

Measuring the Weak Gravitational Lensing Signal from Cosmic Voids

by

Hunter Martin

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2026

© Hunter Martin 2026

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Tereasa Brainerd
Professor, Dept. of Astronomy, Boston University

Supervisor(s): Michael Hudson
Professor, Dept. of Physics and Astronomy, University of Waterloo

Internal Member: Will Percival
Professor, Dept. of Physics and Astronomy, University of Waterloo

Internal-External Member: Marek Stastna
Professor, Dept. of Applied Mathematics, University of Waterloo

Other Member(s): James Taylor
Professor, Dept. of Physics and Astronomy, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

I am the sole author of Chapters 1 through 3, Section 4.2, and Chapter 7 under the supervision of Mike Hudson.

Chapter 5, Sections 4.1 and 4.3, and Appendices A.1 and B are based on the following paper:

Martin, H., Hudson, M. J., Woodfinden, A. et al, 2025. Lensing the darkness: The matter density profile of cosmic voids from UNIONS arxiv:2507.13450 (Submitted to MNRAS, no further review needed)

I was the primary author for the above work under the supervision of Mike Hudson. All work in that paper is my own, with the exception of the void-galaxy cross-correlation measurements shown in colour in Fig. 5.3, which were performed by Alex Woodfinden.

Chapter 6, Section 4.4, and Appendices A.2 and C are based on work that is currently in preparation for publication. As these sections have not been submitted for internal review by the *Euclid* collaboration, all work shown here is preliminary. I am the primary author, and all work presented in these sections is my own under the supervision of Mike Hudson.

Abstract

The field of cosmology is currently in an era principally focused on statistical precision. To achieve greater precision, deeper and wider surveys are actively being developed and conducted to gather more and more information about the surrounding cosmos. Alternatively, there are efforts to develop new statistical probes to better utilize currently-existing data to obtain tighter constraints. Cosmic voids, as vast underdense regions, then represent an ideal candidate to complement the statistical information already extracted from the opposite density extremes: massive luminous galaxies and galaxy clusters. Voids have already seen some success in constraining cosmology through probes like the void size function and void-galaxy cross-correlations. This thesis introduces the matter distribution within cosmic voids as measured by weak gravitational lensing as a new probe that is significantly detectable within current and future data sets. The goal of this demonstration is to justify future efforts in extracting the cosmological information from this newfound signal.

For data currently available, we make use of the large overlap of the Sloan Digital Sky Survey (SDSS) Baryon Oscillation Spectroscopic Survey (BOSS) and the Ultraviolet Near-Infrared Optical Northern Survey (UNIONS). We measure void lensing around BOSS voids and find that we can detect the signal at 6.2σ significance, the most significant detection from spectroscopically-identified voids to date. We additionally are able to significantly detect differences in the void profile with void size between the larger half and smaller half of the void catalogue at 2.3σ . To help perform this measurement, we present and validate a novel method for computing the Gaussian component of the conventional weak lensing covariance, adapted for use with void studies. Comparing the void profile to a measurement of the void-galaxy cross-correlation to test the linearity of the relationship between mass and light, we find good visual agreement between the two, and a galaxy bias factor of 2.45 ± 0.36 , consistent with other works. We additionally assist in future developments of the UNIONS data by running quality control tests for the future photometric redshift data sets. These future releases will provide additional data to make these detections stronger.

For future data, we use the Flagship simulation of the *Euclid* survey to simulate the expected data from the *Euclid* VISible imager (VIS) and Near-Infrared Spectrometer and Photometer (NISF) instruments across an octant of the sky. This octant then roughly corresponds to an equivalent area of the planned second data release of the survey. We extend the methodology and covariance model to a lensing tomography setup by binning voids and sources along the line of sight. We then stack information along source bins. From this, we are able to detect the void lensing profiles with 12σ , 11σ , 7.1σ , and 4.7σ significance across the four different void redshift bins. Scaling the most significant result to the expected areas for the first and final data releases,

we get 6.9σ and 21σ respectively. We additionally find a 4.4σ difference between the lensing profiles of the smallest voids and the largest voids.

Acknowledgements

To Dr. Mike Hudson: Thank you for providing me the opportunity to explore the world of cosmic voids. Your expertise and guidance have proved invaluable time and time again throughout the course of this program. Thank you for taking the time and having patience with me through the many twists and turns these projects have had.

To the many friends I met from amongst the grad students and postdocs during my time at the WCA: Thank you for being there when I needed to rant about my code or about my covariances. I hope I was able to be there for as many of your rants as you were for me.

To Thomas, Roland, and Chris: Thank you for the many nights of gaming and hanging out that helped keep me sane.

To Ahmed: Thank you for being an amazing roommate and for always being there to bounce ideas off of.

Dedication

This thesis is dedicated to my family, specifically my mom, dad, stepdad, and sister.
I would not be in this position if it weren't for all of your unwavering support.

Table of Contents

Examining Committee Membership	ii
Author's Declaration	iii
Statement of Contributions	iv
Abstract	v
Acknowledgements	vii
Dedication	viii
List of Figures	xiii
List of Tables	xv
List of Abbreviations	xvi
1 Introduction	1
1.1 Cosmology	2
1.2 Λ CDM model	5
1.3 Distances	6
1.4 Density Perturbations and Growth of Structure	7

1.5	Statistical Distribution of Matter	9
1.6	Extensions to Λ CDM: Neutrinos and Modified Gravity	10
1.6.1	Modified Gravity	10
1.6.2	Neutrinos	11
1.7	The Goal of Modern Cosmology	12
2	Weak Lensing Formalism	14
2.1	Weak Lensing Theory	14
2.2	Galaxy Shape Measurements	17
2.3	Redshift Distributions	19
3	Cosmic Voids	21
3.1	Idealized Voids	21
3.2	Voids in Observations	23
4	Data	28
4.1	UNIONS Galaxy Shape Catalogue	28
4.2	UNIONS Multiband Photometry and Quality Control Tests	30
4.3	SDSS BOSS	41
4.3.1	Void Finding	41
4.4	Euclid Flagship Simulation	47
4.4.1	Euclid Flagship NISP Void Catalogue	48
4.4.2	Euclid Flagship VIS Sources	49
5	Measuring Lensing from BOSS Voids with UNIONS Sources	54
5.1	Introduction	54
5.2	Methodology	55
5.2.1	Weak Lensing Formalism	55
5.2.2	Random Lensing Voids from PATCHY Mocks	57

