-Satellite (green) correlations, halo model does not fit the shape of the correlation function well (blue and green points are horizontally shifted. We donot fit first four bins since they are affected by fiber collisions). Also BGG-Group are lower than BGG-Sat at small scales (they both effectively measure BGG-Sat at small scales) due to higher RR term in the correlation function, since our randoms ($N_R$) are correlated with the density sample ($N_D$) which is higher for group sample due to presence of BGGs which are absent in satellite sample.

than for $A_I$, this implies that the radial ($r_p$) dependence of the intrinsic alignments on $\sim 300h^{-1}$kpc to tens of $h^{-1}$Mpc scales must be changing as the mass/luminosity change. Thus, a single template for intrinsic alignment radial dependence will be insufficient to model intrinsic alignments for future surveys.

As discussed in Sec. 5.4.5 and Sec. 5.4.4, we do not detect any significant redshift or color evolution of intrinsic alignments amplitudes within the limited range of values probed by our sample (red sequence galaxies in the range $0.16 < z < 0.36$). However, we note that galaxy properties like mass, luminosity and bias also may depend slightly on the redshift and color. Any such evolution of these properties and intrinsic alignments amplitudes is self-consistently explained by the scaling relations discussed above and we do not detect any additional redshift or color dependence.

To summarize, we find that the best (lowest scatter) predictor of the *large-scale* intrinsic alignments signal for a given shape sample is the galaxy luminosity or mass (not its linear bias), whereas the lowest- scatter predictor of the *small-scale* intrinsic alignments signal for a given shape sample is the linear bias (not its luminosity or mass). However, the shape and amplitude of the small-scale intrinsic alignments signal also depends on the choice of density-tracer sample in a non-trivial way, whereas the large-scale signal only depends in a simple (linear) way on the linear bias of the density sample.

### 5.4.10 Comparison with other works

Different studies have used a variety of estimators to measure the intrinsic alignments of galaxies. To compare with their results and also to test for compatibility of different estimators, we present our results using two more estimators, $\langle\theta\rangle$ and $\langle\gamma\rangle$.

$$\langle\gamma\rangle = \frac{S_+D}{SD} \tag{5.36}$$

$$\langle\theta\rangle = \frac{\theta_{SD}}{SD} \tag{5.37}$$

$S_+D$ is as defined in Sec. 5.2.3. $\theta_{SD}$ is the angle between the projected major axis of the shape galaxy and the line joining the pair of galaxies. $\langle\gamma\rangle$ measures the average intrinsic shear of the galaxies, while $\langle\theta\rangle$ measures the tendency of galaxies to align in the direction of other galaxies. In the absence of

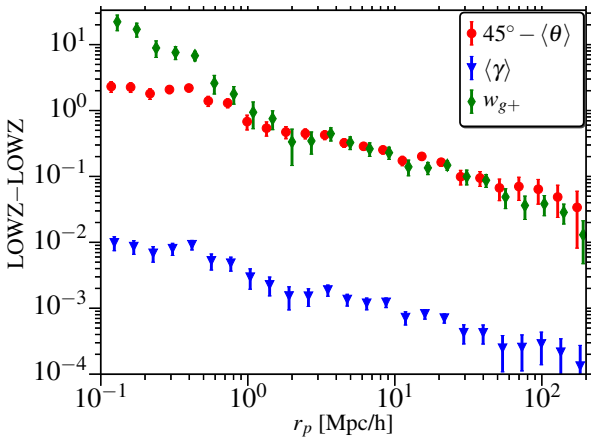



Figure 5.17: Comparison of three different estimators for density-shape intrinsic alignments. We note that the similar amplitude of $45° - \langle\theta\rangle$ and $w_{g+}$ has no physical meaning, since they are very different measurements.

intrinsic alignments, $\langle\gamma\rangle$ will be consistent with zero while $\langle\theta\rangle$ will tend to a value of 45°. In case of radial alignments, $\langle\theta\rangle$ will be less than 45° and $\langle\gamma\rangle$ will be positive.

Both $\langle\theta\rangle$ and $\langle\gamma\rangle$ are measured with a single $\Pi$ bin in the range $[-\Pi_{\rm max}, \Pi_{\rm max}]$. Unless stated otherwise, we take $\Pi_{\rm max} = 100h^{-1}$Mpc.

Figure 5.17 shows our measurement of both $\langle\theta\rangle$ and $\langle\gamma\rangle$ for the full LOWZ sample. There is good agreement with the three estimators in terms of scale dependence, with 1-halo to 2-halo transitions around $r_p \sim 1-2h^{-1}$Mpc clearly observable in all three estimators. The flattening of $45° - \langle\theta\rangle$ and $\langle\gamma\rangle$ at small scales is due to the fact that in taking average, normalization is done by $SD$ while in $w_{g+}$ it is done by $RR$. Since at small scales there are more galaxy pairs than would be expected for a random galaxy distribution, the small scale slope for $w_{g+}$ is steeper than for $45° - \langle\theta\rangle$ and $\langle\gamma\rangle$. The flattening of the signal at $r_p < 0.3h^{-1}$Mpc could also partly be due to contamination from fiber collision galaxies that we have included in our sample, which can potentially bring the signal down if the fiber collision correction is wrong. We note that similar values of $w_{g+}$ and $45° - \langle\theta\rangle$ is a coincidence. They represent very different measurements, and similarities in their magnitudes does not have any physical meaning.

Fig. 5.18 shows the satellite-BGG and BGG-satellite correlations. In this case, we fix $\Pi_{\rm max} = 20h^{-1}$Mpc, so that the signal at small scales, $r_p \lesssim 1h^{-1}$Mpc is only for galaxies within the groups. Our choice of $\Pi$ range here is the same as in Sec. 5.3.3, to select the galaxy pairs within the groups (given that the apparent size of groups along the line of sight is determined by the Fingers of God effect). Sifón et al. (2014) measured the satellite-BGG correlations using spectroscopically selected member for clusters and found a null signal, with $\langle\epsilon_+\rangle = -0.0016 \pm 0.0020$. For our sample, we find satellite radial shear $\langle\gamma_+\rangle = 0.005 \pm 0.001$, for scales $r_p < 1h^{-1}$Mpc. To provide a plausible explanation for this discrepancy, we assume that the power-law relation between halo model amplitude $a_h$ and $r$-band magnitude $M_r$ in Fig. 5.8 can be extrapolated outside of the luminosity range of the LOWZ sample and can also be extended to $\langle\gamma_+\rangle$ measurement for Sat-BCG type correlations. This assumption, coupled with the fact that galaxies used by Sifón et al. (2014) are on average $\sim 1.5$ magnitude fainter than LOWZ galaxies, implies that the expected signal for Sifón et al. (2014) is lower than our $\gamma_+$ by a factor of $\sim 10-50$ (allowing for some possible variations in luminosity scaling), which is well within the size of their errors. We note that disturbed and late type (blue) galaxies in Sifón et al. (2014) sample will further bring down the expected signal, but their galaxies reside in more massive halos, which could push the signal up again. We speculate that, with these two effects acting in opposite direction, the expected signal in Sifón et al. (2014) sample will still be lower than observed in our sample.

For $\langle\theta\rangle$, averaging over $r_p < 1h^{-1}$Mpc, we find $45° - \langle\theta\rangle = 1.5 \pm 0.2$ for satellite-BGG correlations

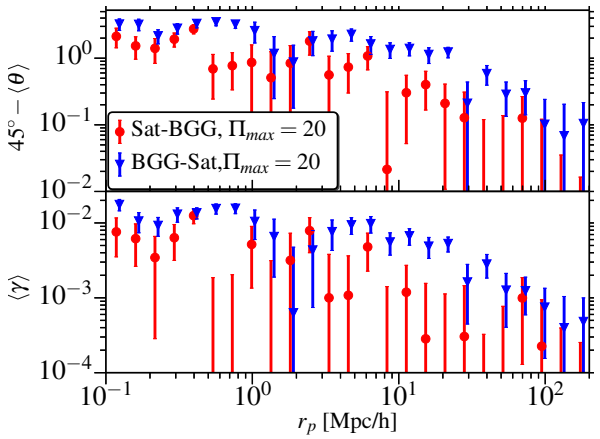



Figure 5.18: $45° - \langle\theta\rangle$ and $\langle\gamma\rangle$ for satellite-BGG (red) and BGG-satellite (blue) correlations. $\Pi_{\rm max}$ is fixed to $20h^{-1}$Mpc so that the small-scale signal comes only from within the same group.

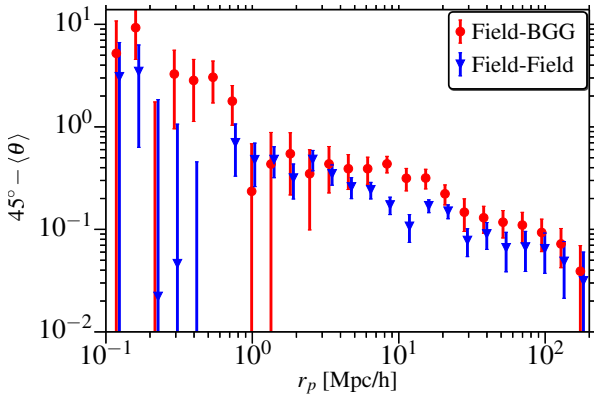



Figure 5.19: $45° - \langle\theta\rangle$ for field galaxies with respect to other field galaxies (blue) and BGGs (red). At scales $r_p \sim 10h^{-1}$Mpc, field galaxies tend to point more towards BGGs than to other field galaxies, which suggests there is some mass dependent sphere of influence.

and $45° - \langle\theta\rangle = 3.0 \pm 0.2$ for BGG-satellite correlations with $\Pi_{\rm max} = 20h^{-1}$Mpc While satellites and BGGs tend to point towards each other, the effect is stronger for BGGs pointing towards satellites rather than the reverse. Though these results suffer from contamination in our group sample due to our choice of CiC parameters (see discussion in Sec. 5.3.3), the results in Fig. 5.18 as function of $r_p$ are robust since most satellites are within $|\Pi| \lesssim 10h^{-1}$Mpc of the BGGs (see Fig. 5.3).

Fig. 5.19 shows a measurement of $45° - \langle\theta\rangle$ for field galaxies, correlated with other field galaxies (blue) and BGGs (red). $45° - \langle\theta\rangle$ is higher for field-BGG than field-field, which suggests that field galaxies have higher tendency to point towards groups than towards other field galaxies. Averaging over $2h^{-1}$Mpc $< r_p < 20h^{-1}$Mpc, we get $45° - \langle\theta\rangle = 0.38 \pm 0.06$ for field-BGG and $45° - \langle\theta\rangle = 0.25 \pm 0.02$ for field-field correlations (errors are correlated since the shape sample is the same in both cases.). The deviation from a strict amplitude rescaling, manifested as a change in shape around $10h^{-1}$Mpc, is likely from filamentary structures, with field galaxies residing in filaments around the groups and hence having higher tendency to align with the groups. This result is qualitatively consistent with the findings of Zhang et al. (2013), who used SDSS DR7 data to reconstruct tidal fields and found that galaxies within filaments tend to have their major axis preferentially aligned with the direction of the filament. More generally, this observation implies that there is a mass dependent sphere of influence within which gravity can align the shapes of galaxies, with BGGs having a larger sphere of influence

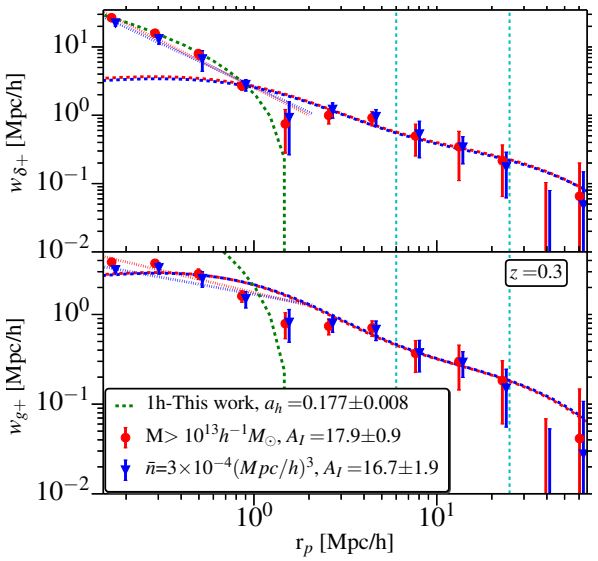



Figure 5.20: $w_{g+}$ and $w_{\delta+}$ measurements from the MB-II hydrodynamical simulation (Tenneti et al., 2014a), at redshift $z = 0.3$ for a halo mass threshold ($M > 10^{13}h^{-1}M_\odot$) sample (blue) and a sample defined by a luminosity threshold to have a fixed comoving number density of $\bar{n} = 3\times10^{-4}$ $(h^{-1}\mathrm{Mpc})^3$. The dashed green lines show our halo model, while the red and blue dashed lines show the best-fit NLA model.

due to the fact that they reside in higher-mass halos.

Fig. 5.20 shows results from Tenneti et al. (2014a) for $w_{g+}$ and $w_{\delta+}$ measurements from the MB-II hydrodynamical simulation, at redshift $z = 0.3$. We show results from two different samples, the first of which is defined by a halo mass threshold $M > 10^{13}h^{-1}M_\odot$ (blue) and the second by a luminosity threshold giving a fixed comoving number density of $\bar{n} = 3 \times 10^{-4}$ $(h^{-1}\mathrm{Mpc})^3$. $w_{g+}$ in Tenneti et al. (2014a) is measured using the shapes of galaxies as defined using the star particles (for those galaxies with $>$ 1000 star particles), with dark matter subhalos as density tracers ($b_D = 0.8$), while for $w_{\delta+}$, dark matter particles are the density tracers ($b_D = 1$). Both $w_{g+}$ and $w_{\delta+}$ are fit simultaneously at linear scales, assuming they have the same $A_I$ but different $b_D$. On non-linear scales, Tenneti et al. (2014a) fit a simple power law. The NLA model intrinsic alignments amplitudes are higher for these simulated samples than for LOWZ. The mass threshold sample is comparable to our $L_1$ and BGG samples in terms of average halo mass, but its intrinsic alignments are higher than what we observe for these samples by a factor of two. The luminosity threshold sample has a comoving number density similar to the full LOWZ sample, but LOWZ has $A_I$ lower by a factor of $\sim 4$.

There are several possibilities that could explain the discrepancy between simulations and observations. The first possibility is some unrealistic aspect of the simulations, including an atypical galaxy shape distribution or a higher degree of alignment between dark matter halo and galaxy shapes than in reality. The second possibility is that this comparison is not completely fair due to the way the galaxy samples were selected in MB-II. For example, the LOWZ sample has a color cut but the simulated samples do not; in real data we are rejecting some massive galaxies that are even slightly blue and including some lower mass red galaxies, but this effect is not included in the simulated samples. Also, a mass-selected sample is unrealistic because our mass tracer (luminosity) is known to have a quite high scatter with mass. Moreover, if this scatter is not present at the same level in the simulated luminosities, that will drastically change the nature of the simulated samples in the direction of having too high average mass and intrinsic alignments. We therefore cannot draw any conclusions from the disagreement in amplitude. However, it is encouraging that despite the difference in selection, the scaling of $w_{g+}$ with $r_p$ in simulations and the LOWZ observations are comparable. This suggests that

the simulations are likely successful in modeling the basic physical processes that cause the intrinsic alignments, even if the magnitude is not quite correct.

## 5.5 Conclusion

We have studied intrinsic alignments in SDSS-III BOSS low redshift (LOWZ) galaxies by combining spectroscopic redshifts from BOSS with shape measurements from Reyes et al. (2012). Using this sample, we have made a robust detection of intrinsic alignments from 0.3 to $200h^{-1}$Mpc scales. We find that the NLA model (linear alignment model with non-linear power spectrum) fits the data well for $r_p > 6h^{-1}$Mpc, though there are significant deviations at smaller scales. At small scales, we also fitted the data with a halo model-based fitting formula, and find that as long as the density tracer sample is fixed, the intrinsic alignments for different shape samples can be described by simply varying the halo model amplitude, i.e. different shape samples do not introduce any significant scale dependence of their own. However, variations in density samples changes the scale dependence on $\lesssim 3h^{-1}$Mpc scales due to changes in the non-linear bias for different samples.

We also studied the luminosity dependence of intrinsic alignments, finding significant evolution of intrinsic alignments with luminosity, with brighter galaxies having stronger intrinsic alignments. For the density sample comprising of LOWZ galaxies, the luminosity dependence of the intrinsic alignments can be well described by a power law fit in luminosity, with power law slope of $\beta = 1.30 \pm 0.27$ for $A_I$ (NLA model amplitude) and $\beta_h = 2.1 \pm 0.4$ for the small-scale halo model fitting function amplitude. While extrapolation beyond the range of luminosities explored here must be done cautiously, this luminosity dependence likely has important implications for future surveys like LSST and Euclid, which will go to much fainter magnitudes than SDSS, and hence overall intrinsic alignments contamination for shear studies in such surveys should go down. This however does not mean that intrinsic alignments contamination will not be important as the statistical error budget for these surveys is also far smaller than for SDSS.

We calculated average halo mass of galaxies in various samples using weak lensing and then studied the mass dependence of intrinsic alignments. We find that the intrinsic alignments amplitude (on both small and large scales) increases with halo mass, and the mass-intrinsic alignments amplitude relations are well described by a power law in mass. This result is consistent with luminosity evolution as well and taken together, both luminosity and mass evolutions imply that brighter and more massive galaxies show stronger intrinsic alignments.

We do not find any significant redshift dependence of intrinsic alignments within the limited range of $0.16 < z < 0.36$. This is consistent with the linear alignment model, which assumes that intrinsic alignments are imprinted at earlier times during the formation of galaxy. Other effects at later times can in principle introduce some redshift dependence, but the effect seems to be too small to be detected in our narrow redshift range.

We also do not find any significant evolution of intrinsic alignments with color, within the red sequence. We split the sample into 5 color bins with each bin containing 20 per cent of the LOWZ sample. There is some luminosity evolution within our color samples, and any intrinsic alignments variation that we see within these samples can be explained by the luminosity scalings.

We identified the galaxies in groups using the CiC technique as described in Sec. 5.3.3 and studied the environment dependence of intrinsic alignments. We find large scale alignments for BGGs in our sample as well as strong alignments of BGGs shapes with satellites. Our results are consistent with the halo model as well as observations from simulations where BGGs are found to be aligned with their host dark matter halos (as traced by satellites), which in turn have large scale alignments as described by linear alignment model. We also detect small scale alignments of satellite shapes, consistent with radial alignments assumed in halo model and observed in simulations. We reiterate that the CiC parameters used in this work were optimized for the higher mass SDSS LRG sample, not the LOWZ sample. This could potentially lead to some contamination in our group sample. However, based on our estimates of the importance of this effect, we expect this to give only minor changes in our overall conclusions about the environment.

Finally, comparison with previous studies using a variety of estimators suggests that our results are either in good agreement or the differences can be explained by differences in the sample selection, most notably the average luminosity of the galaxies.

To summarize, our results suggest that mass and luminosity are the most important factors in determining the intrinsic alignments of red galaxies, with color (within the red sequence) and redshift effects being subdominant. The best predictor (least scatter) for intrinsic alignments at large scales is the halo mass/luminosity of the galaxies, while at small scales, the (large-scale) linear bias is the best predictor. Comparing our halo fitting function and NLA amplitudes to those observed in simulations should help in better understanding of processes involved in galaxy formation and how baryonic matter and dark matter influence each other within the halos. Moreover, our results and constraints on the NLA model and the small-scale fitting function should also be useful in putting priors on models of intrinsic alignments used in mitigation schemes (see for example, King, 2005; Bridle & King, 2007; Joachimi & Bridle, 2010) for upcoming large surveys that aim to use weak lensing to measure the equation of state of dark energy to extremely high precision.

## 5.6 Acknowledgements

This work was supported by the National Science Foundation under Grant. No. AST-1313169. RM was also supported by an Alfred P. Sloan Fellowship. SM is supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan, by the FIRST program "Subaru Measurements of Images and Redshifts (SuMIRe)", CSTP, Japan. We thank Benjamin Joachimi, Alina Kiessling, Marcello Cacciato, Jonathan Blazek, and Cristobal Sifón for useful discussions about this work.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## 5.A Effect of Redshift space distortions on intrinsic alignments

In this appendix we show the effect of redshift space distortions on the intrinsic alignment correlation functions. For simplicity, we will consider configuration space, assume that the plane parallel approximation holds, and that the $z$ direction corresponds to the line-of-sight. The following coordinate transformation maps real space to redshift space,

$$x' = x;\ y' = y;\ z' = z + v\mu;\ , \tag{5.38}$$

where the primed quantities correspond to redshift space, we have assumed units with $H_0 = 1$, $v$ is the magnitude of the velocity at the real space position (positive value corresponds to motion away from the observer) and $\mu$ is the cosine of the angle between the velocity and the line-of-sight.