5.2.3	Analytic Covariance Modelling	58
5.2.4	Void Density Profile Models	64
5.3	Results	70
5.3.1	Void profiles	70
5.3.2	Integrated Density Contrast	78
5.3.3	Galaxy Bias in Underdense Environments	82
5.4	Discussion	83
5.4.1	Comparison with Previous Results	83
5.4.2	Future Work	84
5.5	Conclusions	85
6	Euclid Forecasts	87
6.1	Introduction	87
6.2	Methodology	88
6.2.1	Weak Lensing Methodology	88
6.2.2	Analytical Covariance Model	89
6.2.3	Stacking Source Bins	91
6.2.4	Void Density Profile Model	92
6.2.5	Model Fitting Methodology	93
6.3	Results	94
6.3.1	Integrated Density Contrast	99
6.4	Discussion	103
6.5	Conclusions	105
7	Conclusions and Future Work	106
7.1	Future Work	107
	References	112

APPENDICES	122
A Comparing Covariance Terms	123
A.1 BOSS and UNIONS Covariance	123
A.2 Euclid Flagship Covariance	127
B Investigating the Five-Parameter HSW Model Posterior	131
C Jackknife Verification	136

List of Figures

2.1	Diagram of the gravitational lensing ray tracing	16
3.1	Diagram of the effects of weak lensing on circular source	24
3.2	Examples of voids in an N -body simulation	26
4.1	Overlap of the CMASS and UNIONS survey footprints	29
4.2	UNIONS median limiting u -band magnitude sky plot	33
4.3	UNIONS median limiting g -band magnitude sky plot	34
4.4	UNIONS median limiting i -band magnitude sky plot	35
4.5	UNIONS median limiting r -band magnitude sky plot	36
4.6	UNIONS median limiting WISHES z -band magnitude sky plot	37
4.7	UNIONS median limiting Pan-STARRS z -band magnitude sky plot	38
4.8	UNIONS median redshift sky plot	39
4.9	UNIONS star-galaxy separation	40
4.10	BOSS void radius distribution	45
4.11	BOSS and UNIONS redshift distributions	46
4.12	<i>Euclid</i> Flagship void size distribution	50
4.13	<i>Euclid</i> Flagship tomographic redshift distributions	52
5.1	PATCHY mock void lensing results	59
5.2	Analytic covariance diagonals compared to mock covariance diagonals	63
5.3	Void-galaxy cross-correlation profile measurements and average	67

5.4	BOSS void lensing results	72
5.5	BOSS void lensing results by void size	73
5.6	Corner plot for the BOSS Full catalogue	76
5.7	Corner plots for the BOSS Small and Large catalogues	79
5.8	Galaxy bias as a function of void size	80
6.1	<i>Euclid</i> Flagship tomographic void lensing results	95
6.2	<i>Euclid</i> Flagship void lensing results from the void size catalogues	96
6.3	Corner plot for the <i>Euclid</i> Flagship tomographic catalogues	100
6.4	Corner plot for the <i>Euclid</i> Flagship void size catalogues	101
7.1	Simulated 3D void $\Delta\Sigma$ profiles in modified gravity	109
7.2	Simulated 2D void $\Delta\Sigma$ profiles in modified gravity	109
A.1	Void bias measurement from BOSS void-void clustering	126
A.2	Covariance term comparison for the BOSS Full catalogue	128
A.3	Covariance term comparison for the <i>Euclid</i> Flagship L1 catalogue	130
B.1	Comparing the two posterior region fits to the BOSS Full catalogue	132
C.1	Comparison between the jackknife diagonal and the model covariance diagonal	138

List of Tables

4.1	BOSS void catalogues and statistics	44
4.2	<i>Euclid</i> Flagship void catalogues and statistics	53
5.1	Five-parameter HSW fit statistics for the BOSS catalogues	74
5.2	Void-galaxy fit statistics to the BOSS catalogues	75
5.3	Integrated density contrasts of the BOSS catalogues	81
6.1	<i>Euclid</i> Flagship HSW fit statistics	97
6.2	Integrated density contrasts for the <i>Euclid</i> Flagship catalogues	102

List of Abbreviations

AP Alcock-Paczyński

B+15 [Barreira et al. \(2015\)](#)

BAO Baryonic Acoustic Oscillations

BOSS Baryon Oscillation Spectroscopic Survey

CDM Cold Dark Matter

CfA Centre for Astrophysics

CFHT Canada-France-Hawaii Telescope

CFHTLenS Canada-France-Hawaii Telescope Lensing Survey

CFIS Canada-France Imaging Survey

CMB Cosmic Microwave Background

COBE COsmic Background Explorer

DES Dark Energy Survey

DESI Dark Energy Spectroscopic Instrument

DR Data Release

ELG Emission Line Galaxy

ESMD Excess Surface Mass Density

FLRW Friedmann-Lemaître-Robertson-Walker

GR General Relativity

HOD Halo Occupation Distribution

HSC Hyper-Suprime Camera

HSW [Hamaus et al. \(2014b\)](#)

KiDS Kilo-Degree Survey

LoS Line of Sight

LRG Luminous Red Galaxy

LSS Large-Scale Structure

LSST Large Synoptic Survey Telescope

MCMC Monte Carlo Markov Chain

NGC North Galactic Cap

NISP Near Infrared Spectrometer and Photometer

Pan-STARRS Panoramic Survey Telescope and Rapid Response System

PSF Point Spread Function

Revolver REal-space VOid Locations from surVEy Reconstruction

RMS Root Mean Squared

RSD Redshift Space Distortion

SDSS Sloan Digital Sky Survey

SNR Signal-to-Noise Ratio

SOM Self-Organizing Map

SPHEREx Spectro-Photometer for the History of the Universe, Epoch of Reionization and Ices Explorer