Let us begin by considering the quantity $\gamma_{\rm I}^{\rm real}$ which is the intrinsic alignment of a shape sample of galaxies with respect to matter in real space. Since the intrinsic shear is measured at the position

of the shape sample of galaxies, in practice, we measure the density-weighted intrinsic shear, $\widetilde{\gamma}_{\rm I}^{\rm real}$, where

$$\widetilde{\gamma}_{\rm I}^{\rm real} = \gamma_{\rm I}^{\rm real}(1+\delta_{\rm S}^{\rm real}) \tag{5.39}$$
$$= \gamma_{\rm I}^{\rm real}(1+b_{\rm S}\delta^{\rm lin}) \tag{5.40}$$
$$= \gamma_{\rm I}^{\rm real} + \mathcal{O}(\delta^{\rm lin}\delta^{\rm lin})\,, \tag{5.41}$$

where $\delta_{\rm S}^{\rm real}$ is the overdensity of the shape sample of galaxies and $b_{\rm S}$ its bias with respect to the matter density fluctuation $\delta^{\rm lin}$. The last equality follows from the proportionality $\gamma_{\rm I}^{\rm real} \propto \phi_{\rm p}$ (see Equation 5.2), and $\phi_p \propto \delta^{\rm lin}$ from the Poisson equation. Here $\delta^{\rm lin}$ is the matter density fluctuation. Thus if we restrict to linear order (as we will throughout this appendix), there is no correction to the intrinsic shear due to the density weighted measurement. Therefore, we will not make a distinction between the intrinsic shear and its density weighted counterpart, hereafter.

In this appendix we will work out the relation between $\gamma_{\rm I}^{\rm real}$ and the corresponding redshift space quantity, $\gamma_{\rm I}^{\rm red}$. Since the shear distortion comes from the tidal field at the true position of the galaxy, and is unchanged by redshift-space distortions, the only effect of RSD is to change the apparent position of the galaxy. This will result in its contributing to the 3D correlation function $\xi_{g+}$ or $\xi_{++}$ at a different value of $\Pi$, but without changing the intrinsic shear value itself. $\gamma_{\rm I}^{\rm red}$ and $\gamma_{\rm I}^{\rm real}$ are therefore related by

$$\gamma_{\rm I}^{\rm red}(x',y',z') = \gamma_{\rm I}^{\rm real}(x,y,z)\,. \tag{5.42}$$

The Fourier transform of the left hand side gives

$$\gamma_{\rm I}^{\rm red}(k'_x,k'_y,k'_z) = \int dx'dy'dz'\gamma_{\rm I}^{\rm red}(x',y',z')e^{i(k'_x x'+k'_y y'+k'_z z')}\,, \tag{5.43}$$
$$= \int dx'\,dy'\,dz'\,\gamma_{\rm I}^{\rm real}(x',y',z'-v\mu)e^{i(k'_x x'+k'_y y'+k'_z z')} \tag{5.44}$$
$$= \int dx\,dy\,dz\,\gamma_{\rm I}^{\rm real}(x,y,z)e^{i(k'_x x+k'_y y+k'_z[z+\mu v])}\frac{dz'}{dz} \tag{5.45}$$

where $dz'/dz = 1+\mu\partial_z v$. We can combine this term and expand the exponential as a Taylor series to obtain

$$\frac{dz'}{dz}e^{ik'_z\mu v} = (1+\mu\partial_z v)\sum_{n=0}^{\infty}\frac{i^n(k'_z\mu v)^n}{n!} \tag{5.46}$$
$$= 1+\mu\partial_z v + ik'_z\mu v + \mathcal{O}(\delta^{\rm lin}\delta^{\rm lin}) \tag{5.47}$$
$$= 1-2\mu^2 f\delta^{\rm lin} + \mathcal{O}(\delta^{\rm lin}\delta^{\rm lin}) \tag{5.48}$$

Here we have used $\mu\partial_z v = ik'_z\mu v = -\mu^2\dot{\delta}^{\rm lin} = -\mu^2 f\delta^{\rm lin}$, where the last two equalities are valid at linear order. Therefore,

$$\gamma_{\rm I}^{\rm red}(k'_x,k'_y,k'_z) = \int dx\,dy\,dz\gamma_{\rm I}^{\rm real}(x,y,z)e^{i(k'_x x'+k'_y y+k'_z z)}[1+\mathcal{O}(\delta^{\rm lin})] \tag{5.49}$$
$$= \gamma_{\rm I}^{\rm real}(k'_x,k'_y,k'_z) + \mathcal{O}(\delta^{\rm lin}\delta^{\rm lin}) \tag{5.50}$$

This implies that there is no linear order correction to the intrinsic shear field while going from real to redshift space.

Now consider a density sample characterized by its overdensity $\delta_{\rm D}$. Number density conservation relates the redshift space overdensity to the real space overdensity such that

$$(1+\delta_{\rm D}^{\rm red})dx'dy'dz' = (1+\delta_{\rm D}^{\rm real})dx\,dy\,dz\,, \tag{5.51}$$

Therefore,

$$\delta_{\rm D}^{\rm red} = (1+\delta_{\rm D}^{\rm real})\left[\frac{dz'}{dz}\right]^{-1} - 1 \tag{5.52}$$
$$= (1+\delta_{\rm D}^{\rm real})\left[1-\mu\partial_z v\right] - 1 + \mathcal{O}(\delta\delta) \tag{5.53}$$
$$= \delta_{\rm D}^{\rm real} - \mu\partial_z v + \mathcal{O}(\delta\delta) \tag{5.54}$$
$$= \delta_{\rm D}^{\rm real} + f\mu^2\delta^{\rm lin} + \mathcal{O}(\delta\delta) \tag{5.55}$$
$$= \delta_{\rm D}^{\rm real}\left(1+\frac{f}{b_{\rm D}}\mu^2\right) + \mathcal{O}(\delta\delta) \tag{5.56}$$

The linear order correction of $(1+f/b_{\rm D}\mu^2)$ when going from real coordinates to redshift coordinates is valid both in configuration and Fourier space.

The auto-power spectrum of the intrinsic shear and its cross power spectrum with the galaxy overdensity field in redshift space and real space are therefore related by

$$P_{++}^{\rm red}(\mathbf{k}) = \langle\widetilde{\gamma}_I^{\rm red}(\mathbf{k})\widetilde{\gamma}_I^{\rm red}(\mathbf{k})\rangle \tag{5.57}$$
$$= \langle\gamma_{\rm I}^{\rm real}(\mathbf{k})\gamma_{\rm I}^{\rm real}(\mathbf{k})\rangle + \mathcal{O}(\delta^3) \tag{5.58}$$
$$= P_{++}^{\rm real}(\mathbf{k}) + \mathcal{O}(\delta^3) \tag{5.59}$$
$$P_{g+}^{\rm red}(\mathbf{k}) = \langle\delta_{\rm D}^{\rm red}(\mathbf{k})\widetilde{\gamma}_I^{\rm red}(\mathbf{k})\rangle \tag{5.60}$$
$$= \left(1+\frac{f}{b_{\rm D}}\mu^2\right)\langle\delta_{\rm D}^{\rm real}(\mathbf{k})\gamma_{\rm I}^{\rm real}(\mathbf{k})\rangle + \mathcal{O}(\delta^3) \tag{5.61}$$
$$= \left(1+\frac{f}{b_{\rm D}}\mu^2\right)P_{g+}^{\rm real}(\mathbf{k}) + \mathcal{O}(\delta^3)\,, \tag{5.62}$$

while the galaxy-galaxy power spectrum in real and redshift space are related by

$$P_{gg}^{\rm red}(\mathbf{k}) = \langle\delta_{\rm D}^{\rm red}(\mathbf{k})\delta_{\rm S}^{\rm red}(\mathbf{k})\rangle \tag{5.63}$$
$$= \left(1+\frac{f}{b_{\rm D}}\mu^2\right)\left(1+\frac{f}{b_{\rm S}}\mu^2\right)\langle\delta_{\rm D}^{\rm real}(\mathbf{k})\delta_{\rm S}^{\rm real}(\mathbf{k})\rangle + \mathcal{O}(\delta^3) \tag{5.64}$$
$$= \left(1+\frac{f}{b_{\rm D}}\mu^2\right)\left(1+\frac{f}{b_{\rm S}}\mu^2\right)P_{gg}^{\rm real}(\mathbf{k}) + \mathcal{O}(\delta^3)\,. \tag{5.65}$$

# 6 Intrinsic alignments of BOSS LOWZ galaxies II: Impact of shape measurement methods

Sukhdeep Singh[1], Rachel Mandelbaum[1]

[1]McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA

**Abstract**

Measurements of intrinsic alignments of galaxy shapes with the large-scale density field, and the inferred intrinsic alignments model parameters, are sensitive to the shape measurement methods used. In this paper we measure the intrinsic alignments of the Sloan Digital Sky Survey-III (SDSS-III) Baryon Oscillation Spectroscopic Survey (BOSS) LOWZ galaxies using three different shape measurement methods (re-Gaussianization, isophotal, and de Vaucouleurs), identifying a variation in the inferred intrinsic alignments amplitude at the 40% level between these methods, independent of the galaxy luminosity or other properties. We also carry out a suite of systematics tests on the shapes and their two-point correlation functions, identifying a pronounced contribution from additive PSF systematics in the de Vaucouleurs shapes. Since different methods measure galaxy shapes at different effective radii, the trends we identify in the intrinsic alignments amplitude are consistent with the interpretation that the outer regions of galaxy shapes are more responsive to tidal fields, resulting in isophote twisting and stronger alignments for isophotal shapes. We observe environment dependence of ellipticity, with brightest galaxies in groups being rounder on average compared to satellite and field galaxies. We also study the anisotropy in intrinsic alignments measurements introduced by projected shapes, finding effects consistent with predictions of the nonlinear alignment model and hydrodynamic simulations. The large variations seen using the different shape measurement methods have important implications for intrinsic alignments forecasting and mitigation with future surveys.

## 6.1 Introduction

Weak gravitational lensing (for a review, see Massey et al., 2010; Weinberg et al., 2013), the deflection of light from distant objects by mass in more nearby foregrounds, results in coherent distortions of galaxy shapes that are measured statistically, by averaging over large ensembles of galaxies. It has the power to reveal the dark matter halos in which galaxies and galaxy clusters reside (e.g., Velander et al., 2014; Coupon et al., 2015; Han et al., 2015; Hudson et al., 2015; Zu & Mandelbaum, 2015), to constrain the growth of cosmic structure and thus the nature of dark energy (e.g., Heymans et al., 2013b; Jee et al., 2013; Mandelbaum et al., 2013), and even to constrain the theory of gravity on cosmological scales (e.g., Reyes et al., 2010; Simpson et al., 2013; Pullen et al., 2015b). Intrinsic alignments of galaxies (IA; for a review, see Joachimi et al. 2015; Troxel & Ishak 2015; Kirk et al. 2015; Kiessling et al. 2015), the coherent alignment of galaxy shapes with each other (II) or with the local density field (GI), result in a violation of the assumption that galaxy shapes are not intrinsically correlated and that any observed shape correlations are from gravitational lensing. Thus, intrinsic alignments are an important astrophysical systematic for weak lensing surveys.

Singh et al. (2015) (hereafter Paper I) studied the IA of galaxies in one of the Sloan Digital Sky Survey-III (SDSS-III) Baryon Oscillation Spectroscopic Survey (BOSS) galaxy samples, called LOWZ. This sample consists of Luminous Red Galaxies (LRGs) with a comoving number density of $\sim 3\times10^{-4}(h/\mathrm{Mpc})^3$ from $0.16 < z < 0.36$. The large sample size and high signal-to-noise ratio

enabled not only a detection of IA, but a study of its dependence on galaxy properties such as mass, luminosity and environment, finding strong correlations of IA with the host halo mass and galaxy luminosity (see also Joachimi et al., 2011). A commonly-adopted theoretical model called the linear alignment model (Catelan et al., 2001; Hirata & Seljak, 2003), which relates the galaxy alignments to the tidal field from large-scale structure at the time of galaxy formation, was found to provide a good description of the data for projected separation $r_p > 6\ h^{-1}$Mpc, provided that the non-linear matter power spectrum was used (NLA model; Bridle & King, 2007).

A natural goal of such studies is to predict the intrinsic alignment contamination in weak lensing surveys, and to provide templates that can be used to marginalize over this effect. However, different IA studies in the literature (e.g., Mandelbaum et al., 2006; Hirata et al., 2007; Okumura et al., 2009; Joachimi et al., 2011; Blazek et al., 2011; Hao et al., 2011; Hung & Ebeling, 2012; Li et al., 2013; Chisari et al., 2014; Sifón et al., 2015b; Singh et al., 2015) use different galaxy shape measurement methods and ensemble IA estimators, which makes it difficult to compare them or to combine their results into a single comprehensive view of the subject. For example, Okumura et al. (2009) used isophotal shape measurements from SDSS data release 7 (DR7) to measure shape-shape correlations for the SDSS LRG sample at high significance which were then interpreted by Blazek et al. (2011) in the context of the NLA model. However, in Paper I using re-Gaussianization shapes and a larger BOSS low redshift sample (LOWZ-DR11) from SDSS-III, we measured the shape-shape correlation function to be consistent with zero. Though the two measurements were shown to be statistically consistent in Paper I, the reason for the varying detection significance in the two studies is not well understood, and worth investigation.

IA measurements with different shape measurements may be particularly difficult to compare due to different systematic errors and ranges of galaxy radius probed by different methods. First, galaxy shape measurements are affected by observational systematics such as the point-spread function (PSF), pixel noise, errors in estimation and subtraction of sky level, and blending with the light profiles of nearby galaxies. Different shape measurement methods treat these systematics differently or in some cases ignore them. Improper treatment of systematic effects in galaxy shapes can propagate into IA measurements. For example, Hao et al. (2011) found that using isophotal shapes from the SDSS to measure the alignments of satellite shapes around brightest cluster galaxies (BCGs) that these alignments correlated with the *apparent* magnitude of the BCG. This result was interpreted as a systematic error in isophotal shapes of satellite galaxies from BCG light leaking into the satellite shapes, an effect that is more complicated than a multiplicative bias.

Also, different shape measurement methods use different radial weight functions and thus probe galaxy shapes at different effective radii. Intrinsic variation in the galaxy shapes with radius can thus affect the IA measurements. These could be gradients in the ellipticity with radius (with galaxies being intrinsically more or less round in the outer regions), or isophotal twisting due to the outer parts of galaxies being more aligned with external tidal field than the inner parts. For example, Tenneti et al. (2015) found in hydrodynamic simulations that galaxies become rounder in their outer regions. Velliscig et al. (2015b), on the other hand, found that the projected RMS ellipticity, $e_{\rm rms}$, is lower when measured using star particles within the half-light radius compared to using all star particles (see also Chisari et al., 2015a). In observations, the magnitude and sign of ellipticity gradients have been found to be correlated with the galaxy environment (di Tullio, 1978, 1979; Pasquali et al., 2006). Ellipticity variations using different shape measurements will affect inferences made using ensemble IA estimators that include the ellipticity rather than just the position angle (e.g., Mandelbaum et al., 2006; Hirata et al., 2007; Joachimi et al., 2011; Blazek et al., 2011; Singh et al., 2015).

Several studies have also detected isophote twists in small samples of elliptical galaxies, whereby the position angle in the measured galaxy shape changes with radius (Wyatt, 1953; Abramenko, 1978; Kormendy, 1982; Fasano & Bonoli, 1989; Nieto et al., 1992; Lauer et al., 2005). The isophote twisting may originate from varying triaxiality of galaxies with radii (see, e.g., Romanowsky & Kochanek, 1998) though Kormendy (1982) pointed out that the outer regions of galaxies are more susceptible to tidal fields, which can result in isophote twisting. In simulations, Kuhlen et al. (2007) observed the effects of shape twisting from tidal interaction when measuring the radial alignments of dark matter

subhalos. The radial alignment signal of subhaloes increased monotonically with the radius at which the subhalo shape was measured. If such results also apply to galaxies, then a stronger IA signal from shape measurements that probe the outer regions of galaxies would be expected. In support of this inference, Velliscig et al. (2015a) found using hydrodynamic simulations that using star particles within the half-light radius to define the galaxy shapes results in lower IA signal compared to using all the star particles (see also Chisari et al., 2015a, but note that that work attributes the differences to ellipticity variations rather than to isophote twists).

Finally, we measure only the projected shapes of galaxies, which are insensitive to line-of-sight galaxy alignments. This effect introduces anisotropy in the redshift-space structure of IA[1], and different estimators of IA vary in their sensitivity to this anisotropy. For example, the redshift-space structure of the 3D shape-density cross-correlation function, $\xi_{g+}$, includes the redshift-space structure of the 3D galaxy-galaxy auto-correlation function, $\xi_{gg}$; however, the mean IA shear, $\langle\gamma\rangle$, is independent of $\xi_{gg}$ ($\xi_{g+} = \langle\gamma\rangle(1+\xi_{gg})$, Blazek et al. 2015). Depending on the relative importance of variations in IA and galaxy clustering, $\xi_{g+}$ and $\langle\gamma\rangle$ may have different redshift-space structure. The variations in the sensitivity to redshift-space structure between different IA estimators can complicate a quantitive comparisons between studies using different estimators.

In this paper, we repeat the analysis of Paper I using three different shape measurement methods, and carry out numerous systematics tests, to study the radial and environment dependence of galaxy shapes and of the IA signal. We also use the methodology of Blazek et al. (2011) to understand the origin of differences in their shape-shape correlations ($w_{++}$) measurement and the one in Paper I. Finally, to understand the redshift-space structure of intrinsic alignments, we investigate the IA signals as a function of projected and line of sight separations and compare the results with NLA model predictions.

Throughout we use a standard flat ΛCDM cosmology with $h = 0.7$, $\Omega_b = 0.046$, $\Omega_{DM} = 0.236$, $\Omega_\Lambda = 0.718$, $n_s = 0.9646$, $\sigma_8 = 0.817$ (WMAP9, Hinshaw et al., 2013). All distances are in comoving $h^{-1}$Mpc, though $h = 0.7$ was used to calculate absolute magnitudes and to generate predictions for the matter power spectrum.

## 6.2 Formalism and Methodology

Details of intrinsic alignments models and correlation estimators are given in Paper I. In this section, we briefly summarize the important points.

### 6.2.1 The nonlinear alignment (NLA) model

The linear alignment (LA) model predicts that IA are set at the time of galaxy formation (Catelan et al., 2001), with galaxy shapes being aligned with the tidal fields present during galaxy formation. This assumption allows us to write intrinsic shear in terms of primordial potential $\phi_p$

$$\gamma^I = (\gamma^I_+, \gamma^I_\times) = -\frac{C_1}{4\pi G}(\partial_x^2 - \partial_y^2, \partial_x\partial_y)\phi_p, \tag{6.1}$$

with alignment strength defined by an amplitude parameter $C_1$. In our sign convention, positive (negative) $\gamma^I_+$ indicates alignments along (perpendicular to) the direction of the tidal field while positive (negative) $\gamma^I_\times$ indicates alignments along the direction at 45 (135) degrees from the direction of the tidal field. Assuming a linear galaxy bias relating matter overdensities $\delta_m$ and galaxy densities $\delta_g = b\,\delta_m$, the power spectrum of galaxy-shape and shape-shape correlations can be written as (Hirata

[1]This anisotropy due to projected shapes is also present in real space, but we will use the term redshift-space throughout this paper since measurements are made in redshift-space.

& Seljak, 2003)

$$P_{g+}(\mathbf{k},z) = A_I b \frac{C_1\rho_{\rm crit}\Omega_m}{D(z)} \frac{k_x^2-k_y^2}{k^2} P_\delta^{\rm lin}(\mathbf{k},z) \tag{6.2}$$

$$P_{++}(\mathbf{k},z) = f_{II}\left(A_I \frac{C_1\rho_{\rm crit}\Omega_m}{D(z)} \frac{k_x^2-k_y^2}{k^2}\right)^2 P_\delta^{\rm lin}(\mathbf{k},z) \tag{6.3}$$

$$P_{g\times}(\mathbf{k},z) = A_I b \frac{C_1\rho_{\rm crit}\Omega_m}{D(z)} \frac{k_x k_y}{k^2} P_\delta^{\rm lin}(\mathbf{k},z) \tag{6.4}$$

$P_\delta^{\rm lin}$ is the linear matter power spectrum. $P_{g+}$ ($P_{g\times}$) is the cross-power spectrum between the galaxy density field and the shear component along (at 45° from) the line joining the galaxy pair. $P_{++}$ is the shape-shape correlation with shear component along the line joining the galaxy pair. Following Joachimi et al. (2011), we fix $C_1\rho_{\rm crit} = 0.0134$ and use a dimensionless constant $A_I$ to measure the IA amplitude, which is primarily constrained from density-shape cross-correlations. To allow for departure from the LA model, we have introduced an additional free parameter $f_{II}$, which would be 1 in the case that the LA model is correct. In the NLA model, to extend the linear alignment model to the non-linear regime, the linear matter power spectrum is replaced with the non-linear matter power spectrum (Bridle & King, 2007) using an updated halo-fit model (Smith et al., 2003; Takahashi et al., 2012). As shown by Blazek et al. (2015), the NLA model neglects other terms that are important at the same order; however, in Paper I we found it provided an adequate fit to our intrinsic alignments measurements down to $\sim 6\ h^{-1}$Mpc.

The power spectra in Eqs. (6.2)–(6.4) can be Fourier transformed to obtain the 3D correlation functions as a function of comoving projected separation $r_p$ and line-of-sight separation $\Pi$,

$$\xi_{AB}(r_p,\Pi,z) = \int \frac{{\rm d}^2k_\perp {\rm d}k_z}{(2\pi)^3} P_{AB}(\mathbf{k},z)\left(1+\beta_A\mu^2\right) \left(1+\beta_B\mu^2\right) e^{i(r_p.k_\perp+\Pi k_z)}. \tag{6.5}$$

The Kaiser factor $(1+\beta\mu^2)$ accounts for the effect of linear redshift-space distortions (RSD; Kaiser 1987). As shown in Paper I, for power spectra that include intrinsic shapes, $\beta_{+,\times}=0$ (the shear field is not affected by RSD to first order). For galaxies, $\beta(z) = f(z)/b$, where the linear growth factor $f(z)\sim\Omega_m(z)^{0.55}$ in the $\Lambda$CDM model.

In the data, we measure the projected correlation function, $w_{AB}(r_p)$, which can be obtained by projecting the 3D correlation function in Eq. (6.5) along the line-of-sight separation:

$$w_{AB}(r_p) = \int {\rm d}z\, W(z) \int {\rm d}\Pi\, \xi_{AB}(r_p,\Pi,z). \tag{6.6}$$

$W(z)$ is the redshift window function (Mandelbaum et al., 2011)

$$W(z) = \frac{p_A(z)p_B(z)}{\chi^2(z){\rm d}\chi/{\rm d}z} \left[\int \frac{p_A(z)p_B(z)}{\chi^2(z){\rm d}\chi/{\rm d}z}{\rm d}z\right]^{-1} \tag{6.7}$$

Assuming cross-correlation between samples of galaxies with shapes S and others that trace the density field D with in principle different biases $b_S$ and $b_D$, the different projected correlation functions are

$$w_{gg}(r_p) = \frac{b_S b_D}{\pi^2}\int {\rm d}z\, W(z) \int_0^\infty {\rm d}k_z \int_0^\infty {\rm d}k_\perp \frac{k_\perp}{k_z} P(\mathbf{k},z)\sin(k_z\Pi_{\rm max})J_0(k_\perp r_p)\left(1+\beta_S\mu^2\right)\left(1+\beta_D\mu^2\right)$$

$$w_{g+}(r_p) = \frac{A_I b_D C_1\rho_{\rm crit}\Omega_m}{\pi^2}\int {\rm d}z \frac{W(z)}{D(z)} \int_0^\infty {\rm d}k_z \int_0^\infty {\rm d}k_\perp \frac{k_\perp^3}{(k_\perp^2+k_z^2)k_z} P(\mathbf{k},z)\sin(k_z\Pi_{\rm max})J_2(k_\perp r_p)\left(1+\beta_D\mu^2\right) \tag{6.8}$$

$$w_{++}(r_p) = f_{II}\frac{(A_I C_1 \rho_{\rm crit}\Omega_m)^2}{2\pi^2}\int {\rm d}z \frac{W(z)}{D(z)^2}\int_0^\infty {\rm d}k_z \int_0^\infty {\rm d}k_\perp \frac{k_\perp^5}{(k_\perp^2+k_z^2)^2 k_z} P(\mathbf{k},z)\sin(k_z\Pi_{\rm max}) \times[J_0(k_\perp r_p)+J_4(k_\perp r_p)] \quad (6.9)$$

Whether we fit the three correlation functions ($w_{gg}$, $w_{g+}$, and $w_{++}$) jointly or independently to this model, the galaxy bias, intrinsic alignment amplitude $A_I$, and $f_{II}$ are primarily constrained by $w_{gg}$, $w_{g+}$, and $w_{++}$, respectively. This is due to the different sensitivities of these signals to the three parameters, along with the different signal-to-noise ratio ($S/N$) in the measurements. The $S/N$ degrades with increasing factors of galaxy shape due to the shape noise they add to the correlation function.

### 6.2.2 Correlation function estimators

We calculate cross-correlation functions between sets of galaxies for which we wish to estimate the intrinsic alignments (the "shape sample") and those used to trace the density field (the "density sample"). We use a generalized Landy-Szalay estimator (Landy & Szalay, 1993) to compute the correlation functions:

$$\xi_{gg} = \frac{(S-R_S)(D-R_D)}{R_S R_D} = \frac{SD - R_S D - S R_D}{R_S R_D} + 1$$

$$\xi_{g+} = \frac{S_+ D - S_+ R_D}{R_S R_D}$$

$$\xi_{++} = \frac{S_+ S_+}{R_S R_S},$$

where $S$ and $D$ represent the galaxy counts in the shape and density samples, and $R_S$ and $R_D$ are sets of random points corresponding to these samples. The terms involving shears for the galaxies are

$$S_+X = \sum_{i\in S, j\in X} \gamma_+^{(i)}(j|i),$$
$$S_+S_+ = \sum_{i\in S, j\in S} \gamma_+^{(i)}(i|j)\gamma_+^{(j)}(j|i).$$

Here $\gamma_+^{(i)}(j|i)$ represents the component of the shear for galaxy $i$ along the line joining it to galaxy $j$. Positive (negative) $\gamma_+$ implies radial (tangential) alignments.

Measuring projected correlations involves summation over bins in $\Pi$,

$$w_{AB} = \int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} \xi_{AB}(r_p,\Pi)\,{\rm d}\Pi. \quad (6.10)$$

We use $\Pi_{\rm max} = 100h^{-1}$Mpc and $d\Pi = 10h^{-1}$Mpc.

To calculate the covariance matrices of the $w_{AB}$, we divide the sample into 100 equal-area regions, compute the signal by excluding one region at a time and then compute the jackknife variance from the 100 jackknife samples. See Paper I for more details.

### 6.2.3 Anisotropy

It is well known that redshift-space distortions (RSD) introduce anisotropy in the observed galaxy clustering in redshift space (e.g., Kaiser, 1987; Beutler et al., 2014). In Paper I, it was shown that the IA measurements are not affect by RSD to first order. However, IA are affected by another source of anisotropy, the projected shapes of galaxies, due to which we cannot measure the line-of-sight IA

signal. Hence, the IA signal is expected to fall faster with increasing line-of-sight separation, $\Pi$, than with $r_p$.

To understand this anisotropy, we study the IA signal in $(r_p, \Pi)$ space and compare it with the NLA model prediction. To compute the model predictions, we compute the real part of $\xi_{AB}$ as follows:

$$\begin{aligned}\xi_{gg}(r_p,\Pi) =& \frac{b_S b_D}{2\pi^2}\int \mathrm{d}z\, W(z)\int_0^\infty \mathrm{d}k_z \cos(k_z\Pi)\\ &\int_0^\infty \mathrm{d}k_\perp \frac{k_\perp}{k_z} P(\mathbf{k},z) J_0(k_\perp r_p)\left(1+\beta_S\mu^2\right)\\ &\left(1+\beta_D\mu^2\right) \qquad (6.11)\\ \xi_{g+}(r_p,\Pi) =& \frac{A_I b_D C_1 \rho_{\rm crit}\Omega_m}{2\pi^2}\int \mathrm{d}z \frac{W(z)}{D(z)}\int_0^\infty \mathrm{d}k_z \cos(k_z\Pi)\\ &\int_0^\infty \mathrm{d}k_\perp \frac{k_\perp^3}{(k_\perp^2+k_z^2)k_z} P(\mathbf{k},z) J_2(k_\perp r_p)\\ &\left(1+\beta_D\mu^2\right) \qquad (6.12)\\ \xi_{++}(r_p,\Pi) =& f_{II}\frac{(A_I C_1\rho_{\rm crit}\Omega_m)^2}{4\pi^2}\int \mathrm{d}z \frac{W(z)}{D(z)^2}\int_0^\infty \mathrm{d}k_z\\ &\cos(k_z\Pi)\int_0^\infty \mathrm{d}k_\perp \frac{k_\perp^5}{(k_\perp^2+k_z^2)^2 k_z} P(\mathbf{k},z)\times\\ &[J_0(k_\perp r_p)+J_4(k_\perp r_p)] \qquad (6.13)\end{aligned}$$

The $k_\perp^2/k^2$ terms in Eqs. (6.12) and (6.13) are equivalent to the Kaiser factor with $\beta = -1$:

$$\frac{k_\perp^2}{k^2} = 1 - \frac{k_z^2}{k^2} = 1-\mu^2 \qquad (6.14)$$

Mathematically, the projection effects in $\xi_{g+}$ and $\xi_{++}$ introduce similar factors in the power spectrum as the Kaiser effect. The Kaiser factor with positive $\beta$ leads to compression along the line of sight direction, while a negative prefactor for $\mu^2$ leads to compression along $r_p$. The final shape of the correlation function is also determined by the Bessel functions which are different for different correlation functions. In the absence of RSD and shape projection factors, $J_0$ leads to an isotropic correlation function, while $J_2$ and $J_4$ lead to a peanut-shaped function in $(r_p, \Pi)$ space ($J_2, J_4 \to 0$ for $r_p \to 0$). Following the methodology used to study RSD (see, e.g., Beutler et al., 2014), we also compute the monopole and quadrupole terms for the $\xi_{AB}$ using an expansion in Legendre polynomials.

$$\xi^l(r) = \frac{2l+1}{2}\int \mathrm{d}\mu_r\, L^l(\mu_r)\xi(r,\mu_r) \qquad (6.15)$$

$r = \sqrt{r_p^2+\Pi^2} = s$ (in redshift space) is the 3D separation of the galaxy pair, while $\mu_r = \Pi/r$ is the cosine of the angle between $\Pi$ and $r$.

In the data, we compute $\xi_{AB}$ as a function of $s$ and $\mu$, and apply the transform defined in Eq. (6.15). For theory calculations, we first compute $\xi_{AB}$ in the $(r_p, \Pi)$ plane on a grid, then transform it to the $(r, \mu)$ plane and apply the transform defined in Eq. (6.15). Since the theory prediction is calculated on a grid, there is some noise due to the finite grid size and Fourier ringing. Even though we smooth out the noise, we do not attempt to fit the theory to the observed multipoles. Instead, we use the best-fitting parameters from fits to $w_{AB}$, and simply compare the theory predictions with the data to confirm whether the observed trends in $(r_p, \Pi)$ are consistent with the above model on scales large enough that effects from non-linear clustering and non-linear RSD are not important.

## 6.3 Data

The SDSS (York et al., 2000) imaged roughly $\pi$ steradians of the sky, and the SDSS-I and II surveys followed up approximately one million of the detected objects spectroscopically (Eisenstein et al., 2001; Richards et al., 2002; Strauss et al., 2002). The imaging was carried out by drift-scanning the sky in photometric conditions (Hogg et al., 2001; Ivezić et al., 2004), in five bands ($ugriz$) (Fukugita et al., 1996; Smith et al., 2002) using a specially-designed wide-field camera (Gunn et al., 1998) on the SDSS Telescope (Gunn et al., 2006). These imaging data were used to create the catalogues of shear estimates that we use in this paper. All of the data were processed by completely automated pipelines that detect and measure photometric properties of objects, and astrometrically calibrate the data (Lupton et al., 2001a; Pier et al., 2003; Tucker et al., 2006). The SDSS-I/II imaging surveys were completed with a seventh data release (Abazajian et al., 2009), though this work will rely as well on an improved data reduction pipeline that was part of the eighth data release, from SDSS-III (Aihara et al., 2011); and an improved photometric calibration ('ubercalibration', Padmanabhan et al., 2008).

### 6.3.1 Redshifts

Based on the photometric catalog, galaxies are selected for spectroscopic observation (Dawson et al., 2013), and the BOSS spectroscopic survey was performed (Ahn et al., 2012) using the BOSS spectrographs (Smee et al., 2013). Targets are assigned to tiles of diameter 3° using an adaptive tiling algorithm (Blanton et al., 2003), and the data were processed by an automated spectral classification, redshift determination, and parameter measurement pipeline (Bolton et al., 2012).

We use SDSS-III BOSS data release 11 (DR11; Alam et al., 2015a) LOWZ galaxies, in the redshift range $0.16 < z < 0.36$. The LOWZ sample consists of Luminous Red Galaxies (LRGs) at $z < 0.4$, selected from the SDSS DR8 imaging data and observed spectroscopically in the BOSS survey. The sample is approximately volume-limited in the redshift range $0.16 < z < 0.36$, with a number density of $\bar{n} \sim 3 \times 10^{-4}\ h^3\mathrm{Mpc}^{-3}$ (Manera et al., 2015). We combine the spectroscopic redshifts from BOSS with galaxy shape measurements from Reyes et al. (2012). BOSS DR11 has 225334 LOWZ galaxies within our redshift range. However, Reyes et al. (2012) masks out certain regions that have higher Galactic extinction, leaving us with 173855 galaxies for our LOWZ density sample.

### 6.3.2 Subsamples

To test for the dependence of our results on galaxy properties such as luminosity, color, and redshift, we split the LOWZ sample into subsamples based on these properties following the methodology detailed in Paper I. To summarize, we define four subsamples based on luminosity, $L_1$–$L_4$, with $L_1$ ($L_4$) being the brightest (faintest) subsample. Luminosity cuts are applied in 10 redshift bins, with $L_1$–$L_3$ each containing 20% of the galaxies and $L_4$ containing 40% to improve the $S/N$ for the fainter galaxies. Similarly, we define color subsamples $C_1$–$C_5$, each containing 20% of the galaxies, with $C_1$ being the reddest subsample. For redshift, we define two subsamples, $Z_1$ ($z \in [0.16, 0.26]$) and $Z_2$ ($z \in [0.26, 0.36]$).

We also identify the galaxies in groups using the counts-in-cylinders (CiC) method (Reid & Spergel, 2009) and split the sample into field galaxies (group of one), BGG (brightest group galaxy) and satellites (all non-field and non-BGGs). See Paper I for more details and caveats related to the CiC group identification.

### 6.3.3 Shapes

To measure the intrinsic alignments, we need to measure the shapes of the galaxies to be used in estimates of the ensemble intrinsic shear. One of the main goals of this work is to compare different shape measurement methods and their impact on the measured IA signal. Here we briefly describe the three different shape measurements used in this paper. Though there are many shape measurement methods that are used for weak lensing in the literature (see for example Mandelbaum et al., 2015),

our choice in this work is limited to the methods that (a) are available for SDSS data and (b) have been previously used in intrinsic alignment studies in the SDSS. However, we attempt to derive more general conclusions that could be applicable to other shape measurement methods by considering the essential properties of these methods (e.g., the choice of radii that they are sensitive to within the galaxy light profiles).

Note that we do not have usable shape estimates from all methods for all the galaxies. In order to do a fair comparison between different methods, the final shape sample in this work only contains the galaxies for which shapes are available from all three methods. Due to differences in the sky coverage of different shapes (see Sec. 6.3.3) we mask out some additional area on the sky, also reducing the size of our density sample. The final shape (density) sample has 122513 (131227) galaxies, making it smaller and somewhat intrinsically brighter on average than the shape (density) sample used in Paper I.

### Re-Gaussianization Shapes

Re-Gaussianization shapes were used in the IA study done in Paper I. These shape measurements are described in more detail in Reyes et al. (2012). Briefly, these shapes are measured using the re-Gaussianization technique developed by Hirata & Seljak (2003). The algorithm is a modified version of ones that use "adaptive moments" (equivalent to fitting the light intensity profile to an elliptical Gaussian), determining shapes of the PSF-convolved galaxy image based on adaptive moments and then correcting the resulting shapes based on adaptive moments of the PSF. The re-Gaussianization method involves additional steps to correct for non-Gaussianity of both the PSF and the galaxy surface brightness profiles (Hirata & Seljak, 2003). The components of the distortion are defined as

$$(e_+, e_\times) = \frac{1-(b/a)^2}{1+(b/a)^2}(\cos 2\phi, \sin 2\phi), \tag{6.16}$$

where $b/a$ is the minor-to-major axis ratio and $\phi$ is the position angle of the major axis on the sky with respect to the RA-Dec coordinate system. The ensemble average of the distortion is related to the shear as

$$\gamma_+, \gamma_\times = \frac{\langle e_+, e_\times \rangle}{2\mathcal{R}} \tag{6.17}$$

$$\mathcal{R} = 1 - \frac{1}{2}\langle e_{+,i}^2 + e_{\times,i}^2 - 2\sigma_i^2 \rangle \tag{6.18}$$

where $\sigma_i$ is the per-component measurement uncertainty of the galaxy distortion, and $\mathcal{R}$ is the shear responsivity representing the response of an ensemble of galaxies with some intrinsic distribution of distortion values to a small shear (Kaiser et al., 1995; Bernstein & Jarvis, 2002). In Paper I we used an $\mathcal{R}$ value for the entire SDSS shape sample from Reyes et al. (2012), $\mathcal{R} \sim 0.87$, whereas using Eq. (6.18) we find $\mathcal{R} \sim 0.925$ for the LOWZ sample, which we use in this paper. In order to compare with Paper I, the results in that paper should be rescaled by $0.87/0.925 \approx 0.94$. For different subsamples, we calculate $\mathcal{R}$ values separately using Eq. (6.18), though the variations between our subsamples are $\lesssim 1\%$, well below the statistical errors.

### Isophotal Shapes

Isophotal shapes are measured in general by fitting an ellipse to an outer isophote of the observed galaxy light profiles. SDSS DR7 provides isophotal shapes[2] of galaxies using the isophote corresponding to a surface brightness of 25 mag/arcsec$^2$. These have been used for several measurements of intrinsic alignments in the SDSS (Okumura et al., 2009; Hao et al., 2011; Zhang et al., 2013). The LRG intrinsic alignments measurements by Okumura et al. (2009) were later interpreted by Blazek et al. (2011) in the context of the NLA model. Hao et al. (2011) measured satellite alignments in

[2]http://classic.sdss.org/dr7/ algorithms/classify.html

a sample of galaxy clusters. They found that the satellite alignment signal depended on the BCG apparent magnitude, which they attributed to contamination from BCG light leaking into satellite isophotal shapes. In addition, isophotal shapes are not corrected for the PSF. Due to the unreliability of these isophotal measurements, they were not included in the SDSS data release 8 (DR8) or subsequent data releases. Thus, the isophotal shapes we use for this study do not cover the part of the LOWZ area coverage for which photometric measurements were made during DR8.