UNIONS Ultraviolet Near-Infrared Optical Northern Survey

VIDE Void IDentification and Examination

VIS VISible imager

WISHES Wide Imaging with Subaru HSC of the Euclid Sky

ZOBOV ZOnes Bordering On Voidness

Chapter 1

Introduction

The publication of Einstein’s General Theory of Relativity (GR; [Einstein, 1916](#)) laid the foundations of our modern understanding of cosmology. While at the time he was a proponent of a static universe through the inclusion of the cosmological constant Λ , this theoretical framework laid the groundwork for the development of the theory of the Big Bang and an expanding universe through the independent derivations of [Friedmann \(1922\)](#), and [Lemaître \(1931\)](#). An expanding universe was later evidenced by the findings of [Hubble \(1929\)](#), who observed the famous Hubble Law, wherein galaxies were found to be moving away from the observer at rates proportional to their distance. The discovery of the Cosmic Microwave Background (CMB) radiation ([Penzias & Wilson, 1965](#)) as an early-time relic of the Big Bang helped to confirm the model further and provides a wealth of cosmological information even today. Future data sets, such as COBE measuring the anisotropies of the CMB ([Smoot et al., 1992](#)), or the first large systematic galaxy redshift survey, CfA ([Huchra et al., 1983](#)), ushered in a new wave of cosmological probes by studying the Universe through statistical distributions.

Today, we are in an era of precision cosmology. Developments and advancements in telescopes have produced the ability to measure dimmer and dimmer luminous objects, over wider and deeper patches of the sky, both from the ground and from space. These surveys capture light from larger and larger numbers of galaxies, numbering in the hundreds of millions to billions expected in future surveys. These surveys observe bright objects in such large quantities that statistical detections of percent-level effects (such as the distortion of galaxy shapes via weak gravitational lensing) are possible with great precision. The results of these measurements have revealed many lingering questions about our cosmological models, such as the nature of the “dark” sectors of our Universe: dark matter and dark energy. To probe deeper into our cosmology to find these answers, we can either develop new surveys, like *Euclid* ([Euclid Collaboration et al., 2025](#)), LSST ([Ivezić et al., 2019](#)), DESI ([DESI Collaboration et al., 2016](#)), SPHEREx ([Doré et al., 2014](#)), and Roman

(Eifler et al., 2021), or develop new methodology to extract more information from existing surveys, such as through improvements of existing tools or through entirely new measurements. This thesis helps to address the latter as well as forecasts what this new methodology can measure in the context of these new surveys. While the majority of analyses of galaxy surveys focus on the luminous objects we can see, a burgeoning field has been developing to exploit the empty regions, voids, for their cosmological information. The goal of this thesis is to measure the significance that we can detect the weak gravitational lensing signal from voids, which is a measure of the correlation of background galaxy shapes caused by the gravitational potential of the foreground object. We also wish to forecast the strength of the detection for future surveys. This effort will lay the groundwork to motivate the development of tools to extract cosmological information from this new signal by demonstrating that future data sets will be able to constrain this signal with great precision.

This thesis is organized as follows. This chapter will introduce cosmological theory and standard models and describe some of the available tools used to describe the distribution of matter. Chapter 2 introduces the theory of weak gravitational lensing, including techniques used to measure galaxy shapes from these surveys. Chapter 3 introduces cosmic voids, how they formed, and how they are observed. The chapter will also motivate the existence of the void lensing signal and previous efforts made to detect this signal. Chapter 4 describes the data sets used in Chapters 5 and 6, including detailing efforts made to assist in the development of future releases of the Chapter 5 data. Chapter 5 then develops the methodology and provides a measurement of the void lensing signal from currently-available data. Chapter 6 extends this analysis and forecasts what the expected significance of the void lensing detection could be from simulations of the upcoming *Euclid* survey. Finally, Chapter 7 summarizes the main conclusions and details future steps to extract information this novel cosmological probe.

1.1 Cosmology

The study of cosmology is guided by a fundamental principle: that there is no special place in the Universe (homogeneity) and no special direction (isotropy). While plainly not true on small scales (the Milky Way would represent a special location), we find that this principle holds on sufficiently large scales. This is reinforced by the temperature of the CMB, which finds a consistent temperature today of 2.73 K everywhere, with deviations at the order of 10^{-5} K (Fixsen, 2009). Since this temperature is related to the distribution of mass, we find that the Universe is remarkably homogeneous and isotropic. This principle of homogeneity and isotropy, combined with GR to describe how masses interact on cosmological scales, is then sufficient to derive not only the behaviour of the matter distribution internal to the Universe, but also the dynamics of

the Universe itself.

To describe distances in an expanding, homogeneous, and isotropic spacetime, we use the Friedmann-Lemaître-Robertson-Walker (FLRW) metric

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - Kr^2} + S_K^2(r)(d\theta^2 + \sin^2\theta d\phi^2) \right], \quad (1.1)$$

where K describes the curvature of the geometry, and

$$S_K(r) = \begin{cases} \frac{1}{\sqrt{K}} \sin(\sqrt{K}r) & \text{if } K > 0 \\ r & \text{if } K = 0 \\ \frac{1}{\sqrt{|K|}} \sinh(\sqrt{|K|}r) & \text{if } K < 0 \end{cases} \quad (1.2)$$

and $a(t)$, called the scale factor, quantifies how a distance is expanded or contracted through time, normalized to a value of one today. The evolution of this scale factor is then described by the Friedmann equation