We use the isophotal shape parameters `isoA` (semi-major axis, $a$), `isoB` (semi-minor axis, $b$) and `isoPhi` (major axis position angle, $\phi$) from the SDSS DR7 sky server. We get good isophotal shape measurements for $\sim 297,000$ galaxies ($\sim 71\%$) in the full LOWZ sample, of which 122513 are in the redshift range ($0.16 < z < 0.36$) that we use for final analysis. The ellipticity is defined as

$$(\varepsilon_+, \varepsilon_\times) = \frac{1 - b/a}{1 + b/a}(\cos 2\phi, \sin 2\phi). \tag{6.19}$$

The ensemble average of the ellipticity is an estimator for the shear:

$$\gamma_+, \gamma_\times = \langle \varepsilon_+, \varepsilon_\times \rangle. \tag{6.20}$$

### de Vaucouleurs Shapes

The de Vaucouleurs shapes are defined by fitting galaxy images to a de Vaucouleurs profile[3] (Stoughton et al., 2002),

$$I(r) = I_0 \exp\{-7.67[r/r_{\rm eff}]^{1/4}\} \tag{6.21}$$

with half-light radius $r_{\rm eff}$. The profile is allowed to have arbitrary axis ratio and position angle, and is truncated beyond $7r_{\rm eff}$ to go smoothly to zero beyond $8r_{\rm eff}$. It is convolved with a double Gaussian approximation to the PSF model (for more details, see Stoughton et al., 2002). To reduce computation time, the fitting uses pre-computed tables of models (Lupton et al., 2001a), resulting in discretization in the resulting model parameters (Stoughton et al., 2002). The fit yields axis ratios and position angles that can be used to define ellipticities as in Eq. (6.19).

Several systematics in shape measurements depend on the apparent size of the galaxies. The apparent size can quantified using the circularized radius of the galaxy, $R_{\rm circ}$, defined as

$$R_{\rm circ} = R_{\rm deVauc,r} \times \sqrt{\frac{b}{a}} \tag{6.22}$$

We use the de Vaucouleurs profile fits in the $r$ band to get the circularized radius. $R_{\rm deVauc,r}$ is the de Vaucouleurs semi-major axis of the galaxy, and $b/a$ is the minor to major axis ratio for the de Vaucouleurs fit (already used to calculate the ellipticity).

Li et al. (2013) used de Vaucouleurs shapes to measure the intrinsic alignment signal for the higher-redshift BOSS CMASS sample. One of their systematics tests consisted of randomly permuting the position angle measurements, yet after this permutation, they found that large scale correlations of galaxy position angles persist. They attributed these correlations to the effects of cosmic variance and survey geometry, though these systematics can also arise from the incorrect PSF correction which can introduce such large scale correlations assuming that the PSF is roughly coherent across much of the survey. The SDSS PSF has a preferred direction determined by the scan direction, and thus it is possible for coherent systematics to be present at some level even after randomly permuting the galaxies. We will explore systematic effects in the de Vaucouleurs shapes in detail in Sec. 6.4.

### PSF shapes

PSF models are defined at the position of stars (Lupton et al., 2001a; Stoughton et al., 2002), then interpolated to arbitrary positions. For each frame, the PSF is expanded into Karhunen-Loéve (K-L) basis using stars in the frame and its surrounding neighbours. The K-L models can then be

[3]https:// www.sdss3.org/dr10/algorithms/magnitudes.php

interpolated to the positions of galaxies to provide an image of the PSF at those locations. We use `mE1psf_r` and `mE2psf_r` (defined using adaptive moments) from the SDSS sky server, and rotate them to the coordinate system defined by right ascension and declination. Since these quantities are defined as in Eq. (6.19), we correct them as in Eq. (6.17) using $\mathcal{R} = 1$ to get the PSF shear, $\gamma_{\rm PSF}$.

**Shape measurement systematics**

As described earlier, isophotal shapes are not corrected for PSF effects, de Vaucouleurs shapes are approximately corrected using a double Gaussian PSF model, while re-Gaussianization uses the full PSF model. The PSF causes an overall rounding of galaxy shapes, which if left uncorrected can introduce a multiplicative bias $m$. However, if the (coherent) PSF anisotropies leak into the galaxy shapes, they produce an additive term $a\gamma_{\rm PSF}$ that contributes spurious shape correlations. Thus, individual galaxy shear estimates $\gamma_{\rm obs}$ include a purely random component (shape noise) along with contributions from intrinsic shear $\gamma_I$ and systematics:

$$\gamma_{\rm obs} = (1+m)\gamma_I + a\gamma_{\rm PSF}.$$

These enter the ensemble intrinsic alignments observables as

$$\langle g\gamma_{\rm obs}\rangle = (1+m)\langle g\gamma_I\rangle + a\langle g\gamma_{\rm PSF}\rangle = (1+m)\langle g\gamma_I\rangle \tag{6.23}$$

$$\langle \gamma_{\rm obs}\gamma_{\rm obs}\rangle = (1+m)^2\langle \gamma_I\gamma_I\rangle + a^2\langle \gamma_{\rm PSF}\gamma_{\rm PSF}\rangle. \tag{6.24}$$

Eqns. (6.23) and (6.24) correspond to $\xi_{g+}$ and $\xi_{++}$, respectively, and we have assumed that $\langle\gamma_I\gamma_{\rm PSF}\rangle = \langle g\gamma_{\rm PSF}\rangle = 0$. That is, the intrinsic shear and PSF anisotropies are uncorrelated (since they arise due to completely different physics), and likewise the PSF anisotropies are uncorrelated with galaxy overdensities. Incorrect PSF correction can clearly bias the IA measurements, with multiplicative bias mimicking the behavior of IA amplitude $A_I$ and additive bias adding in a spurious term to shape-shape correlations such as in $\xi_{++}$. A simple way to detect the effect of additive bias is to compute cross-correlations between the PSF and galaxy shapes and normalize by the PSF shape auto-correlation function, which gives an approximately scale-independent ratio. To summarize this in one number, we define

$$A_{\rm PSF} = \frac{\int {\rm d}r_p \langle\gamma_{\rm obs}\gamma_{\rm PSF}\rangle / \langle\gamma_{\rm PSF}\gamma_{\rm PSF}\rangle}{\int {\rm d}r_p} \sim a \tag{6.25}$$

Multiplicative bias, on the other hand, is more difficult to detect using the data alone, and requires detailed characterization for each shape measurement method using simulations, which is beyond the scope of this work.

Besides PSF-related systematics, several other effects such as incorrect sky subtraction and deblending can also lead to spurious alignment signals. Some of these systematics lead to biases similar to those already discussed, while many others could be revealed by systematic tests that we present in Sec. 6.4.2.

**Physical effects**

Aside from systematics, there are physical reasons why different shape measurements can give different results. Several studies of bright elliptical galaxies have found evidence for ellipticity gradients and isophote twisting with radius (see, e.g., Wyatt, 1953; di Tullio, 1978, 1979; Kormendy, 1982; Fasano & Bonoli, 1989; Nieto et al., 1992; Romanowsky & Kochanek, 1998; Lauer et al., 2005; Pasquali et al., 2006). di Tullio (1978, 1979) first found environment-dependence of ellipticity gradients. Galaxies for which the ellipticity increases (decreases) with radius are found preferentially in dense (isolated) environments. Pasquali et al. (2006) confirmed these trends in a sample of 18 elliptical galaxies from the *Hubble Space Telescope* (HST) Ultra-deep Field (UDF). Wyatt (1953) also found evidence of isophote twisting in elliptical galaxies (see also Fasano & Bonoli, 1989; Nieto et al., 1992; Lauer et al., 2005). These observations are consistent with bright ellipticals being triaxial systems with varying

triaxiality as a function of radius, which in projection appears as isophote twisting (e.g., Romanowsky & Kochanek, 1998). However, external tidal fields can also influence the galaxy shapes leading to isophote twisting (Kormendy, 1982), which could mean that the measured intrinsic alignments depend on the effective radius within the galaxy used for measuring the shape. IA measured with shapes that are weighted towards the galaxy centers (outer regions) will be weaker (stronger).

Several studies have also looked at variations in galaxy/halo ellipticity and IA response with the radius using simulations (Schneider et al., 2013; Tenneti et al., 2014b; Velliscig et al., 2015b,a). The consensus has been that halo or galaxy shapes get rounder with increasing radius (for a contradictory result, see Velliscig et al., 2015b) but the outer regions also show stronger IA, primarily because of isophotal twisting.

Different radial weighting in the various shape measurements can in principle allow us to test these variations from data, although interpretation of our results will be complicated by the possibility of systematics such as PSF contamination as discussed earlier. The re-Gaussianization shapes assign higher weights to inner regions of galaxy profiles, while de Vaucouleurs profiles have a comparatively somewhat broader coverage in radius, and isophotal shapes only measure the outer regions of the galaxy light profiles. If tidal fields lead to isophote twisting, making outer regions of the galaxy more intrinsically aligned, we should expect higher IA signal for isophotal shapes followed by de Vaucouleurs and re-Gaussianization shapes.

## 6.4 Results

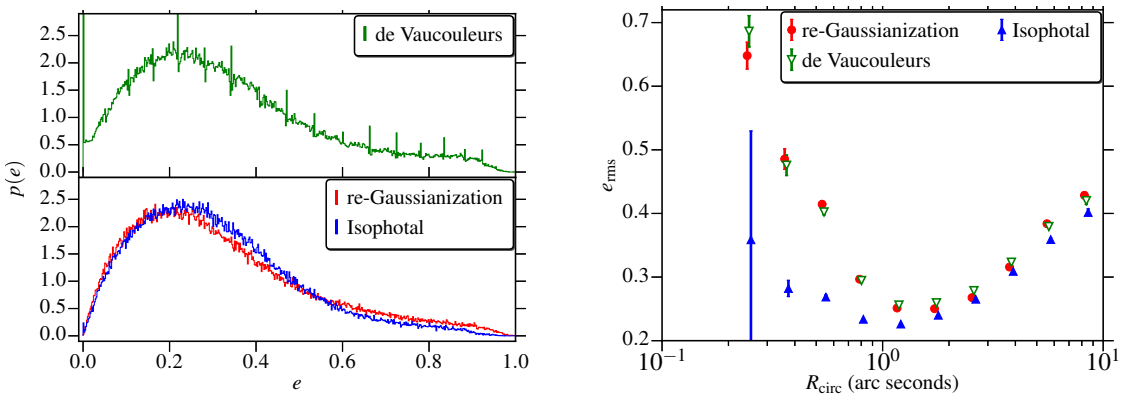
To begin, we explore the features of the galaxy shape distributions (Sec. 6.4.1) and carry out basic systematics tests of intrinsic alignments two-point correlation functions (Sec. 6.4.2) using the different shape measurements. The results in these subsections provide some basic context that will be useful when interpreting the intrinsic alignments results in later subsections. Then we confirm the equivalence of intrinsic alignment two-point statistics using distortions (Eq. 6.16) vs. ellipticities (Eq. 6.19) in Sec. 6.4.3, and compare intrinsic alignments for different shape measurement methods and IA estimators (Sec. 6.4.4). Our tests of the redshift-space $(r_p, \Pi)$ structure of the IA two-point correlations are in Sec. 6.4.5. Finally, we compare with other studies in Sec. 6.4.6.

### 6.4.1 Ellipticity from different shape measurement methods

In this section we study the distortion, $e$, as defined in Eq. (6.16), using re-Gaussianization, isophotal, and de Vaucouleurs shapes. Fig. 6.1a shows the probability distribution $p(e)$, and Fig. 6.1b shows the per-component RMS distortion, $e_{\rm rms}$, as a function of circularized apparent radius $R_{\rm circ}$ for each method. The $p(e)$ for the de Vaucouleurs shapes in Fig. 6.1a has periodic spikes that are likely caused by discretization of the model parameters during the fitting procedure (Lupton et al., 2001a; Stoughton et al., 2002). Similar quantization features are seen in the position angle distribution (not shown). The de Vaucouleurs shapes have the highest RMS distortion ($e_{\rm rms} = 0.261$), followed closely by re-Gaussianization ($e_{\rm rms} = 0.256$), with isophotal shapes ($e_{\rm rms} = 0.241$) giving a noticeably lower RMS distortion. Since the three methods measure galaxy shapes at different effective radii, the straightforward interpretation of this trend is that the galaxy ellipticity first increases and then decreases with radius. However, the quantization in the de Vaucouleurs shapes combined with the fact that (as shown in Sec. 6.4.2) the de Vaucouleurs shapes have significant additive PSF bias casts doubt on this interpretation. It may, however, be a valid interpretation of the differences between re-Gaussianization and isophotal shapes.

The trend for re-Gaussianization vs. isophotal shapes is qualitatively in agreement with hydrodynamic simulations (Tenneti et al., 2015), where the axis ratio increases with radius.

The lower values of $e_{\rm rms}$ in isophotal shapes is primarily driven by the deficit of galaxies with $e > 0.7$, as shown in Fig. 6.1a. This could also be due to some systematic effect such as multiplicative bias from the PSF, which tends to make galaxy shapes rounder and will be more important for more elongated galaxies. The trends in $e_{\rm rms}$ in Fig. 6.1b are consistent with the presence of some




(a) (b)

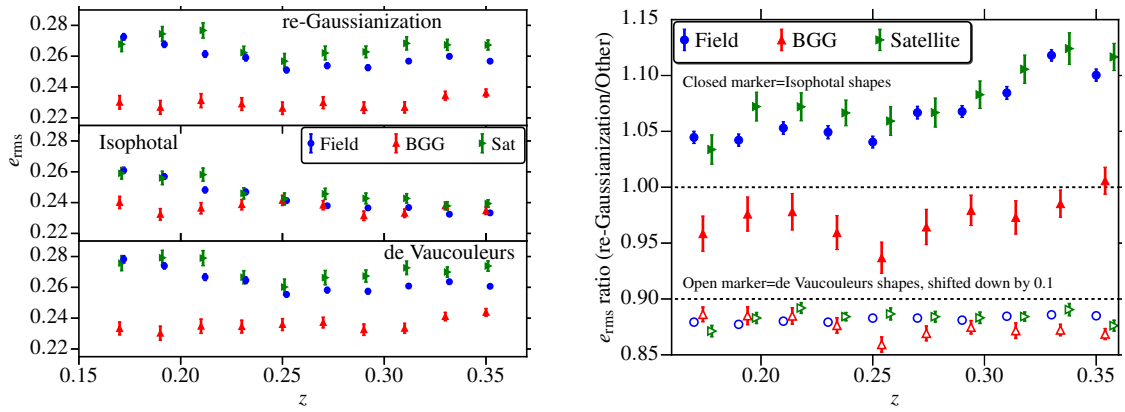
Figure 6.1: (a) Probability distribution of $e$ for different shape measurement methods. The periodic spikes in the de Vaucouleurs $p(e)$ are discussed in the text of Sec. 6.4.1. (b) The RMS ellipticity per shape component, $e_{\rm rms}$, for different shape measurement methods as a function of circularized apparent radius $R_{\rm circ}$ (Eq. 6.22).

systematic bias. The PSF effects become more important for smaller galaxies (smaller $R_{\rm circ}$), thus making them appear rounder and thereby reducing the incidence of high-ellipticity objects. The fact that the isophotal shapes, which have no PSF correction at all, have a lower $e_{\rm rms}$ and that this difference between isophotal and other shapes gets progressively more pronounced for galaxies with a small apparent size is a clear signature of PSF dilution systematics. Note that (for a fixed total flux), the effect of pixel noise increases measurement errors preferentially for the smaller sizes, leading to an increase in the $e_{\rm rms}$ for smaller galaxies for all measurement methods. While we have attempted to subtract the measurement errors, these are known to be underestimated (Reyes et al., 2012) and thus incompletely removed. Fully disentangling true physical effects like ellipticity gradients from systematic effects requires detailed analysis of different shape measurements using simulations, which is beyond the scope of this work. Our statements about radial evolution in ellipticity using isophotal shapes are only valid under the assumption that they are not significantly biased by the PSF or other systematics. In the context of the discussion so far, this is a reasonable assumption since $\sim 90\%$ of galaxies in our sample have $R_{\rm circ} > 1''$.

As described in Sec. 6.3.3, few studies have observed the ellipticity trends in small samples of elliptical galaxies (see for eg. di Tullio, 1978, 1979; Pasquali et al., 2006). di Tullio (1978, 1979) first identified environment dependence of ellipticity gradients. Galaxies that have increasing ellipticity with radius are found preferentially in dense environments like clusters and groups, while those that have decreasing ellipticity with radius were generally isolated though they were also found in groups. We study these trends in the much larger LOWZ sample by identifying BGGs, satellites and field galaxies. Fig. 6.2a shows the $e_{\rm rms}$ for these different environment subsamples for all three shape measurements, and Fig. 6.2b shows the ratio of $e_{\rm rms}$ from re-Gaussianization shapes to isophotal and de Vaucouleurs shapes as a function of redshift. For both re-Gaussianization and de Vaucouleurs shapes, BGGs are rounder than field and satellites galaxies, while the trend is less clear for isophotal shapes. Doing the comparison between different shapes as shown in Fig. 6.2b, BGGs get more elliptical with increasing radius while satellites and field galaxies get rounder, consistent with the observations of di Tullio (1979).

### 6.4.2 Systematics in two-point functions

In this section we show some systematics tests in the intrinsic alignment two-point correlation functions. Fig. 6.3 shows the density-shape ($w_{g+}$) and shape-shape ($w_{++}$) correlations using large $\Pi$




(a) (b)

Figure 6.2: (a) The $e_{\rm rms}$ as a function of redshift for galaxies in different environments using different shape measurement methods. BGGs are rounder than satellite and field galaxies using all three methods, though the differences are much less significant using isophotal shapes. (b) Ratio of $e_{\rm rms}$ using re-Gaussianization shape to isophotal or de Vaucouleurs shapes for galaxies in different environments. BGGs are relatively more elliptical in isophotal shapes while satellites and field galaxies are more round using isophotal shapes. The differences are dominated by ellipticity differences using re-Gaussianization shapes as shown in (a). The variations in ellipticity between re- Gaussianization and de Vaucouleurs shapes are nearly identical across all galaxy environments (open markers in (b)).

| Correlation | Shape | $\chi^2$ | $p-$value |
|---|---|---|---|
| $w_{g+}$ | re-Gaussianization | 17.9 | 0.29 |
| $w_{g+}$ | Isophotal | 17.4 | 0.31 |
| $w_{g+}$ | de Vaucouleurs | 20.4 | 0.19 |
| $w_{++}$ | re-Gaussianization | 9.8 | 0.79 |
| $w_{++}$ | Isophotal | 15.2 | 0.43 |
| $w_{++}$ | de Vaucouleurs | 31.6 | 0.02 |

Table 6.1: $\chi^2$ and $p-$values for $w_{g+}$ and $w_{++}$ calculated using large $\Pi$ separations ($|\Pi| \in [200, 500]$ $h^{-1}$Mpc) and different shape measurements. All signal are consistent with a null detection except for $w_{++}$ with de Vaucouleurs shapes.

separations ($|\Pi| \in [200, 500]$ $h^{-1}$Mpc). For such large separations, we do not expect any contribution from intrinsic alignments, so a deviation of the signals from zero is more likely from additive systematics due to the PSF (in the case of $w_{++}$) or, on small scales, incorrect sky subtraction. These signals are mostly consistent with zero (see Table 6.1 for $\chi^2$ and $p$ values) and thus do not indicate the presence of systematics, except for $w_{++}$ with de Vaucouleurs shapes on large scales. The slight negative signals in $w_{g+}$ are consistent in magnitude and $r_p$-scaling with the presence of a small lensing signal (tangential shape alignments) given the redshift separations between the galaxy pairs and the known lensing signals from Paper I.

As discussed in Sec. 6.3.3, additive PSF contamination in galaxy shape measurements can affect IA results, particularly $w_{++}$, if not corrected properly. To better understand this contamination, we directly calculate $\langle\gamma_{\rm obs}\gamma_{\rm PSF}\rangle$ cross-correlations using PSF shapes at the positions of galaxies. Fig. 6.4 shows the cross-correlations of different shape measurements with the PSF shape, $w_{++}^{\rm sys}$, with de Vaucouleurs shapes showing strong correlations. These are likely the cause of the non-zero $w_{++}$ for large $\Pi$ in Fig. 6.3. Isophotal and re-Gaussianization shapes, on the other hand, show much lower levels of contamination, which suggests that these two are not strongly affected by additive PSF errors. Fig. 6.5 shows $A_{\rm PSF}$ defined in Eq. (6.25) for different galaxy subsamples and shape measurement

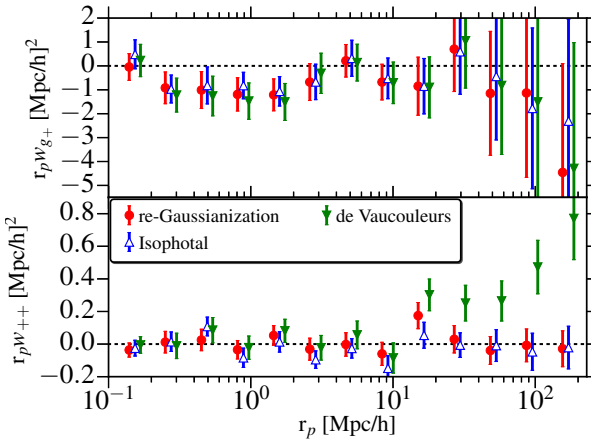



Figure 6.3: Intrinsic alignment correlation functions using integration only over large line-of-sight separations ($|\Pi| \in [200, 500]h^{-1}$Mpc) for the full LOWZ sample, with different shape measurement methods. The top and bottom show $r_p w_{g+}$ and $r_p w_{++}$, respectively.

methods. This figure confirms our conclusion that de Vaucouleurs shapes have stronger additive PSF bias while isophotal and re-Gaussianization shapes have much smaller levels of contamination. Using the full LOWZ sample, we find $A_{\rm PSF} = -0.019 \pm 0.015$, $-0.049 \pm 0.017$, and $-0.29 \pm 0.02$ for re-Gaussianization, isophotal, and de Vaucouleurs shapes. Our value of $|A_{\rm PSF}|$ for re-Gaussianization is smaller than that in Mandelbaum et al. (2015), who found $A_{\rm PSF} \sim -0.1$; however, those results were for a simulated galaxy sample extending to much lower $S/N$ and resolution, which is expected to have relatively stronger additive systematics compared to LOWZ.

It is quite interesting and somewhat counter-intuitive that the de Vaucouleurs shapes, which include PSF correction with an approximate PSF model, have a much larger $A_{\rm PSF}$ than the isophotal shapes, which do not. We propose that the reason for this is that the de Vaucouleurs shapes are weighted towards the central part of the galaxy light profile and thus rely on small scales (which are highly sensitive to the PSF). Thus, a small error in the PSF model can be very important. In contrast, the isophotal shapes use such a low surface-brightness isophote that they correspond to much larger scales, where the impact of the PSF is much smaller, and even without correction, $A_{\rm PSF}$ can be quite small.

There is no strong dependence of the additive PSF contamination on galaxy properties such as luminosity and color. The contamination in de Vaucouleurs shapes does show significant redshift dependence, with the higher redshift sample, Z2, showing stronger contamination; a similar but much less significant effect is present in isophotal and re-Gaussianization shapes. This is likely due the fact that higher redshift galaxies have a smaller apparent size, making the contamination from PSF anisotropy more important for such galaxies. There are hints of luminosity dependence to the additive PSF contamination for both de Vaucouleurs and isophotal shapes, though these trends are not very significant. $|A_{\rm PSF}|$ is expected to be larger for lower luminosity samples since those galaxies generally have smaller apparent size, making PSF contamination more important.

We caution that the tests in this section do not rule out multiplicative bias, which could change the IA two-point correlation functions in a way that is degenerate with the IA amplitude $A_I$.

### 6.4.3 Ellipticity definition

To confirm the consistency of intrinsic alignments results using the two different ellipticity definitions in Eq. (6.16) and (6.19), we calculate the density-shape correlation function $w_{g+}$ with isophotal and de Vaucouleurs shapes using both definitions, modifying the $w_{g+}$ estimator appropriately. Fig. 6.6 shows the ratio of the inferred $A_I$ for different subsamples from the NLA model fits to $w_{g+}$ using

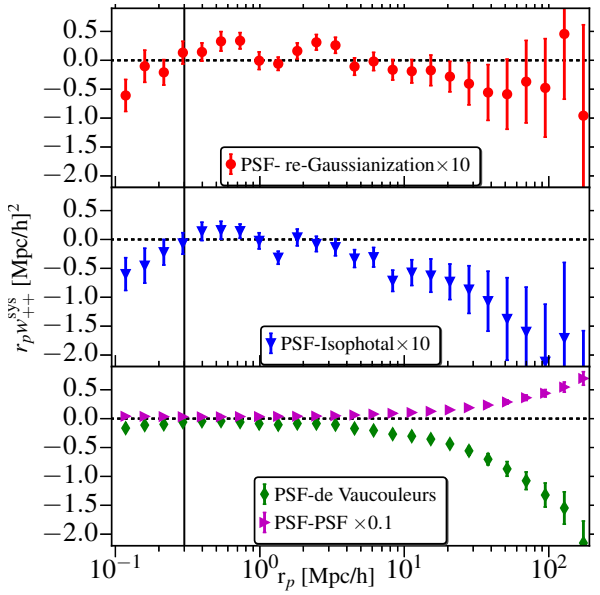



Figure 6.4: Galaxy shape vs. PSF shape cross-correlations using the full LOWZ sample with different shape measurement methods, as a way to reveal additive PSF contamination in the galaxy shapes. The signal for re-Gaussianization and isophotal shapes has been multiplied by a factor of 10 in this plot, while the signal for PSF-PSF correlations is suppressed by a factor of 10. The black solid line marks the SDSS fiber collision limit.

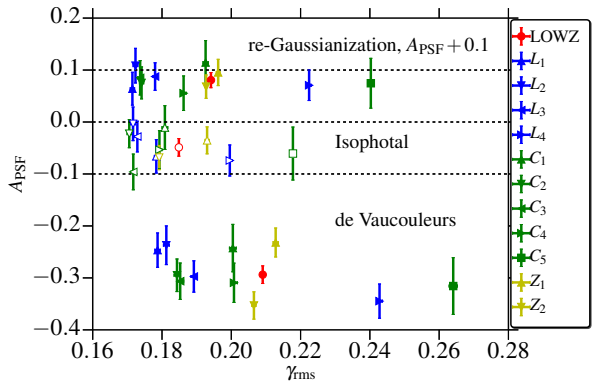



Figure 6.5: Galaxy shape vs. PSF shape cross-correlation amplitude, $A_{\rm PSF}$ (Eq. 6.25), using isophotal (open markers), re-Gaussianization (filled markers, $A_{\rm PSF}+0.1>0$, shifted up for clarity) and de Vaucouleurs (filled markers, $A_{\rm PSF}<-0.1$) shapes. re-Gaussianization and isophotal shapes show relatively low and similar amplitudes while de Vaucouleurs shapes have a large value of $A_{\rm PSF}$, indicative of significant additive bias from the PSF anisotropy.

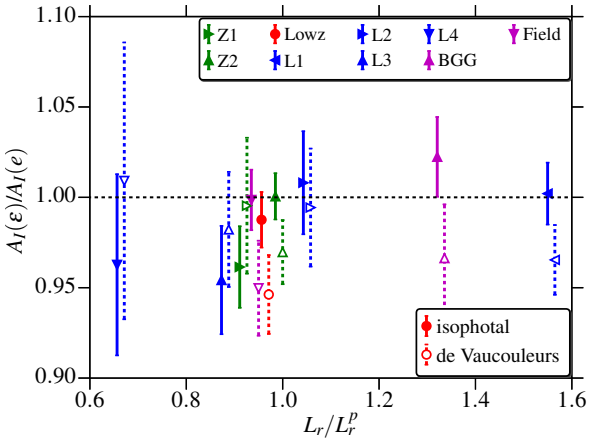



Figure 6.6: Ratio of IA amplitude, $A_I$, obtained via NLA model fits to $w_{g+}$ calculated using shears from two different ellipticity definitions, $e$ (Eq. 6.16) and $\varepsilon$ (Eq. 6.19), for isophotal (solid lines) and de Vaucouleurs (dashed lines) shapes. Both definitions give consistent results across all luminosities, redshifts, and environments.

the two shear estimators. $A_I$ measured using both definitions are consistent within $1\sigma$ for isophotal shapes, and there are no strong dependences on galaxy properties such as color or luminosity. In principle, this is precisely as expected; however, differences could arise due to systematic errors in the RMS ellipticity calculation if the error estimates are incorrect. For the rest of the paper, we use $\varepsilon$ (Eq. 6.19) as our galaxy shape estimator for isophotal and de Vaucouleurs shapes.

### 6.4.4 IA with different shape measurements

In this section we present the IA measurements using the different shape measurement methods discussed in Sec. 6.3.3.

Fig. 6.7a shows the $w_{g+}$ measurement for the full LOWZ sample using all three shape measurement methods. Isophotal shapes give the highest IA amplitude, followed by de Vaucouleurs and re-Gaussianization shapes. Fig. 6.8 shows the amplitude trends more clearly, where we have plotted the ratio of $A_I$ for isophotal and de Vaucouleurs shapes with respect to re-Gaussianization shapes, for various subsamples defined by different galaxy properties. Isophotal (de Vaucouleurs) shapes give a higher $A_I$ by $\sim 40\%$ (20%) compared to re-Gaussianization shapes, with no clear trends with galaxy properties like luminosity, color, and redshift.

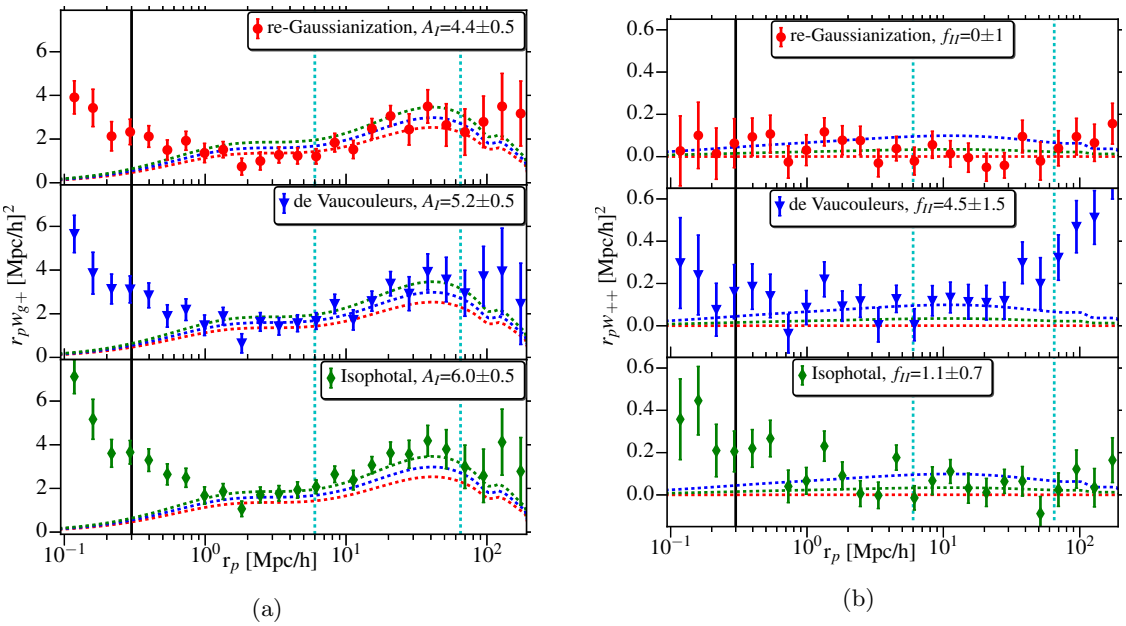
To further understand the effects of different shape measurement methods, we measure two more IA estimators, $\langle\gamma\rangle$ and $\langle\theta\rangle$, defined as

$$\langle\gamma\rangle = \frac{S_+ D}{SD} \tag{6.26}$$

$$\langle\theta\rangle = \frac{\theta_{SD}}{SD} \tag{6.27}$$

Both $\langle\gamma\rangle$ and $\langle\theta\rangle$ are calculated in a single $\Pi$ bin with $\Pi \in [-100, 100]h^{-1}$Mpc. In the absence of intrinsic alignments, $\langle\gamma\rangle = 0$ and $\langle\theta\rangle = 45°$, while in presence of IA, $\langle\gamma\rangle > 0$ and $\langle\theta\rangle < 45°$ Fig. 6.9 shows the ratio of $\langle\gamma\rangle$ and $45° - \langle\theta\rangle$ measured using isophotal and de Vaucouleurs shapes with respect to re-Gaussianization shapes. $\langle\gamma\rangle$ is higher for isophotal shapes followed by de Vaucouleurs shapes, as expected from $w_{g+}$ results. $45° - \langle\theta\rangle$ is also higher.

There are a few possible explanations for these differences in IA amplitudes with different shape measurement methods. The first possibility, as discussed in Sec. 6.3.3 and observed in Sec. 6.4.2, is a systematic error from incorrect PSF removal. Isophotal shapes are not corrected for the PSF, while de




(a) (b)

Figure 6.7: (a) The projected galaxy density-shape correlation function $w_{g+}$ (Eq. 6.8) using the full LOWZ sample and different shape measurement methods, isophotal (blue), re-Gaussianization (red) and de Vaucouleurs shapes (green), along with the best-fitting NLA models. For comparison, the best-fitting NLA models for all three shape measurement methods are plotted on all panels. Isophotal shapes show the highest IA signal, followed by de Vaucouleurs and re-Gaussianization shapes (note that error bars are correlated between the different shape measurement methods). The solid black line shows the SDSS fiber collision limit, and the dashed cyan lines show the range of $r_p$ used for the NLA model fitting. (b) Same as (a), but for the projected shape-shape correlation function $w_{++}$ (Eq. 6.9).

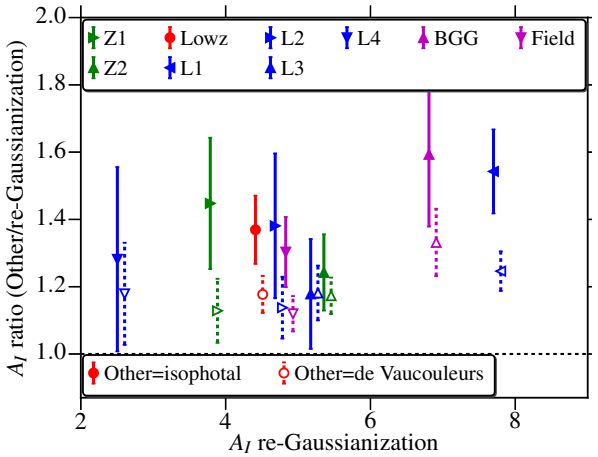



Figure 6.8: Comparison of NLA model amplitude $A_I$ for isophotal and de Vaucouleurs vs. re-Gaussianization shape measurements. Isophotal (de Vaucouleurs) shapes consistently give a higher amplitude by $\sim 40\%$ (20%) compared to re-Gaussianization shapes. The de Vaucouleurs results have been shifted horizontally for clarity.

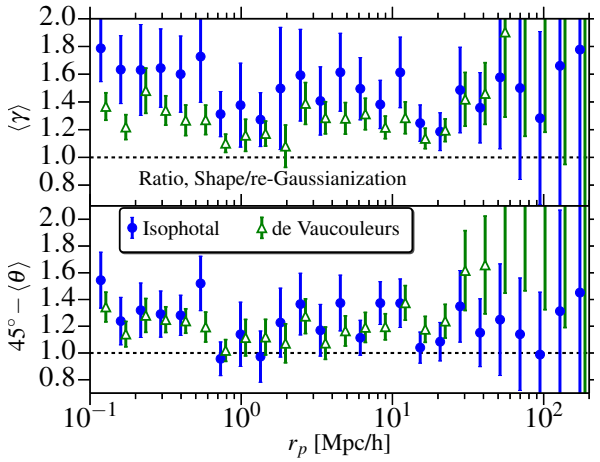



Figure 6.9: Ratio of mean intrinsic shear, $\langle\gamma\rangle$ (Eq. 6.26) and mean alignment angle, $45^\circ - \langle\theta\rangle$ (Eq. 6.27) using different shape measurement methods.

Vaucouleurs shapes are only approximately corrected. As shown in Sec. 6.4.2, we have a clear detection of additive PSF bias in de Vaucouleurs shapes, though isophotal and re-Gaussianization shapes have much lower levels of contamination. The additive bias should drop out of $w_{g+}$ calculations, under the assumption that it is not correlated with the positions of other galaxies ($\langle g\gamma_{\rm PSF}\rangle \sim 0$ in Eq. 6.23). We have not ruled out the presence of multiplicative bias, which could cause an apparent change in $A_I$, but is extremely difficult to rule out using the data alone. However, the simplest possible interpretation of how multiplicative bias should affect these results is not valid: if multiplicative bias is responsible for the lower RMS ellipticities using isophotal shapes compared to other methods in Fig. 6.1b, then $A_I$ should be the lowest using isophotal shapes, not the highest.

Another possibility as discussed in Sec. 6.3.3 is the presence of some physical effect such as isophote twisting and/or ellipticity gradients. Since $w_{g+}$ is an ellipticity-weighted measure of IA, gradients in the ellipticity can also lead to higher $w_{g+}$. However, as discussed in Sec. 6.4.1, isophotal shapes have lower $e_{\rm rms}$ than re-gaussianization and de Vaucouleurs shapes. This suggests that if the galaxies have the same large-scale alignments with the tidal field at all radii, but an ellipticity gradient that modifies $e_{\rm rms}$, then isophotal shapes should have lowest $w_{g+}$ amplitude, not the highest. Moreover, ellipticity gradients alone cannot explain a stronger alignment angle $\langle\theta\rangle$ for isophotal shapes, as shown in Fig. 6.9.

The other physical effect, isophote twisting, is related to the fact that outer regions of galaxies may be more susceptible to the local and large-scale tidal fields and thus show stronger alignments. If this effect is important, isophotal shapes should have the highest IA amplitude, followed by de Vaucouleurs and finally re-Gaussianization shapes, consistent with our results. The results in Fig. 6.9 are also consistent with this physical effect. Given that PSF-related systematics should cancel out of $\langle\theta\rangle$ (since the alignment is being calculated with respect to galaxy positions that are not aligned with respect to the PSF), and multiplicative biases also cannot modify the alignment angle, it is difficult to interpret that finding as anything other than isophote twisting.

Finally, Fig. 6.7b shows the shape-shape ($w_{++}$) correlation functions. The $w_{++}$ detection using isophotal and re- Gaussianization shapes is not very significant, and most of the signal using de Vaucouleurs shapes can be attributed to additive PSF bias as shown in Sec. 6.4.2. The detection significance of our isophotal measurements is not consistent with that of Blazek et al. (2011), who reported a high-significance detection of $w_{++}$ using measurements of Okumura et al. (2009) with isophotal shapes. While this could also be due to the differences in sample definition, our results in Paper I suggest that even with an appropriately bright subset of LOWZ, our results may not agree. Thus, we will address this discrepancy in greater detail in Sec. 6.4.6 after we have discussed the

anisotropy of IA signals, which turns out to be an important factor in this difference.

### 6.4.5 Anisotropy of IA

As discussed in Sec. 6.2.3, IA measurements suffer from anisotropy in $(r_p, \Pi)$ introduced by the projected shapes, which do not allow measurement of the IA signal along the line of sight. To enable a study of this effect, Fig. 6.10 shows $\xi_{gg}$, $\xi_{g+}$, and $\xi_{++}$ as a function of $r_p$ and $\Pi$, as well as the corresponding monopole and quadrupole measurements defined in Sec. 6.2.3. The top row shows the galaxy clustering measurements as well as the model predictions. Non-linear theory predictions with the Kaiser formula for RSD match the data for $r_p \gtrsim 5h^{-1}$Mpc and $s \gtrsim 20h^{-1}$Mpc, though there are deviations below these scales due to nonlinear RSD and the Finger-of-God effect.

The middle row in Fig. 6.10 shows $\xi_{g+}$ measurements using isophotal shapes. Again, the NLA model fits the data well for $r_p \gtrsim 5h^{-1}$Mpc and $s \gtrsim 20h^{-1}$Mpc with deviations at small scales due to nonlinear RSD and clustering. The strong anisotropy introduced by the projected shapes is clearly visible in the shape of the peanut-shaped contours, where the signal drops off quickly with $\Pi$. The NLA model incorporates this anisotropy and is consistent with the data. The redshift-space structure of $\langle\gamma\rangle$ and $\langle\theta\rangle$ (shown in Fig. 6.11) is very similar to that of $\xi_{g+}$.