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho(a) - \frac{K}{a^2}, \quad (1.3)$$

where $\rho(t)$ is the total energy density contained within the Universe at a given time t . Defining the Hubble parameter,

$$H = \frac{\dot{a}}{a}, \quad (1.4)$$

the curvature dependence of the Friedmann equation can be used to find the critical value of the total energy density which results in a cosmology with zero curvature. Setting K to zero and solving for ρ results in

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}. \quad (1.5)$$

We can then define the parameter $\Omega = \rho/\rho_{\text{crit}}$ to simplify the curvature definition to whether or not the sum of Ω_x for all types of energy density x is equal to one. Solving for K in Eqn. 1.3 results in

$$K = H^2 \left(\sum_x \Omega_x - 1 \right) \quad (1.6)$$

The total energy density is comprised of separate components which scale with the expansion differently. The three broad classes of this energy density are matter, radiation, and vacuum energy. To derive how these different densities scale with the expansion of the Universe, we can

write down a conservation law on the basis that heat from the Universe's expansion cannot leave the system (i.e. the expansion is adiabatic). This results in

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P) , \quad (1.7)$$

where P is pressure. Assuming a perfect fluid, where

$$P = w\rho , \quad (1.8)$$

for some equation of state parameter w , the solution to Eq. 1.7 is then

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

Matter in the Universe exerts a negligible pressure on cosmological scales, therefore $w = 0$ and $\rho \propto a^{-3}$. Radiation exerts radiation pressure such that $w = 1/3$, therefore $\rho \propto a^{-4}$. Vacuum energy, which is an intrinsic property of space, has a density that is expected to be constant with expansion, therefore $\dot{\rho} = 0$, which corresponds to $w = -1$. This then gives us the scale factor dependence of each component of energy density.

We can now rewrite the Friedmann equation to more explicitly separate the expansion dependence of each energy density term. Expanding Ω into each component, and expressing them in terms of their values today (denoted by a subscript 0), we find

$$\frac{H^2}{H_0^2} = \Omega_{m,0}a^{-3} + \Omega_{r,0}a^{-4} + \Omega_{\Lambda,0} + \Omega_{K,0}a^{-2} , \quad (1.10)$$

where we additionally define an effective curvature term, which can be redefined in terms of the deviation of the total energy density Ω from one

$$\Omega_{K,0} = -\frac{K_0}{H_0^2} = \Omega - 1 . \quad (1.11)$$

This form of the Friedmann equation therefore relates the expansion history of the Universe to the relative values of the component energy densities at a given time.

As the Universe expands, per the FLRW metric in Eq. 1.1, it stretches the space coordinates with it, meaning that a cube at a later time will take up more volume than a cube at an earlier time. This is the intuitive reason why matter density scales with volume (a^{-3}), as matter content does not change with expansion. The wavelength of light is similarly affected, stretching with expansion. This effect adds the additional scaling factor for radiation energy density (a^{-4}) as the energy of a photon is directly proportional to its wavelength. This also has the observational

effect of redshifting the light. We define redshift z in terms of the ratio of the wavelength of the light at observation (i.e. current time, or $a_{\text{observed}} = 1$) compared to the wavelength of light at some initial time.

$$1 + z = \frac{\lambda_{\text{observed}}}{\lambda_{\text{emit}}} = \frac{a_{\text{observed}}}{a_{\text{emit}}} = \frac{1}{a}. \quad (1.12)$$

Therefore the scale factor has a direct observational effect in the redshifting of light. It is through the measurements of this redshift effect from [Hubble \(1929\)](#) that led cosmologists to begin to adopt the expanding Universe model.

1.2 Λ CDM model

The current model for cosmology is one that is defined by three principal epochs of expansion, where each of the three non-curvature Ω terms dominate. Further, there is expected to be a period of exponential expansion at the very beginning of the Universe’s timeline. This period, called inflation ([Guth, 1981](#)), necessarily produces a flat Universe, fixing the fine-tuning problem where the smallest deviations from $\Omega = 1$ would have produced vastly different Universes by the current time. The two components of the Λ CDM name correspond to the the “dark” sectors of the Universe: dark matter and dark energy. From examinations of the rotation curves of galaxies ([Rubin & Ford, 1970](#)), motions of nearby galaxies ([Ostriker & Peebles, 1973](#)), differences in X-ray and optical cluster mass estimates ([Bahcall et al., 1995](#)), and mass-to-light ratios of galaxies ([Faber & Gallagher, 1979](#)), it has been demonstrated that there is a “missing mass” that needs to exist in order for our understanding of gravity and cosmology to hold. This missing matter is called dark matter, and it has been found to be the dominant portion of the matter distribution, with visible, luminous matter in the form of baryons making up about 5 per cent of the Universe’s energy budget compared to the 31 per cent for all types of matter ([Planck Collaboration, 2020](#)). The age of the oldest stars and galaxies being older than larger structures like clusters placed limits on the viability of hot dark matter models, which suppress structure growth on small scales. In these models, superstructures would need to fragment into smaller components in a “top down” method of structure formation, but current evidence supports a “bottom up” hierarchy, where smaller objects like stars form first then amass into larger structures. Further, it was found through the weak gravitational lensing signal of merging clusters ([Clowe et al., 2006](#)) that dark matter does not self-interact, consistent with cold dark matter (CDM) models. Additionally, studies of distances to nearby supernovae ([Riess et al., 1998](#); [Perlmutter et al., 1999](#)) have revealed the Universe’s expansion to be accelerating, revealing another dark sector of the Universe to be powering the accelerated expansion: dark energy. Further, this dark energy is expected to be synonymous with the vacuum energy from the cosmological constant Λ , and

its energy density is therefore constant with expansion. These measurements of an accelerated expansion prompted the inclusion of the cosmological constant into the model. Subsequent CMB anisotropy measurements in the following years (de Bernardis et al., 2000) supported a flat Universe model, where the total energy is equivalent to the critical energy density. Since the total matter component is only 31 per cent of the critical energy density, this necessitates that dark energy is the dominant source of energy density today, at about 69 per cent (Planck Collaboration, 2020). However, recent cosmological results hint that the dark energy equation of state might vary with time (Adame et al., 2025).