The bottom row in Fig. 6.10 shows $\xi_{++}$ measurements using isophotal shapes. To display NLA model predictions, we use the best-fitting parameters from fitting $w_{g+}$, with $f_{II} = 1$ (solid lines) and $f_{II} = 2$ (dashed lines). The two-dimensional contours in Fig. 6.10e suggest that the data prefer the model with $f_{II} = 2$. However, these contours are quite noisy, so Fig. 6.10e is not a reliable test of the validity of the model. In Fig. 6.10f, we show a clear detection of the monopole for $\xi_{++}$, with the data again preferring a higher amplitude than predicted by NLA model ($f_{II} > 1$). This discrepancy could either be from the effects of non-linear physics that is not included in the NLA model (Blazek et al., 2015), or from additive PSF contamination. Even though additive PSF contamination was shown to be low for isophotal shapes ($|A_{PSF}| \sim 0.05$), the contamination in $\xi_{++}$ could still be strong enough to increase the observed $\xi_{++}$ amplitude. In Fig. 6.12, we show the $\xi_{++}$ monopole for the full LOWZ sample and the brightest subsample, $L_1$, along with the predicted PSF contamination. For large scales, $s > 30h^{-1}$Mpc, the PSF contamination alone can account for the observed $\xi_{++}$ signal for the full LOWZ sample. At smaller scales, the PSF effects are subdominant, which make it unlikely to be responsible for the higher than predicted amplitude. Both LOWZ and $L_1$ samples prefer a higher amplitude; based on $\Delta\chi^2$, the $p-$value$= 0.01$ $(0.15)$ for LOWZ ($L_1$). However, the $f_{II} = 1$ model is not ruled out by either sample, with a $p$-value of 0.07 (0.44) for LOWZ ($L_1$). With PSF effects being subdominant, we conclude that the discrepancy between NLA and data is primarily due to non-linear physics (e.g. non-linear clustering, non-linear RSD), which become important for scales $s < 30h^{-1}$Mpc and generally increase the amplitude (Blazek et al., 2015).

The $\xi_{++}$ quadrupole moment in both the data and the NLA model prediction is close to zero. $\xi_{++}$ is relatively more isotropic (Croft & Metzler, 2000) than $\xi_{g+}$ as some large-scale modes along the line-of-sight can introduce similar IA in projected galaxy shapes with larger separation along $\Pi$. These correlations show up in the shape-shape correlation function but not in the density-shape correlations. Mathematically, these terms are sourced by the $J_0$ term combined with $(1-\mu^2)^2$ (see Eq. 6.13) which is more extended along $\Pi$, while the $J_2$ and $J_4$ terms are more extended along $r_p$. The combination of $J_0$ and $J_4$ terms makes $\xi_{++}$ relatively more isotropic than $\xi_{g+}$.

As a final exploration of the anisotropy of intrinsic alignments, Fig. 6.13 shows $\xi_{gg}$ and $\xi_{g+}$ measured using the MassiveBlack-II (MB-II) cosmological hydrodynamic simulation (Khandai et al., 2015) at $z = 0.3$, along with NLA model predictions. The top row shows the signal without RSD, along with model predictions with $\beta = 0$. Bottom row shows both signal and model with RSD effects included. The effects of RSD are clearly visible in $\xi_{gg}$. On the other hand, the variations in $\xi_{g+}$ from top to bottom row are far less pronounced, consistent with our findings that $\xi_{g+}$ is not affect by RSD to first order, $\beta_\gamma = 0$ (see Paper I for derivation). Qualitatively, the simulation results produce the features seen in the data and show good agreement with the model within the limitations of validity of the model (which does not include nonlinear RSD or nonlinear galaxy bias, resulting in small-scale discrepancies). At large scales ($r \gtrsim 20h^{-1}$Mpc) cosmic variance plays an important role

due to the limited size of the simulation box ($100h^{-1}$Mpc), making it difficult to compare the model and simulation data at these scales.

### 6.4.6 Comparison with other studies

As discussed in Paper I and Sec. 6.4.4, there is an apparent discrepancy in detection significance of our $w_{++}$ measurements with those of Blazek et al. (2011). Blazek et al. (2011) used a projection of 3D results from Okumura et al. (2009), who measured $\xi_{++}$ ($C_{11}$ in their notation) as function of redshift-space separation $s$ using isophotal shapes. We hereafter use $C_{11}$ to refer to $\xi_{++}(s)$, to distinguish it from our measurement of $\xi_{++}(r_p, \Pi)$. Assuming isotropy of $\xi_{++}(s)$, Blazek et al. (2011) projected the $C_{11}$ measurement onto $(r_p, \Pi)$ space in order to calculate $w_{++}$. Their error estimates were based on generating 1000 random realizations of $C_{11}$, assuming Gaussian and independent errors. To test whether any parts of this procedure (either the signal or error estimation) could lead to differences in the estimated $S/N$ compared to our procedure, we repeat their analysis for our LOWZ sample, using isophotal shapes.

Fig. 6.14 shows our measurement of $C_{11}(s)$. With isophotal shapes, we have a significant detection of shape-shape correlations (as do Okumura et al. 2009 for LRGs), whereas the detection is not significant using re-Gaussianization shapes. We do not attempt a quantitative comparison with their results, since differences in sample selection complicate the comparison.

Using this $C_{11}(s)$ measurement with isophotal shapes, we repeat the procedure of Blazek et al. (2011) to get $w_{++}(r_p)$. We project the $C_{11}$ measurement onto the $(r_p, \Pi)$ plane, using 200 $\Pi$ bins with $\Pi \in [-100, 100]h^{-1}$Mpc and $d\Pi = 1h^{-1}$Mpc. Fig. 6.15 compares the $w_{++}$ obtained using different methods. The $w_{++}$ for the LOWZ sample calculated using the Blazek et al. (2011) method of projecting $C_{11}(s)$ does have a higher detection significance compared to the $w_{++}$ using our $\xi_{++}(r_p, \Pi)$ projection. This result is most likely due to the fact that $C_{11}$ is effectively the monopole term, which we detect with high significance (Fig. 6.10f). $C_{11}$ and the monopole have different amplitudes due to an additional $(2l+1)/2$ factor in the multipole moments (see Eq. 6.15). The differences in $S/N$ between $C_{11}$ and $w_{++}$ primarily come from the different sensitivity of the two estimators to the large $\Pi$ regions that contribute little signal but significant noise. In accordance with this explanation, we find that using $\Pi_{\rm max} \sim 30h^{-1}$Mpc when computing $w_{++}$ does give a statistically significant measurement (not shown). To compare error estimates, we repeat the Blazek et al. (2011) analysis but obtain errors using 100 jackknife regions instead of using 1000 random realizations of $C_{11}(s)$. The errors obtained using both methods are consistent. Note that since the LRG sample used by Okumura et al. (2009) is brighter than the LOWZ sample, the amplitude for the Blazek et al. (2011) $w_{++}$ measurement is expected to be higher than for LOWZ. Also, the choice of $d\Pi = 1\ h^{-1}$Mpc leads to an approximately constant $w_{++}$ for $r_p < 1h^{-1}$Mpc since all bins are dominated by $C_{11}$ values from $s \sim 1h^{-1}$Mpc, given that the signal drops exponentially with increasing $s$ and bins with large $\Pi$ do not contribute much. Hence, the apparently constant $w_{++}$ at $r_p < 1h^{-1}$Mpc in Fig. 6.15 is not physical.

In Fig. 6.15 we also show the theory prediction from the best-fitting NLA model to $w_{g+}$. As in Fig. 6.12, the data prefer a higher amplitude than predicted by the NLA model, likely due to the effects of non-linear physics beyond the NLA model. This result is inconsistent with the findings of Blazek et al. (2011), who found consistent IA amplitudes from both $w_{g+}$ and $w_{++}$. Their sample, however, is similar to our $L_1$ sample, which as shown in Fig. 6.12 is consistent with $f_{II} = 1$.

To summarize, we have investigated three possible reasons for the apparent difference in $w_{++}$ detection significance in Blazek et al. (2011) compared to Paper I. We find that the primary causes are (a) use of isophotal rather than re-Gaussianization shapes (increased signal), and (b) the projection of $C_{11}(s)$ rather than $\xi_{++}(r_p, \Pi)$ (which ignores noisy higher multipoles), but conclude that their error estimate procedure is in agreement with the jackknife procedure. Hence the different $S/N$ in $w_{++}$ arises from differences in signal estimation (the first of which increases the signal, the second of which lowers the noise).

Several other studies have used de Vaucouleurs and isophotal shapes to measure IA using the bright, nearby SDSS Main sample and the fainter, more distant BOSS CMASS samples (Hao et al., 2011; Li et al., 2013; Zhang et al., 2013). These samples have different redshifts and galaxy properties

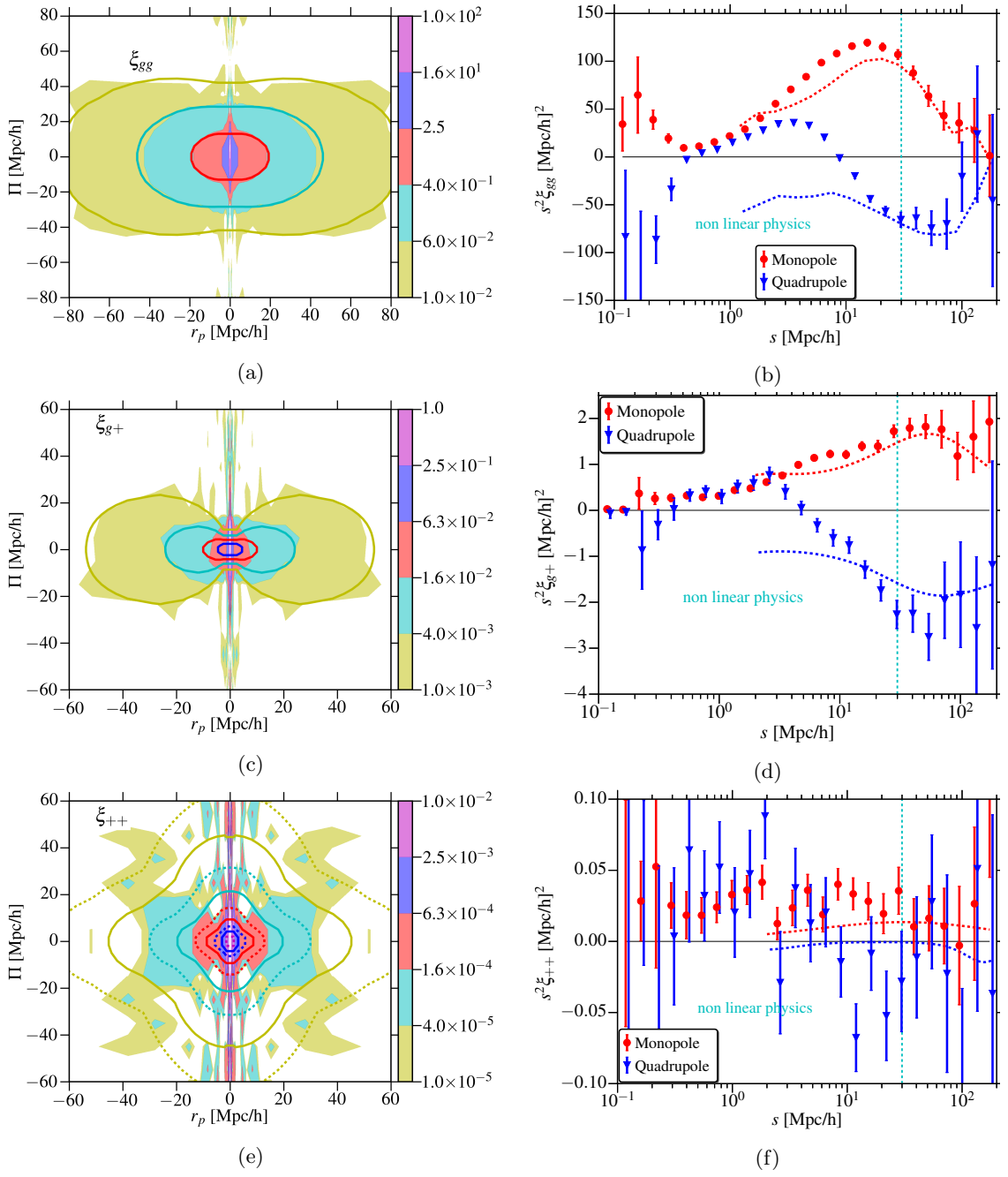
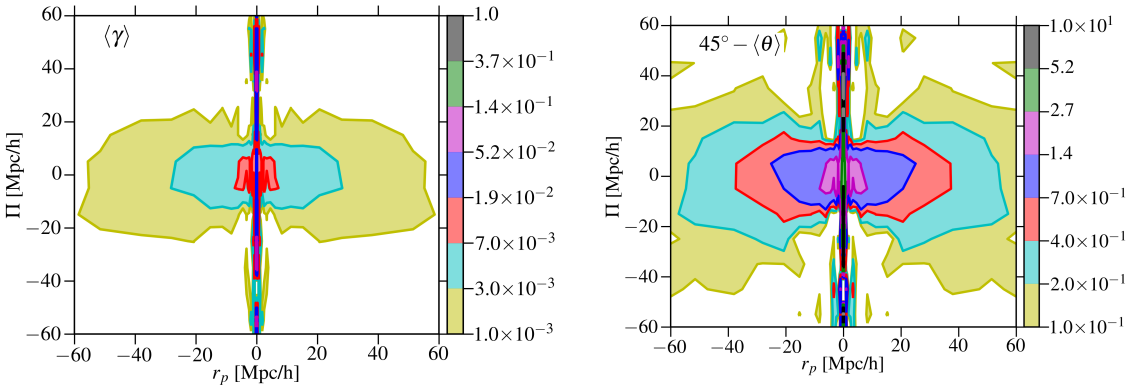



Figure 6.10: The 3D galaxy-galaxy correlation function $\xi_{gg}$ (top row, Eq. 6.11), galaxy density-shape correlation function $\xi_{g+}$ (middle row, Eq. 6.12) and shape-shape correlation function $\xi_{++}$ (bottom row, Eq. 6.13) as a function of $r_p, \Pi$ (left column, reflected about $r_p = 0$) and their multipole moments (right column). All plots use isophotal shapes. In the left column, the filled contours are showing the data, while solid lines are the theory predictions corresponding to the outer edge of the filled contours. The right column shows monopole and quadrupole measurements as a function of three dimensional redshift space separation ($s = \sqrt{r_p^2 + \Pi^2}$ [Mpc/$h$]). The points are measurements from the data, while dashed lines are NLA model predictions. Theory predictions in both columns are from the best-fitting models to $w_{gg}$ and $w_{g+}$, with $f_{II} = 1$ in $\xi_{++}$. The dashed lines in (e) show predictions with $f_{II} = 2$. The linear models with non-linear power spectrum are consistent with data for $s \gtrsim 30\ h^{-1}$Mpc, below which significant deviations are expected due to non-linear RSD and, on even smaller scales, non-linear galaxy bias.




(a) (b)

Figure 6.11: Similar to the left column of Fig. 6.10. (a) $\langle\gamma\rangle$ (Eq. 6.26) as a function of $r_p, \Pi$. (b) $45^\circ - \langle\theta\rangle$ (Eq. 6.27) as a function of $r_p, \Pi$. Both (a) and (b) use isophotal shapes and the same binning as in Fig. 6.10, but there are no theory curves on these plots.

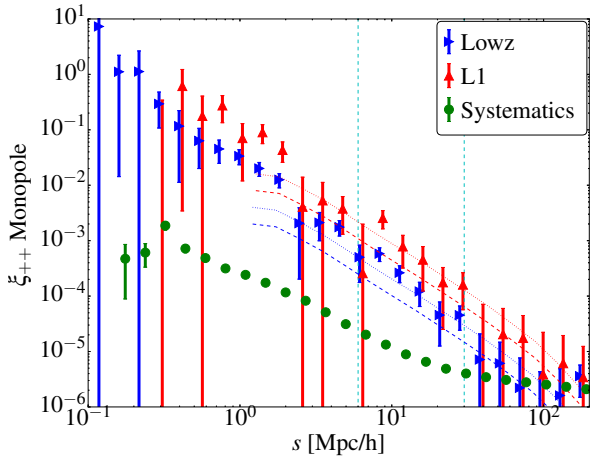



Figure 6.12: The 3D shape-shape correlation function $\xi_{++}$ (Eq. 6.13) monopole moment using isophotal shapes for LOWZ (blue) and the $L_1$ (red) subsample, along with theory predictions from the best-fitting NLA model to $w_{g+}$ with $f_{II} = 1$ (dashed lines) and $f_{II} = 2$ (dotted lines). Green points also show the monopole of the expected additive PSF contamination based on the analysis in Sec. 6.4.2, $\xi_{++}^{\rm sys} \approx A^2_{\rm PSF,isoph} \times \xi_{++}^{\rm PSF\text{-}PSF}$. Vertical lines at $s = 6$ and $30h^{-1}$Mpc show the limits used to calculate $\chi^2$ values to assess the goodness of the model.

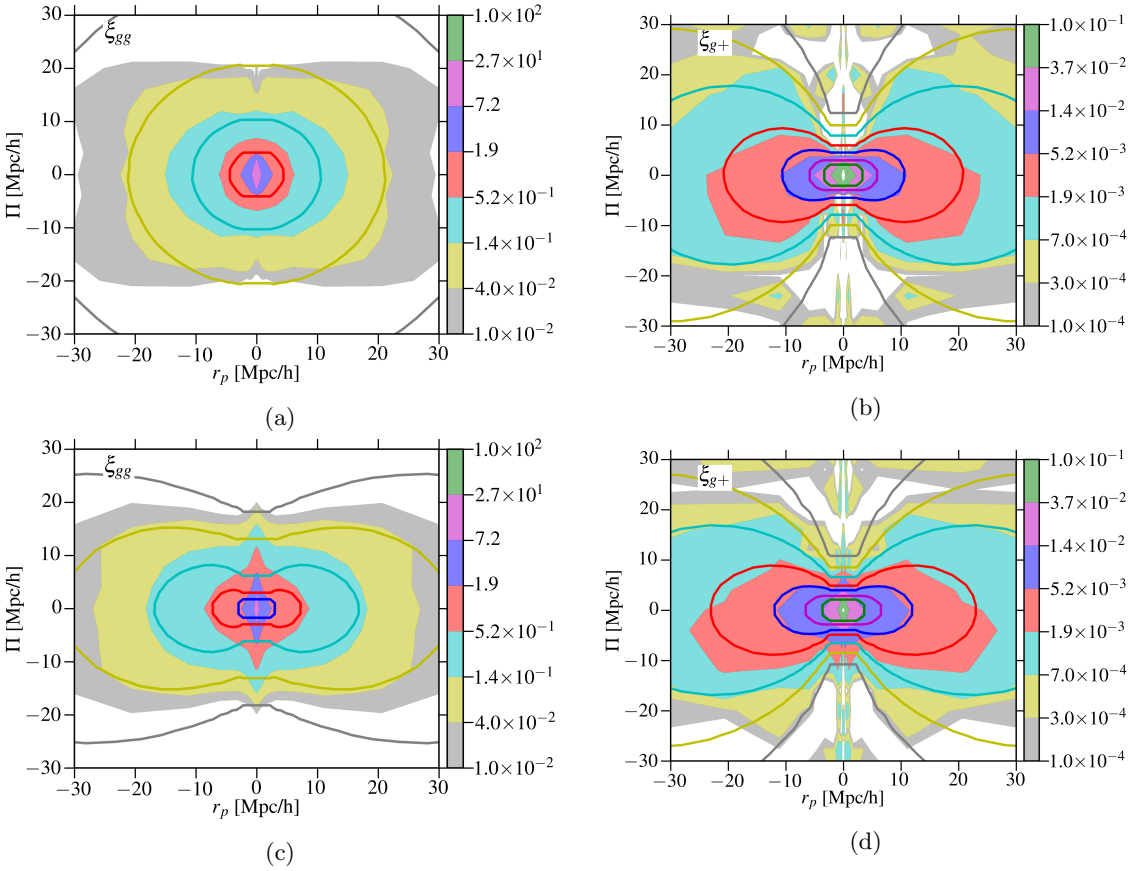



Figure 6.13: The 3D galaxy-galaxy correlation function $\xi_{gg}$ (Eq. 6.11) and galaxy density-shape correlation function $\xi_{g+}$ (Eq. 6.12) contour plots from the MB-II simulation, similar to Fig. 6.10a and Fig. 6.10c. The top row shows the signal and model without RSD, while the bottom row includes RSD. We used linear bias $b = 0.8$ and $A_I = 6$ for the theory predictions. Note that the small scale behavior in (a,b) is an artifact from binning, with logarithmic bins in $r_p$ and linear bins in $\Pi$ with $d\Pi = 5h^{-1}$Mpc.