1.3 Distances

Measuring accurate distances to astronomical objects is critically important in cosmology. As telescope images are effectively two-dimensional projections, we have to indirectly deduce the line-of-sight (LoS) distances to objects. In the very nearby Universe, where Hubble expansion is subdominant, this can be done by studying objects with standardized intrinsic properties, such as standard candles. However, at very large distances, our understanding of how far away an astronomical object is becomes dependent on our knowledge of the underlying cosmology.

To calculate the distance from an observer to an object in an expanding Universe at a redshift z , consider that photons travel along a null geodesic, where $ds = 0$, radially towards the observer, therefore $d\phi = d\theta = 0$. Applying these conditions to the FLRW metric in a flat Universe results in

$$dr = \frac{cdt}{a(t)}. \quad (1.13)$$

Plugging in the definition of the Hubble constant along with $dz = -da/a^2$, which follows from the definition of redshift in Eq. 1.12, we can convert to distance as a function of redshift

$$dr = \frac{cda}{a^2 H(z)} = \frac{cdz}{H(z)}. \quad (1.14)$$

Therefore, we define the *comoving* distance, χ , as

$$r = \chi = \int_0^z \frac{cdz}{H(z)}. \quad (1.15)$$

This is referred to as a comoving distance because the “ruler” grows with the expansion of the Universe, and therefore the measured distance does not depend on the expansion history. This can be converted to a proper distance by accounting for this expansion along one dimension, which results in simply multiplying the comoving distance by a factor of $a(t)$.

The angular diameter distance relates how far away an object is with its angular size. It is given by

$$D_A = \frac{\chi}{1+z}. \quad (1.16)$$

This notion of distance will primarily be used in this thesis to determine the distances between voids and galaxy shapes to determine the radial bins they belong in.

1.4 Density Perturbations and Growth of Structure

We now turn our attention to the physics and behaviour of the matter density and how they relate to the structure of the cosmic web that we observe today. A Universe that follows the cosmological principle of isotropy and homogeneity at all scales necessarily cannot form structure within the Universe, as there is no special point for gravity to act upon to grow an initial density perturbation. Instead, we necessarily must consider a Universe that generated a set of initial matter density perturbations. In inflationary cosmology models, it is predicted that the exponential growth during the inflationary phase of the Universe's expansion "froze-in" a set of density perturbations caused by quantum fluctuations. These density perturbations then act as a base for gravity to amass the surrounding matter onto. We define the perturbations in the matter density field using an overdensity parameter δ

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}. \quad (1.17)$$

We further define its Fourier space counterpart as it is frequently easier to discuss growth of structure in Fourier space

$$\delta(\mathbf{k}) = \frac{1}{(2\pi)^3} \int d\mathbf{x} \delta(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}}. \quad (1.18)$$

At early times, the density fluctuations are very small, therefore this growth of structure can be modelled by linear perturbation theory. In the linear regime, the growth of matter fluctuations is given by

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\rho_b\delta = 0, \quad (1.19)$$

where the second term represents damping of growth by the Hubble expansion, and ρ_b is the background unperturbed matter density. The third term is the growth of the density perturbation due to the gravitational potential described by the Poisson equation. The solution to this equation has two components, a growing mode (δ_+) and a decaying mode (δ_-). At cosmological time scales, it is expected that the decaying mode is negligible, leaving only the growing mode. The

growing mode is then (Mo et al., 2010)

$$\delta_+ \propto H(z) \int_z^\infty (1+z') \left(\frac{H_0}{H(z')} \right)^3 dz'. \quad (1.20)$$

This growing mode, referred to as the linear growth factor, then dictates how overdensities initially grow, and how underdensities initially empty. We then define the linear growth rate f as

$$f(\Omega_m) = \frac{d \ln \delta_+}{d \ln a}. \quad (1.21)$$

This quantity is relevant for correcting for Redshift Space Distortions (RSDs), where additional redshifting occurs due to components of velocity (called peculiar velocity) that arise from local gravitational potential and not the Hubble expansion. Therefore, observations of the clustering of galaxies necessarily take place in what is referred to as ‘redshift space’, which is anisotropic, and typically require removing the RSD effect to reproduce the expected isotropic ‘real space.’ To a good approximation for Λ CDM,

$$f = \Omega_m^{0.55} \quad (1.22)$$

The linear growth rate relates the peculiar velocities to the matter density

$$\nabla \cdot \mathbf{v} = -H(z)f(\Omega_m)\delta. \quad (1.23)$$

At later times, as the density seeds continue to grow, linear perturbation theory begins to break down. In this regime, a useful approximation extends the linear perturbation theory into quasi-linear regimes. The Zel’dovich approximation (Zel’dovich, 1970) extends the linear modelling by not considering the evolution of the densities themselves, but by modelling the trajectories of individual particles. The assumption is that particles will travel in straight lines under the influence of a displacement field that is modelled from the initial conditions. As the model moves forward in time, the magnitude of the displacement field is scaled by the growth of the density perturbations assuming they grow as in linear theory. The direction of the displacement field does not change because the location of density perturbations is assumed to not change. The approximation is found to hold as long as the particle trajectories do not cross. These crossings usually occur when a density perturbation undergoes shell-crossing. This occurs when two matter shells surrounding a density perturbation collide with one another. In an overdensity, this occurs near the centre as matter collapses in on itself, while in an underdensity this shell-crossing occurs at the boundary of the void, where matter shells closer to the centre move outwards faster than shells near the boundary. In regions where the Zel’dovich model breaks down, non-linear models are then typically needed to model growth of structure, such as spherical collapse models in the case of overdensities. However it has been found that the Zel’dovich approximation is still a useful model for describing the matter distribution inside idealized spherical underdensities (Bromley & Geller, 2025).

1.5 Statistical Distribution of Matter

Because the initial density distribution was formed from quantum fluctuations, the initial density field can be modelled as a Gaussian random field. Because Gaussian distributions are specified solely by their mean and variance, and because of homogeneity the mean of the density field is expected to be zero, the matter distribution is defined by the two point function called the matter power spectrum

$$P(k) = \langle \delta(k)\delta(k') \rangle. \quad (1.24)$$

Further, in a Gaussian random field, the density values at each scale k are independent and thus uncorrelated with each other, therefore

$$P(k) = \langle |\delta(k)|^2 \rangle. \quad (1.25)$$

The inverse Fourier transform of the power spectrum is the two-point correlation function

$$\xi(x) = \langle |\delta(x)|^2 \rangle = \frac{1}{2\pi^2} \int_0^\infty \Delta^2(k) \frac{\sin kx}{kx} d\ln k, \quad (1.26)$$

where $x = |\mathbf{x} - \mathbf{x}'|$ and

$$\Delta^2(k) = \frac{k^3}{2\pi^2} P(k) \quad (1.27)$$

is the dimensionless power spectrum. Note that the correlation function only depends on the separation between two points due to isotropy. The two-point correlation function measures how strongly clustered a field or a set of particles are. It is defined as the excess probability (excess when compared to a random sample) of finding an object at a distance x away. The correlation function can be extended to an N -point correlation function, but if the matter distribution of the Universe is fully Gaussian, then the expectation is that the three-point and higher correlation functions are zero. If these higher-order terms are measured and non-zero, then it hints towards potential non-Gaussianities in the matter distribution, either from primordial non-Gaussianities at the largest scales, or smaller-scale non-Gaussianities caused by late-time non-linear physics.

We can use the power spectrum to derive a useful cosmological parameter, σ_8 . This parameter quantifies the clumpiness of the matter distribution, by measuring the variance of the matter field smoothed in $8 h^{-1}$ Mpc spheres. The value of the sphere radius is arbitrary and is set by historic precedent as the scale that approximately delineates between linear and non-linear regimes. The higher the value of σ_8 , the more strongly mass is consolidated into its high-density regions. The lower the value, the more homogeneous the Universe becomes. We define this parameter in terms of the power spectrum by

$$\sigma_8^2 = \int_0^\infty \Delta^2(k) |\tilde{W}_{\text{th}}(k)|^2 d\ln k. \quad (1.28)$$

where $\tilde{W}_{\text{th}}(k)$ is the Fourier transform of the top-hat window function with $8 h^{-1}$ Mpc radius, which is the operator that smooths the dimensionless power spectrum Δ .

Because the majority of the matter distribution is dark matter, and therefore not directly visible, we are limited to measuring the clustering of matter through the clustering of galaxies. Galaxies are biased tracers of the underlying matter distribution, meaning that they have an excess probability of being found near density peaks. On large scales, where the fluctuations are small, the bias can be written as a simple multiplicative factor, b_g (Kaiser, 1984),

$$\delta_g = b_g \delta_m. \quad (1.29)$$

Therefore, to first order, the galaxy-galaxy correlation can be written in terms of the matter-matter correlation

$$\xi_g = b_g^2 \xi_m. \quad (1.30)$$

This formalism applies to any biased tracer of matter, including voids. Further, the bias factor is dependent on properties of the tracer sample. Certain subsets of galaxies, such as Luminous Red Galaxies (LRGs) are more biased than others, like Emission Line Galaxies (ELGs), due to typically being found in larger mass dark matter halos.

1.6 Extensions to Λ CDM: Neutrinos and Modified Gravity

1.6.1 Modified Gravity

While General Relativity has been an incredibly successful model in describing the physics on Solar System scales, it is not a perfect model. As described in Shankaranarayanan & Johnson (2022), there are four principal motivations to modify GR. Firstly, GR remains relatively untested on large cosmological scales. This means that a GR description of gravity with additional scalar, vector, or tensor fields is not yet ruled out. Further, the Solar System represents a weak gravity regime. In strong gravity regimes, such as near curvature singularities at the centres of black holes, GR can break down and require higher-order curvature considerations, which are important at Planck length scales (Brandenberger et al., 1993). Additionally, modified gravity models have been proposed to attempt to explain open questions with the Λ CDM model, including the mechanism behind the existence of late time expansion acceleration and early time inflation due to dark energy (see e.g., Hu & Sawicki, 2007). Lastly, a GR description of gravity is the only of four fundamental forces that has yet to be explained by a consistent, unified framework, leading the development and exploration of gravity models that are consistent with quantum mechanics. As a result many classes of modified gravity models have been proposed, each

tackling some of these open problems. However, each modified gravity model must be able to reproduce the accuracy with which GR describes Solar System physics. This often involves incorporating screening mechanics in regions of high density that force suppression of the GR modifications in overdense environments, such as Chameleon (Khoury & Weltman, 2004) or Vainshtein (Vainshtein, 1972). These screening mechanics will not be present in underdensities like cosmic voids, making them ideal candidates to study modified gravity models.

Observationally, the effect of modified gravity models like $f(R)$ (Hu & Sawicki, 2007) and normal-branch Dvali-Gabadadze-Porrati braneworld (Dvali et al., 2000, nDGP) in cosmological regimes that are relatively unscreened typically result in an enhanced form of gravity, which results in enhanced structure formation. This then results in enhanced galaxy clustering and lensing measurements. However, when performing measurements around overdense structure, the existence of screening mechanisms implies that modified gravity models can only be probed around the outskirts of dark matter halos. It is therefore of interest to study modified gravity models in low-density environments to remove the screening effects (Falek et al., 2018). The enhanced gravitational effects in unscreened environments from different modified gravity models typically result in different scale-dependencies. $f(R)$ models, for example, are typically stronger on smaller scales due to a resulting scale-dependence in the growth factor, while nDGP models have scale-independent growth, but the Vainshtein screening mechanism suppresses small-scale enhancements (Alam et al., 2021). Further modified gravity models, such as Galileon gravity (Nicolis et al., 2009) present pathways that affect both the growth of structure and the gravitational effect on relativistic particles, like photons. These models will then have an additional effect on gravitational lensing potentials, increasing the strength of the weak gravitational lensing signal directly in addition to the enhancement from the increased matter clustering (Barreira et al., 2015; Su et al., 2023).