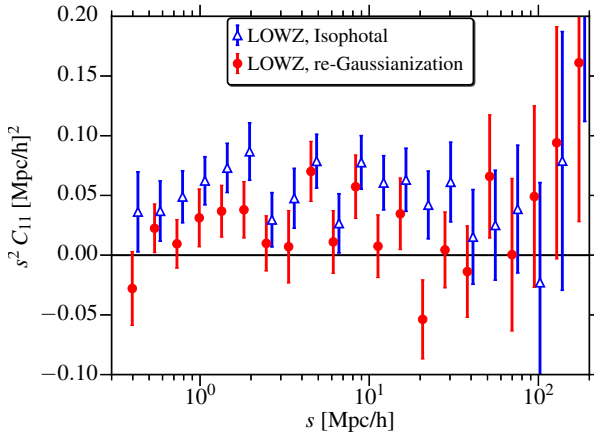



Figure 6.14: The 3D shape-shape correlation function as a function of redshift-space separation $s$, $C_{11}(s)$ (or $\xi_{++}(s)$), for the entire LOWZ sample, using re-Gaussianization and isophotal shapes.

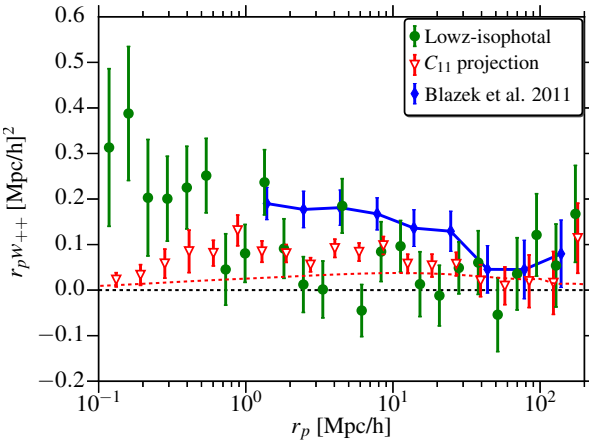



Figure 6.15: Comparison of $w_{++}$ measured in different ways. The green points show the signal after measuring $\xi_{++}$ in $(r_p, \Pi)$ space and projecting along $\Pi$ using our standard methodology, while the red points show the signal obtained by projecting $C_{11}(s)$ following the approach of Blazek et al. (2011). The flatness of $w_{++}$ using the second approach at $r_p \lesssim 1h^{-1}$Mpc is due to the choice of bin sizes ($d\Pi = 1h^{-1}$Mpc); see text for details. The signal for brighter LRGs from Blazek et al. (2011) signal is plotted in green for reference, but should not be compared in detail with our results due to differences in sample definition.

compared to LOWZ. As a result, the galaxy detection have different $S/N$ and resolution compared to the PSF, which should modify the observational systematics in galaxy shapes. Thus, our results about differences in systematics and in the IA amplitude using these different shape measurement methods cannot be used to make definitive statements about systematic effects in those studies.

## 6.5 Conclusions

In this work, we have studied SDSS-III BOSS LOWZ galaxy shapes and intrinsic alignments using three different shape measurement methods: re-Gaussianization (PSF-corrected, weighted towards inner regions of galaxies), isophotal (based on the shape of a low surface brightness isophote at large radius, not PSF-corrected), and de Vaucouleurs shapes (from fitting a de Vaucouleurs model to the light profile using an approximate PSF model, using a grid-based procedure that results in some quantization of model parameters). Different shape measurement methods give a different ellipticity for galaxies, which in the absence of systematic error implies a radial gradient in the galaxy shapes, with the shapes becoming rounder on average at large radius. These variations in ellipticity seem to depend on the galaxy environment, with brightest group galaxies (BGGs) actually becoming more elliptical with radius but satellites and field galaxies becoming rounder. The overall sign of the ellipticity gradients is consistent with hydrodynamic simulations (Tenneti et al., 2015) and the environment trends are consistent with those seen using a small sample of elliptical galaxies (di Tullio, 1978, 1979). We caution, however, that the isophotal shapes on which these conclusions rest are not corrected for the PSF. Biases from the PSF (of which we see hints in Fig. 6.1b and 6.12) and other observational systematics (e.g., contamination from light in nearby galaxies around BGGs) can alter these conclusions.

Tests for systematics do not reveal a level of multiplicative or additive bias in either re-Gaussianization or isophotal shapes that could significantly change the measured intrinsic alignment statistics for LOWZ galaxies on scales up to a few tens of Mpc. In contrast, tests for additive systematic errors in the shape-shape correlation functions reveal that the de Vaucouleurs shapes are significantly affected by additive PSF bias. The magnitude of the systematic in the de Vaucouleurs shapes is consistent

with $\sim 30\%$ of the PSF anisotropy leaking into the galaxy shapes as an additive term.

A comparison of the density-shape correlations ($w_{g+}$) using the different shape measurement methods revealed that isophotal (de Vaucouleurs) shapes give $\sim 40$ (20)% higher NLA model amplitude, $A_I$, compared to re-Gaussianization shapes. Since isophotal shapes are slightly rounder on average, this finding cannot be easily explained in terms of multiplicative bias. These differences in the IA results may imply isophote twisting of galaxy shapes to make the outer regions more aligned with the tidal field, consistent with theoretical predictions (Kormendy, 1982; Kuhlen et al., 2007). We emphasize that our conclusions may not carry over to studies that use significantly fainter or less well-resolved galaxy samples (e.g., the BOSS CMASS sample), where systematic errors are likely to be more important. Use of a suite of systematics tests as shown in this paper can be helpful to reveal problems; however, the issue of multiplicative bias will be difficult to completely resolve from the data alone.

We also studied the anisotropy of IA signal as a function of $r_p$ and $\Pi$, finding that NLA model predictions and those from hydrodynamic simulations are consistent with the observations. The projection factor from 3D to 2D shapes is the dominant source of anisotropy in the 3D density-shape correlations $\xi_{g+}$, resulting in peanut-shaped contours in the $(r_p, \Pi)$ plane.

Finally, we investigated the difference in shape-shape intrinsic alignments ($w_{++}$) detections in Blazek et al. (2011) vs. Paper I, and identified two significant sources of differences: use of isophotal rather than re-Gaussianization shapes, and estimation of the signal via projection of $C_{11}(s)$ under the assumption of isotropy in the $(r_p, \Pi)$ plane rather than via direct projection of $\xi_{++}(r_p, \Pi)$.

Our results have implications for intrinsic alignments forecasting, i.e., the prediction of IA contamination in weak lensing measurements, and for intrinsic alignments mitigation with future surveys. First, in the case of forecasting, the relevant point is the large variation (up to 40%) in the NLA model amplitude that is inferred from $w_{g+}$ measurements using different shape measurement methods. What is the appropriate amplitude to use for forecasts? We argue that the relevant one to use depends on the shape measurement method used for estimation of shear in the weak lensing survey for which forecasts are being done. If using a more centrally-weighted method, our results suggest that a lower IA amplitude is more appropriate and will give more accurate forecasts.

Second, mitigation of IA may involve joint modeling of IA and lensing in measurements of shape-shape, galaxy-shape, and galaxy-galaxy correlation functions (e.g., Joachimi & Bridle, 2010). Joint modeling efforts require a model for intrinsic alignments as a function of separation, redshift and galaxy properties such as luminosity, with priors on these model parameters. Our results suggest that the model parameters should have broad enough priors chosen within a range selected specifically taking into account the radial weighting of the shape measurement method used for shear estimation. In fact, in the limit that the weak lensing analysis is carried out using two shear estimation methods for a consistency check, it is plausible that the best-fitting IA parameters inferred using methods that are more or less centrally-weighted may differ.

Finally, the estimators for measuring IA should also be chosen carefully. For shape-shape correlations, $C_{11}$ is probably a better estimator than $w_{++}$. In general, future studies should look at the full structure of IA in the $(r_p, \Pi)$ or $(r, \mu)$ plane and tailor the estimator accordingly. Though outside the scope of our work in SDSS, it might also be useful for future studies to use shape measurements that probe different effective radii in a controlled manner, to better quantify the effects of isophotal twisting and ellipticity gradients in the measured galaxy alignments.

While further investigation into multiplicative biases and ellipticity gradients is beyond the scope of this work, we have provided a first attempt at reconciling studies about intrinsic alignments in SDSS that use different shape measurement methods. The large differences that we uncovered suggest that this issue is one that future weak lensing surveys cannot afford to ignore when forecasting intrinsic alignment contamination of weak lensing signals.

## Acknowledgments

This work was supported by the National Science Foundation under Grant. No. AST-1313169. RM was also supported by an Alfred P. Sloan Fellowship. We thank Jonathan Blazek, Uroš Seljak, Robert Lupton, Elisa Chisari, and Shadab Alam for useful discussions about this work. We thank the anonymous referee for helpful suggestions, and we thank Ananth Tenneti for sharing the MB-II shape catalog.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

# 7 Conclusions

In chapters 2–4 we presented galaxy-lensing cross correlations, their implications for cosmological models, theory of gravity and also performed several tests for systematics errors. In chapters 5–6, we also presented the measurements of intrinsic alignments, their dependence on various galaxy properties and measurement systematics. In this chapter we present our final conclusions and discuss some possible directions for future work.

## 7.1 Weak lensing

In chapter 2 and appendix 4.B, we presented measurements of cross-correlating spectroscopic galaxy samples from SDSS-BOSS survey with the lensing maps from SDSS and Planck (CMB) surveys. Using these measurements we constrained the matter density in the universe to $\sim 10\%$ level while marginalizing over galaxy bias and relative bias in the CMB and galaxy lensing measurements. In the near future a direct application of these measurements will be to perform a more detailed cosmological analysis (work in progress at the time of writing) and constrain the growth of structure at the redshift of lens samples ($z \sim 0.27, 0.6$). Using the measurements at small scales, we also measured the mean halo mass of the lens galaxies but with CMB lensing these constraints are limited due to resolution of Planck CMB lensing maps. Halo mass constraints from galaxy-galaxy lensing are used in chapter 5 to understand the evolution of intrinsic alignments with the halo mass. These measurements highlight the value of galaxy-lensing cross-correlations to study galaxy physics and our results highlighted the potential power to CMB lensing for such applications with some improvement in resolution from the upcoming surveys. We also constrained the cosmic distance ratio, $\mathcal{R}$, to about 10% level though this measurement is not competitive with other cosmological probes due to low redshift of galaxy samples used. However, a powerful application of $\mathcal{R}$ is to test and constrain relative biases in different lensing maps as it is not very sensitive to cosmology. In addition to many useful applications of galaxy-lensing cross correlations, results in chapter 2 highlight the power of doing joint analysis of different probes which allows many internal consistency tests to be performed and help isolate the impact of systematics.

In chapter 3, we did a detailed study of covariance of galaxy-lensing cross correlations and the impact of different estimators on the covariance matrices. Using mock samples and theoretical arguments, we accentuated the importance of subtracting measurement around random points from the signal to improve the covariance. Our results provided justification (similar to Landy & Szalay (1993)) to use mean zero quantities in general when computing n-point functions of cosmological tracers. We also studied the impact of various sources of covariance (shot noise, cosmic variance, systematics) and showed that the shot noise and cosmic variance dominate with some contribution from systematics. Contributions from other sources, especially the super sample covariance and non-gaussian covariance, are subdominant (though it is important for clustering at small scales, $r_p \lesssim 5h^{-1}$Mpc). We showed that different methods of estimating covariance, jackknife and sub-sampling using data or covariance from several mock survey realizations, give consistent results as long as mean zero quantities are used. These results also validated our use of jackknife method to estimate covariances in other chapters. While our results in this chapter are applicable to lensing measurements from SDSS, the methodology we developed will also be useful for future lensing studies to carefully analyze and characterize the behavior of covariance in the measurements.

In chapter 4, we used the measurements from chapter 2 to test the equality of newtonian and curvature potential by measuring the $E_G$ parameter. Our results are consistent with predictions from $\Lambda$CDM model at $\lesssim 1.5\sigma$ levels. However, we did find some tensions in the results with galaxy-lensing amplitude being lower than the predictions using Planck$\Lambda$CDM model (hints of this were also observed in chapter 2) and galaxy bias from clustering measurements. We performed several checks for systematic errors, most notably testing the impact of photometric redshifts in appendix 4.E and testing for the impact of non-linear physics. Our results suggest that no single systematic can

explain the entire discrepancy, though a combination of systematics can still lead to observed level of discrepancies in lensing and CMB measurements.

With the increasing precision of data from ongoing and upcoming weak lensing surveys, the methodology we have developed for analyzing the weak lensing measurements with a particular focus on tests for systematics will have broad applications. For example, Schaan et al. (2016) showed that joint analysis of LSST with CMB stage IV experiments will enable percent level constraints on cosmology while also providing internal checks to control systematics in shear estimation (they assumed no systematics in CMB lensing for their analysis). Cross correlations of CMB lensing with large scale structures at high redshifts (beyond reach of galaxy lensing surveys), using data from deep surveys such as WFIRST, provides another avenue to study the growth of structure at high redshifts, important for dark matter and dark energy constraints. Extending $E_G$ to higher redshifts with increasing precision also enables strong tests of gravity (Pullen et al., 2015b). We also highlighted several calibration issues in these measurements which will have to be calibrated out with even higher precisions.

Our results also suggest several avenues for future research, with focus on controlling the impact of systematics. We need to develop better models for photometric redshifts, especially using the clustering method. As shown in appendix 4.E, clustering photo-z are sensitive to the priors on galaxy bias, especially redshift evolution of bias of photometric galaxies which can be hard to constrain given complicated selection functions. With the help of simulations we can attempt to construct the realistic survey mocks and get more realistic priors on bias. There is also a pressing requirement to construct better models for non-linear physics given increasing precision of upcoming surveys. In the context of $E_G$ measurements, we showed that the biases due to non-linear physics can be modeled at precision of about $\sim 2\%$, which may not be enough for the upcoming surveys such as LSST and EUCLID, which aim for percent level measurements. We require better models for non-linear matter clustering and a better understanding of galaxy formation and evolution to understand galaxy bias and links between galaxies and matter. We also need better understanding of the covariance in the measurements. While we did a detailed analysis for several terms in the covariance matrix, our analysis was essentially limited to gaussian fields. With the reduced contribution from shot noise terms in future datasets, the contribution from the connected (non-gaussian) terms in the cosmic variance will become important. While these can be estimated using a large suite of mock samples, improving theoretical models is necessary for better understanding and also faster generation of covariance matrices.

## 7.2 Intrinsic Alignments

In chapter 5 we presented a detailed study of Intrinsic alignments (IA) of galaxies for LOWZ sample. We detected IA over a large range of scales, $0.1 < r_p < 200h^{-1}$Mpc, tested the validity of linear alignment model, $5 < r_p < 70h^{-1}$Mpc (for $r_p > 70h^{-1}$Mpc we are limited by size of jackknife regions to estimate the covariance). At small scales, $r_p < 5h^{-1}$Mpc non-linear physics is not captured by the model and we modeled these scales using the halo model from Schneider & Bridle (2010). These measurements are also modeled by Blazek et al. (2015) using the NFW profile (NFW is not full model but works because sample has low satellite fractions and measurement is dominated by central-satellite pairs). We also studied the dependence of IA on various galaxy properties. General conclusions is that brighter galaxies which are preferentially in more massive and higher bias halos, show stronger alignments. We did not find any redshift evolution of IA, though the redshift baseline of our sample is quite small.

In chapter 6, we also tested the IA model in two dimensional space, studying the anisotropy introduced by projected shapes and redshift space distortions. Linear alignment model successfully predicts the anisotropies and we also showed that different treatment of anisotropies can explain the discrepancies in measurements of shape-shape correlations with other studies. We compared different methods of shape measurements and their impact on IA measurements. IA amplitude was highest for isophotal shapes followed by de Vaucouleurs and re-gaussianization shapes. We did find some additive systematics in de Vaucouleurs shapes but tests did not suggest presence of any residual

systematics in isophotal and re-gaussianization shapes, though our tests are not exhaustive. Since different measurements weigh galaxy profiles differently, our measurements suggest twisting of galaxy shapes with outer regions (favored by isophotal shapes) being more strongly aligned with the tidal fields.

Our observations leave several open questions which are important for IA modeling and mitigation for weak lensing analysis and also for understanding galaxy physics. While the environment dependence (or the dependence on halo mass/galaxy brightness) are well established now from simulations and observations, exact physical mechanisms causing these dependences are not well understood yet and require more detailed studies of IA, especially from simulations. Another open question is studying the redshift dependence of IA. Current studies of redshift dependence use narrow redshift baselines or by the quality of data. We are currently extending the redshift baseline of our measurement using the BOSS CMASS sample though our results are limited by the quality of imaging from SDSS. Using deeper imaging from ongoing surveys, e.g. HSC and DES, will improve the quality of these tests. The dependence of IA on galaxy radii (with outer regions being more aligned) also requires further inspection. We stress that our current observations are not guaranteed to be independent of systematics, since isophotal shapes from SDSS are not very well characterized and are known to have systematics problems. Also, since central galaxies (which dominate LOWZ sample) reside well inside the halos, the processes that can lead to shape twisting need to be understood better.

# Bibliography


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

Abazajian K. N., et al., 2016, preprint, (arXiv:1610.02743)

Abbott T., et al., 2016, Phys.Rev.D, 94, 022001

Abramenko B., 1978, Ap&SS, 54, 323

Ade P. A. R., et al., 2014, Physical Review Letters, 113, 021301

Ahn C. P., et al., 2012, ApJS, 203, 21

Aihara H., et al., 2011, ApJS, 193, 29

Alam S., et al., 2015a, ApJS, 219, 12

Alam S., Ho S., Vargas-Magaña M., Schneider D. P., 2015b, MNRAS, 453, 1754

Alam S., et al., 2016a, preprint, (arXiv:1607.03155)

Alam S., Ho S., Silvestri A., 2016b, MNRAS, 456, 3743

Alam S., Miyatake H., More S., Ho S., Mandelbaum R., 2017, MNRAS, 465, 4853

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Anderson L., et al., 2014, MNRAS, 441, 24

Baldauf T., Smith R. E., Seljak U., Mandelbaum R., 2010, Phys.Rev.D, 81, 063531

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

Baxter E. J., et al., 2015, ApJ, 806, 247

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

Bernstein G., Jain B., 2004, ApJ, 600, 17

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

Betoule M., et al., 2014, A&A, 568, A22

Beutler F., et al., 2012, MNRAS, 423, 3430

Beutler F., et al., 2014, MNRAS, 443, 1065

Beutler F., et al., 2016, preprint, (arXiv:1607.03150)

Bhattacharya S., Habib S., Heitmann K., Vikhlinin A., 2013, ApJ, 766, 32

Bianchini F., et al., 2015, preprint, (arXiv:1511.05116)

Blake C., et al., 2016, MNRAS, 456, 2806

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

Blazek J., McQuinn M., Seljak U., 2011, J. Cosmology Astropart. Phys., 5, 10

Blazek J., Mandelbaum R., Seljak U., Nakajima R., 2012, J. Cosmology Astropart. Phys., 5, 41

Blazek J., Vlah Z., Seljak U., 2015, J. Cosmology Astropart. Phys., 8, 15

Bleem L. E., et al., 2012, ApJ, 753, L9

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Bolton A. S., et al., 2012, AJ, 144, 144

Bridle S., King L., 2007, New Journal of Physics, 9, 444

Buddendiek A., et al., 2016, MNRAS, 456, 3886

Caminha G. B., et al., 2016, A&A, 587, A80

Cappi A., 1995, A&A, 301, 6

Carlson J., Reid B., White M., 2013, MNRAS, 429, 1674

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

Chisari N. E., Dvorkin C., 2013, J. Cosmology Astropart. Phys., 12, 29

Chisari N. E., Mandelbaum R., Strauss M. A., Huff E. M., Bahcall N. A., 2014, MNRAS, 445, 726

Chisari N. E., Dunkley J., Miller L., Allison R., 2015a, MNRAS, 453, 682

Chisari N., et al., 2015b, MNRAS, 454, 2736

Chisari N., et al., 2016, MNRAS, 461, 2702

Chisari N. E., et al., 2017, preprint, (arXiv:1702.03913)

Clampitt J., et al., 2016, preprint, (arXiv:1603.05790)

Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Phys.Rep., 513, 1

Cole S., et al., 2005, MNRAS, 362, 505

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

Cooray A., Sheth R., 2002, Phys.Rep., 372, 1

Coupon J., et al., 2015, MNRAS, 449, 1352

Croft R. A. C., Metzler C. A., 2000, ApJ, 545, 561

Csabai I., Dobos L., Trencséni M., Herczegh G., Józsa P., Purger N., Budavári T., Szalay A. S., 2007, Astronomische Nachrichten, 328, 852

Cuesta A. J., et al., 2016, MNRAS, 457, 1770

Das S., et al., 2011, Physical Review Letters, 107, 021301

Das S., Errard J., Spergel D., 2013, preprint, (arXiv:1311.2338)