1.6.2 Neutrinos

The existence of a cosmic neutrino background similar in scale to the cosmic microwave background is predicted from the Big Bang model. Furthermore, the standard model of particle physics predicts neutrinos to be massless, but this is at odds with neutrino flavour oscillation experiments. These experiments predict that at least two of the three neutrino mass states have non-relativistic mass. Further, solar neutrino studies in particular constrain the squared mass difference, $\Delta m_{1,2}^2$ between mass states 1 and 2 (e.g. Aharmim et al., 2013). Meanwhile, reactor and atmospheric neutrino studies constrain the squared mass difference $|\Delta m_{2,3}^2|$ between mass states 2 and 3 (e.g. Adey et al., 2018). Combining these two results allows for understanding of the hierarchy of neutrino mass states. However, these experiments are insensitive to the absolute

scale of neutrino mass, as well as which state is the most massive (due to the absolute value of $\Delta m_{2,3}^2$ constraints). This leaves two possible hierarchies of neutrino masses, either Normal Hierarchy, where the mass of state 3 is greater than the other two, or Inverted Hierarchy, where the mass of state 3 is less than the other two (Lesgourgues & Pastor, 2006). These two hierarchies differ in their predictions of the total sum of neutrino masses. Therefore, constraining the sum of the mass of the neutrinos would be an important discovery to test the standard model.

Laboratory experiments are limited in their ability to constrain this parameter. Since the cosmic neutrino background is expected to exist, if these neutrinos are in some part massive, then they will have some cosmological signature (see Lesgourgues & Pastor, 2006, for a review of cosmological effects of massive neutrinos). As they do not directly interact with light, they will appear as a form of hot dark matter – an invisible form of matter with high thermal velocities that act to smooth out and dampen neutrino matter perturbations smaller than their free streaming length (Agarwal & Feldman, 2011). However, since the mass of neutrinos comprise a small fraction of the total mass compared to the mass of dark matter, this effect is small within overdensities like dark matter halos. However, neutrinos do have an effect on the expansion history of the Universe, while not contributing much to the matter clustering. Since neutrinos at the time of their decoupling from the early Universe plasma are still relativistic, they act as an additional component of radiation that pushes the matter-radiation equality time to later times (Lesgourgues & Pastor, 2006). After becoming non-relativistic at a later time, within scales smaller than their free streaming length, neutrinos contribute to the expansion rate while not contributing to the matter clustering and growth of structure. This, in turn, means that CDM perturbations do not grow as efficiently, as not all of the surrounding matter is being pulled into the density seeds (Lesgourgues & Pastor, 2006). On scales larger than the free streaming length, however, neutrinos act similarly to CDM, and the matter clustering is indistinguishable between massive and massless neutrino cosmologies (Agarwal & Feldman, 2011). These neutrino effects will be observable in probes of matter distributions, including the matter power spectrum and weak lensing (Abazajian et al., 2011).

1.7 The Goal of Modern Cosmology

The goal of modern, low-redshift cosmology is thus to test the validity and determine the values of the cosmological parameters of the currently accepted standard model, Λ CDM. This is done by primarily measuring the values of the set of those parameters (e.g., H_0 , Ω_m , σ_8) across as many possible scales and as many possible redshift ranges to achieve greater and greater precision. When comparing results between low-redshift probes such as supernovae or galaxy lensing and clustering to high-redshift results like the CMB, we find significant tensions in two

of the parameters, H_0 and σ_8 . We note in particular for H_0 that due to this tension its value is typically written as

$$H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}, \quad (1.31)$$

where h carries the unknown value of Hubble constant. For this reason, distances are typically written in terms of $h^{-1} \text{ Mpc}$. In addition to the tensions, there are many unknowns that have cosmological implications, such as the mass hierarchy of neutrino particles or the deviations from General Relativity as represented by modified gravity models. Therefore, cosmologists also frequently test extensions to the ΛCDM models that incorporate possible models of these effects. It is therefore of great interest to not only increase the size and scope of our surveys to gather more and higher quality data, but to also develop new cosmological probes that can provide independent insight using existing data.

Cosmic voids are an ideal candidate for cosmological study for this reason. Voids, defined as underdense regions of the Universe, can exploit the empty regions of current and upcoming galaxy surveys to extract additional information from these data sets. Two choices of void probes have already published cosmological constraints from data: the void size function ([Contarini et al., 2023](#)), and the void-galaxy cross-correlation ([Hamaus et al., 2020](#); [Nadathur et al., 2020](#); [Woodfinden et al., 2022](#)). The goal of this thesis is to demonstrate that another potential void probe is detectable in current and future galaxy surveys with enough significance to motivate the future development of tools and methodologies needed to extract its cosmological information. The void statistic we wish to investigate is the matter profile of cosmic voids as measured by weak gravitational lensing.

Chapter 2

Weak Lensing Formalism

This chapter will introduce the theory and practical measurement techniques for measuring weak gravitational lensing signals for general mass distributions. We will focus primarily on the correlations between galaxies and lenses, and not on similar applications of lensing such as cosmic shear, which correlates background galaxy shapes with themselves. Weak lensing has historically been a valuable tool to examine the distribution of the underlying matter density in the Universe, as it is one of the few signals that are sensitive to the total matter distribution rather than a set of luminous biased tracers. As such, many wide galaxy surveys have been designed with maximizing galaxy shape detections in mind, such as KiDS ([de Jong et al., 2013](#)), DES ([Flaugher et al., 2015](#)), and UNIONS ([Gwyn et al., 2025](#)) in recent history.