Das S., et al., 2014, J. Cosmology Astropart. Phys., 4, 014

Dawson K. S., et al., 2013, AJ, 145, 10

Desjacques V., Jeong D., Schmidt F., 2016, preprint, (arXiv:1611.09787)

Diego J. M., et al., 2015, MNRAS, 446, 683

Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108

Dodelson S., 2003, Modern cosmology

Dodelson S., Schneider M. D., 2013, Phys.Rev.D, 88, 063537

Efstathiou G., Lemos P., 2017, preprint, (arXiv:1707.00483)

Efstathiou G., et al., 2002, MNRAS, 330, L29

Eisenstein D. J., et al., 2001, AJ, 122, 2267

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Erben T., et al., 2013, MNRAS, 433, 2545

Faltenbacher A., Li C., Mao S., van den Bosch F. C., Yang X., Jing Y. P., Pasquali A., Mo H. J., 2007, ApJ, 662, L71

Faltenbacher A., Jing Y. P., Li C., Mao S., Mo H. J., Pasquali A., van den Bosch F. C., 2008, ApJ, 675, 146

Fasano G., Bonoli C., 1989, A&AS, 79, 291

Feldmann R., et al., 2006, MNRAS, 372, 565

Fischer P., et al., 2000, AJ, 120, 1198

Freedman W. L., Madore B. F., 2010, ARA&A, 48, 673

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748

Giannantonio T., Percival W. J., 2014, MNRAS, 441, L16

Giannantonio T., et al., 2016, MNRAS, 456, 3213

Gil-Marín H., et al., 2016a, MNRAS, 460, 4188

Gil-Marín H., et al., 2016b, MNRAS, 460, 4188

Gillis B. R., et al., 2013, MNRAS, 431, 1439

Golse G., Kneib J.-P., Soucail G., 2002, A&A, 387, 788

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Grieb J. N., et al., 2016, preprint, (arXiv:1607.03143)

Grieb J. N., et al., 2017, MNRAS, 467, 2085

Gunn J. E., et al., 1998, AJ, 116, 3040

Gunn J. E., et al., 2006, AJ, 131, 2332

Hamilton A. J. S., 1992, ApJ, 385, L5

Hamilton A. J. S., 1997, MNRAS, 289, 285

Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188

Han J., et al., 2015, MNRAS, 446, 1356

Hand N., et al., 2015, Phys.Rev.D, 91, 062001

Hao J., Kubo J. M., Feldmann R., Annis J., Johnston D. E., Lin H., McKay T. A., 2011, ApJ, 740, 39

Harnois-Déraps J., et al., 2016, MNRAS,

Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399

Heymans C., et al., 2006, MNRAS, 371, L60

Heymans C., et al., 2013a, MNRAS, 432, 2433

Heymans C., et al., 2013b, MNRAS, 432, 2433

Hikage C., Mandelbaum R., Takada M., Spergel D. N., 2013, MNRAS, 435, 2345

Hildebrandt H., et al., 2010, A&A, 523, A31

Hildebrandt H., et al., 2016, preprint, (arXiv:1606.05338)

Hildebrandt H., et al., 2017, MNRAS, 465, 1454

Hinshaw G., et al., 2013, ApJS, 208, 19

Hirata C., Seljak U., 2003, MNRAS, 343, 459

Hirata C. M., Seljak U., 2004, Phys.Rev.D, 70, 063526

Hirata C. M., et al., 2004, MNRAS, 353, 529

Hirata C. M., Mandelbaum R., Ishak M., Seljak U., Nichol R., Pimbblet K. A., Ross N. P., Wake D., 2007, MNRAS, 381, 1197

Hirata C. M., Ho S., Padmanabhan N., Seljak U., Bahcall N. A., 2008, Phys.Rev.D, 78, 043520

Hoekstra H., Yee H. K. C., Gladders M. D., 2002, ApJ, 577, 595

Hoekstra H., Yee H. K. C., Gladders M. D., 2004, ApJ, 606, 67

Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS,

449, 685

Hogg D. W., 1999, ArXiv Astrophysics e-prints,

Hogg D. W., Finkbeiner D. P., Schlegel D. J., Gunn J. E., 2001, AJ, 122, 2129

Hu W., Dodelson S., 2002, ARA&A, 40, 171

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702

Hu W., Okamoto T., 2002, ApJ, 574, 566

Hu W., Sugiyama N., 1996, ApJ, 471, 542

Hu W., White M., 2001, ApJ, 554, 67

Hu W., DeDeo S., Vale C., 2007a, New Journal of Physics, 9, 441

Hu W., Holz D. E., Vale C., 2007b, Phys.Rev.D, 76, 127301

Hudson M. J., et al., 2015, MNRAS, 447, 298

Hung C.-L., Ebeling H., 2012, MNRAS, 421, 3229

Ivezić Ž., et al., 2004, Astronomische Nachrichten, 325, 583

Jaffe A. H., et al., 2001, Physical Review Letters, 86, 3475

Jain B., Khoury J., 2010, Annals of Physics, 325, 1479

Jain B., Taylor A., 2003, Physical Review Letters, 91, 141302

Jee M. J., Tyson J. A., Schneider M. D., Wittman D., Schmidt S., Hilbert S., 2013, ApJ, 765, 74

Jennings E., Baugh C. M., Pascoli S., 2011, ApJ, 727, L9

Joachimi B., Bridle S. L., 2010, A&A, 523, A1

Joachimi B., Schneider P., 2008, A&A, 488, 829

Joachimi B., Mandelbaum R., Abdalla F. B., Bridle S. L., 2011, A&A, 527, A26

Joachimi B., Semboloni E., Bett P. E., Hartlap J., Hilbert S., Hoekstra H., Schneider P., Schrabback T., 2013a, MNRAS, 431, 477

Joachimi B., Semboloni E., Hilbert S., Bett P. E., Hartlap J., Hoekstra H., Schneider P., 2013b, MNRAS, 436, 819

Joachimi B., et al., 2015, preprint, (arXiv:1504.05456)

Joudaki S., et al., 2016, preprint, (arXiv:1610.04606)

Joudaki S., et al., 2017, MNRAS, 465, 2033

Kaiser N., 1987, MNRAS, 227, 1

Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460

Kang X., van den Bosch F. C., Yang X., Mao S., Mo H. J., Li C., Jing Y. P., 2007, MNRAS, 378, 1531

Keck Array T., et al., 2016, preprint, (arXiv:1606.01968)

Khandai N., Di Matteo T., Croft R., Wilkins S. M., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2014,

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

Kiessling A., et al., 2015, preprint, (arXiv:1504.05546)

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

Kilbinger M., et al., 2013, MNRAS, 430, 2200

Kilbinger M., et al., 2017, preprint, (arXiv:1702.05301)

King L. J., 2005, A&A, 441, 47

Kirk D., et al., 2015, preprint, (arXiv:1504.05465)

Kirk D., et al., 2016, MNRAS, 459, 21

Kitching T. D., et al., 2015, preprint, (arXiv:1512.03627)

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Komatsu E., et al., 2011, ApJS, 192, 18

Kormendy J., 1982, in Martinet L., Mayor M., eds, Saas-Fee Advanced Course 12: Morphology and Dynamics of Galaxies. pp 113–288

Krause E., Eifler T., Blazek J., 2016, MNRAS, 456, 207

Kuhlen M., Diemand J., Madau P., 2007, ApJ, 671, 1135

Kwan J., et al., 2016, preprint, (arXiv:1604.07871)

LSST Science Collaboration et al., 2009, preprint, (arXiv:0912.0201)

Lahav O., Suto Y., 2004, Living Reviews in Relativity, 7, 8

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Lauer T. R., et al., 2005, AJ, 129, 2138

Laureijs R., et al., 2011, preprint, (arXiv:1110.3193)

Lawrence E., Heitmann K., White M., Higdon D., Wagner C., Habib S., Williams B., 2010, ApJ, 713, 1322

Leauthaud A., et al., 2012, ApJ, 744, 159

Leauthaud A., et al., 2017, MNRAS, 467, 3024

Leonard C. D., Ferreira P. G., Heymans C., 2015, J. Cosmology Astropart. Phys., 12, 051

Lewis A., Bridle S., 2002a, Phys.Rev.D, 66, 103511

Lewis A., Bridle S., 2002b, Phys. Rev., D66, 103511

Lewis A., Challinor A., 2006, Phys.Rep., 429, 1

Li C., Jing Y. P., Faltenbacher A., Wang J., 2013, ApJ, 770, L12

Li Y., Hu W., Takada M., 2014, Phys.Rev.D, 89, 083519

Link R., Pierce M. J., 1998, ApJ, 502, 63

Liu J., Hill J. C., 2015, Phys.Rev.D, 92, 063517

Liu J., Ortiz-Vazquez A., Hill J. C., 2016, preprint, (arXiv:1601.05720)

Lupton R. H., Gunn J. E., Ivezić Z., Knapp G. R., Kent S., Yasuda N., 2001a, in ASP Conf. Ser. 238: Astronomical Data Analysis Software and Systems X. pp 269–278

Lupton R., Gunn J. E., Ivezić Z., Knapp G. R., Kent S., 2001b, in Harnden Jr. F. R., Primini F. A., Payne H. E., eds, Astronomical Society of the Pacific Conference Series Vol. 238, Astronomical Data Analysis Software and Systems X. p. 269 (arXiv:astro-ph/0101420)

Madhavacheril M., et al., 2015, Physical Review Letters, 114, 151302

Mandelbaum R., et al., 2005, MNRAS, 361, 1287

Mandelbaum R., Hirata C. M., Ishak M., Seljak U., Brinkmann J., 2006, MNRAS, 367, 611

Mandelbaum R., Seljak U., Hirata C. M., 2008, J. Cosmology Astropart. Phys., 8, 6

Mandelbaum R., et al., 2011, MNRAS, 410, 844

Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544

Mandelbaum R., et al., 2014, ApJS, 212, 5

Mandelbaum R., et al., 2015, MNRAS, 450, 2963

Mandelbaum R., Wang W., Zu Y., White S., Henriques B., More S., 2016, MNRAS, 457, 3200

Mandelbaum R., et al., 2017, preprint, (arXiv:1705.06745)

Manera M., et al., 2013, MNRAS, 428, 1036

Manera M., et al., 2014,

Manera M., et al., 2015, MNRAS, 447, 437

Massey R., et al., 2007, ApJS, 172, 239

Massey R., Kitching T., Richard J., 2010, Reports on Progress in Physics, 73, 086901

Massey R., et al., 2013, MNRAS, 429, 661

Masters K. L., et al., 2011, MNRAS, 418, 1055

Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint, (arXiv:1303.4722)

Miyatake H., et al., 2015, ApJ, 806, 1

Miyatake H., Madhavacheril M. S., Sehgal N., Slosar A., Spergel D. N., Sherwin B., van Engelen A., 2016, preprint, (arXiv:1605.05337)

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution

Mohammed I., Seljak U., Vlah Z., 2016, preprint, (arXiv:1607.00043)

Moradinezhad Dizgah A., Durrer R., 2016, J. Cosmology Astropart. Phys., 9, 035

More S., Kravtsov A. V., Dalal N., Gottlöber S., 2011, ApJS, 195, 4

More S., Miyatake H., Mandelbaum R., Takada M., Spergel D., Brownstein J., Schneider D. P., 2014,

More S., Miyatake H., Mandelbaum R., Takada M., Spergel D. N., Brownstein J. R., Schneider D. P., 2015, ApJ, 806, 2

Mukhanov V., 2005, Physical Foundations of Cosmology, doi:10.2277/0521563984.

Nakajima R., Mandelbaum R., Seljak U., Cohn J. D., Reyes R., Cool R., 2012, MNRAS, 420, 3240

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Newman J. A., 2008, ApJ, 684, 88

Nieto J.-L., Bender R., Poulain P., Surma P., 1992, A&A, 257, 97

Okumura T., Jing Y. P., Li C., 2009, ApJ, 694, 214

Padmanabhan N., et al., 2008, ApJ, 674, 1217

Padmanabhan N., White M., Cohn J. D., 2009, Phys.Rev.D, 79, 063523

Parejko J. K., et al., 2013, MNRAS, 429, 98

Pasquali A., et al., 2006, ApJ, 636, 115

Pereira M. J., Bryan G. L., Gill S. P. D., 2008, ApJ, 672, 825

Perlmutter S., et al., 1999, ApJ, 517, 565

Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezić Ž., 2003, AJ, 125, 1559

Planck Collaboration et al., 2014a, A&A, 571, A16

Planck Collaboration et al., 2014b, A&A, 571, A17

Planck Collaboration et al., 2015a, preprint, (arXiv:1502.01589)

Planck Collaboration et al., 2015b, preprint, (arXiv:1502.01591)

Planck Collaboration et al., 2015c, preprint, (arXiv:1502.01597)

Pourtsidou A., 2016, MNRAS, 461, 1457

Pullen A. R., Alam S., He S., Ho S., 2015a, preprint, (arXiv:1511.04457)

Pullen A. R., Alam S., Ho S., 2015b, MNRAS, 449, 4326

Pullen A. R., Alam S., He S., Ho S., 2016, MNRAS, 460, 4098

Reid B. A., Spergel D. N., 2009, ApJ, 698, 143

Reid B. A., Seo H.-J., Leauthaud A., Tinker J. L., White M., 2014, MNRAS, 444, 476

Reid B., et al., 2016, MNRAS, 455, 1553

Reyes R., Mandelbaum R., Seljak U., Baldauf T., Gunn J. E., Lombriser L., Smith R. E., 2010, Nature, 464, 256

Reyes R., Mandelbaum R., Gunn J. E., Nakajima R., Seljak U., Hirata C. M., 2012, MNRAS, 425, 2610

Richards G. T., et al., 2002, AJ, 123, 2945

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 2016, ApJ, 826, 56

Rodríguez-Torres S. A., et al., 2016, MNRAS,

Romanowsky A. J., Kochanek C. S., 1998, ApJ, 493, 641

Ross A. J., et al., 2012, MNRAS, 424, 564

Samushia L., et al., 2013, MNRAS, 429, 1514

Samushia L., et al., 2014, MNRAS, 439, 3504

Sánchez A. G., et al., 2014, MNRAS, 440, 2692

Sanchez A. G., et al., 2016, preprint, (arXiv:1607.03147)

Satpathy S., et al., 2016, preprint, (arXiv:1607.03148)

Schaan E., Takada M., Spergel D. N., 2014, Phys.Rev.D, 90, 123523

Schaan E., Krause E., Eifler T., Doré O., Miyatake H., Rhodes J., Spergel D. N., 2016, preprint, (arXiv:1607.01761)

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Schneider M. D., Bridle S., 2010, MNRAS, 402, 2127

Schneider P., van Waerbeke L., Kilbinger M., Mellier Y., 2002, A&A, 396, 1

Schneider M. D., et al., 2013, MNRAS, 433, 2727

Scoccimarro R., Zaldarriaga M., Hui L., 1999, ApJ, 527, 1

Seljak U., et al., 2005, Phys.Rev.D, 71, 043511

Sheldon E. S., et al., 2004, AJ, 127, 2544

Sherwin B. D., et al., 2012, Phys.Rev.D, 86, 083006

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Shirasaki M., Takada M., Miyatake H., Takahashi R., Hamana T., Nishimichi T., Murata R., 2016, preprint, (arXiv:1607.08679)

Sifón C., Hoekstra H., Cacciato M., Viola M., Köhlinger F., van der Burg R., Sand D., Graham M. L., 2014,

Sifón C., et al., 2015a, MNRAS, 454, 3938

Sifón C., Hoekstra H., Cacciato M., Viola M., Köhlinger F., van der Burg R. F. J., Sand D. J., Graham M. L., 2015b, A&A, 575, A48

Simet M., Battaglia N., Mandelbaum R., Seljak U., 2017, MNRAS, 466, 3663

Simpson F., et al., 2013, MNRAS, 429, 2249

Singh S., Mandelbaum R., 2016, MNRAS, 457, 2301

Singh S., Mandelbaum R., More S., 2015, MNRAS, 450, 2195

Singh S., Mandelbaum R., Brownstein J. R., 2016a, preprint, (arXiv:1606.08841)

Singh S., Mandelbaum R., Seljak U., Slosar A., Vazquez Gonzalez J., 2016b, preprint, (arXiv:1611.00752)

Skibba R. A., van den Bosch F. C., Yang X., More S., Mo H., Fontanot F., 2011, MNRAS, 410, 417

Smee S. A., et al., 2013, AJ, 146, 32

Smith J. A., et al., 2002, AJ, 123, 2121

Smith R. E., et al., 2003, MNRAS, 341, 1311

Smith K. M., Zahn O., Doré O., 2007, Phys.Rev.D, 76, 043510

Spergel D. N., et al., 2003, ApJS, 148, 175

Spergel D., et al., 2013, preprint, (arXiv:1305.5422)

Springel V., et al., 2005, Nature, 435, 629

Steigman G., 2010, preprint, (arXiv:1008.4765)

Stoughton C., et al., 2002, AJ, 123, 485

Strauss M. A., et al., 2002, AJ, 124, 1810

Takada M., Hu W., 2013, Phys.Rev.D, 87, 123504

Takada M., Spergel D. N., 2014, MNRAS, 441, 2456

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152

Taylor J. E., et al., 2012, ApJ, 749, 127

Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928

Tenneti A., Singh S., Mandelbaum R., Di Matteo T., Feng Y., Khandai N., 2014a,

Tenneti A., Mandelbaum R., Di Matteo T., Feng Y., Khandai N., 2014b, MNRAS, 441, 470

Tenneti A., Mandelbaum R., Di Matteo T., Kiessling A., Khandai N., 2015, MNRAS, 453, 469

Tenneti A., Mandelbaum R., Di Matteo T., 2016, MNRAS, 462, 2668

Tenneti A., Gnedin N. Y., Feng Y., 2017, ApJ, 834, 169

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

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., George M. R., Leauthaud A., Bundy K., Finoguenov A., Massey R., Rhodes J., Wechsler R. H., 2012, ApJ, 755, L5

Tinker J. L., Leauthaud A., Bundy K., George M. R., Behroozi P., Massey R., Rhodes J., Wechsler R. H., 2013, ApJ, 778, 93

Tojeiro R., et al., 2014, MNRAS, 440, 2222

Troxel M. A., Ishak M., 2015, Phys.Rep., 558, 1

Tucker D. L., et al., 2006, Astronomische Nachrichten, 327, 821

Vallinotto A., 2012, ApJ, 759, 32

Velander M., et al., 2014, MNRAS, 437, 2111

Velliscig M., et al., 2015a, preprint, (arXiv:1507.06996)

Velliscig M., et al., 2015b, MNRAS, 453, 721

Viola M., et al., 2015, MNRAS, 452, 3529

Wake D. A., et al., 2006, MNRAS, 372, 537

Wang L., Reid B., White M., 2014, MNRAS, 437, 588

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys.Rep., 530, 87

White M., 2014, MNRAS, 439, 3630

Will C. M., 2014, Living Reviews in Relativity, 17

Wyatt Jr. S. P., 1953, AJ, 58, 50

York D. G., et al., 2000, AJ, 120, 1579

Zaldarriaga M., Seljak U., 1999, Phys.Rev.D, 59, 123507

Zhang P., Liguori M., Bean R., Dodelson S., 2007, Physical Review Letters, 99, 141302

Zhang Y., Yang X., Wang H., Wang L., Mo H. J., van den Bosch F. C., 2013, ApJ, 779, 160

Zheng Z., et al., 2005, ApJ, 633, 791

Zu Y., Mandelbaum R., 2015, MNRAS, 454, 1161

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

de Putter R., Doré O., Das S., 2014, ApJ, 780, 185

de la Torre S., et al., 2013, A&A, 557, A54

di Tullio G., 1978, A&A, 62, L17

di Tullio G. A., 1979, A&AS, 37, 591

van Daalen M. P., White M., 2017, preprint, (arXiv:1703.05326)

van Engelen A., et al., 2012, ApJ, 756, 142

van Engelen A., Bhattacharya S., Sehgal N., Holder G. P., Zahn O., Nagai D., 2014, ApJ, 786, 13

van Uitert E., Hoekstra H., Schrabback T., Gilbank D. G., Gladders M. D., Yee H. K. C., 2012, A&A, 545, A71

van Uitert E., et al., 2016, MNRAS, 459, 3251

van Uitert E., et al., 2017, preprint, (arXiv:1706.05004)

van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, MNRAS, 430, 725