2.1 Weak Lensing Theory

Gravitational lensing is a phenomenon where light from a background light source is deflected by gravitational potential from a foreground mass. This deflection in the path of photons creates distortions and magnifications in the image. Depending on the strength of the distortion, which partly depends on the projected angular separation of the source from the foreground mass (referred to as a lens), we classify lensing as either strong or weak. When applied to background galaxies, strong lensing events are relatively rarer and typically include arc-like distortions and multiple-imaging where it becomes plainly visible that an individual background galaxy shape has been distorted. Weak gravitational lensing, on the other hand, typically involves distortions that are indistinguishable from a typical galaxy shape. However, since these distortions across multiple background galaxies arise from a common foreground mass distribution, their distortions are then

correlated. This means that weak gravitational lensing can be measured statistically, either by measuring the correlation of background shapes with foreground positions (i.e. galaxy-galaxy lensing) to study properties of the foreground lens, or by measuring the correlation of shapes with themselves (i.e. cosmic shear) to study the distribution and clustering of all foreground matter. Further, since lensing is sensitive to the gravitational potential in the lens plane, it is therefore sensitive to the combination of dark matter and baryonic matter. This means that lensing is one of the primary methods for mapping the distribution of dark matter in the cosmic web.

We can write the angular displacement (α) of a photon from its original source angle (β) to its image angle (θ) using the lens equation

$$\beta = \theta - \alpha. \quad (2.1)$$

See Fig. 2.1 for a diagram of a typical source-lens system. Each component of α is then related to the projected gravitational potential (Φ) along the line of sight from observer to source (S)

$$\alpha^i = \frac{1}{c^2} \frac{\partial \Phi}{\partial \theta^i}. \quad (2.2)$$

From this, we can consider a single lens and use the Poisson equation to relate the unprojected potential to the mass density. We can additionally assume that the physical thickness of the lens is far smaller than the distances between the observer, lens, and source (the thin lens approximation) and that we can integrate the deflection along an unperturbed path (the Born approximation), which holds true as long as the deflection angle is small. From these assumptions we arrive at

$$\Phi(\theta) = \frac{c^2}{\Sigma_{\text{crit}}} \int \Sigma(\theta') \ln|\theta - \theta'| d^2\theta', \quad (2.3)$$

where we define the prefactor as a critical surface mass density

$$\Sigma_{\text{crit}} = \frac{c^2 D_S}{4\pi G D_{LS} D_L}. \quad (2.4)$$

Here, D_L is the angular diameter distance from the observer to the lens, D_S is the same distance to the source, and D_{LS} is the angular diameter distance between the lens and the source.

To determine the distortion effect on an extended source, we can calculate the Jacobian of the lens equation

$$\frac{\partial \beta^i}{\partial \theta^j} = \begin{pmatrix} 1 - \kappa - \gamma_1 & \gamma_2 \\ \gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \quad (2.5)$$

where we define

$$\kappa = \frac{1}{2c^2} \left(\frac{\partial^2 \Phi}{\partial \theta_1^2} + \frac{\partial^2 \Phi}{\partial \theta_2^2} \right) \quad \gamma_1 = \frac{1}{2c^2} \left(\frac{\partial^2 \Phi}{\partial \theta_1^2} - \frac{\partial^2 \Phi}{\partial \theta_2^2} \right) \quad \gamma_2 = \frac{1}{2c^2} \frac{\partial^2 \Phi}{\partial \theta_1 \partial \theta_2}, \quad (2.6)$$

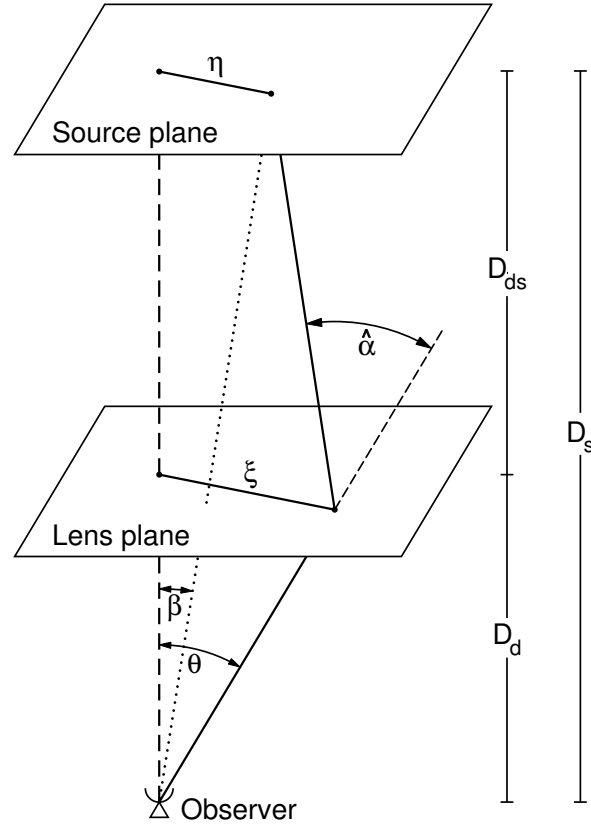


Figure 2.1: Ray tracing diagram showing the angular components of the gravitational lensing effect. β is the angle to the original source, $\hat{\alpha}$ is the lensing deflection angle, and θ is the angle to the observed lensed image. The angular difference between θ and β represented by α in Eqn. 2.1 is not labelled, but is the angle with its vertex at the observer between the dotted and solid line. Image credit: [Bartelmann & Schneider \(2001\)](#)

as the convergence and two components of shear, respectively. When applied to a galaxy shape, a lensing distortion with only positive convergence will magnify the image. A lensing distortion with only positive γ_1 will stretch the major axis and shrink the minor axis. Finally, a lensing distortion with only positive γ_2 will stretch an axis tilted 45 degrees from the major axis, and shrink along the corresponding perpendicular axis. However, these distortion effects take place on an axis system centred on the source. If we wish to detect the correlation between these shapes and a foreground matter distribution (such as in the case of galaxy-galaxy lensing or void-galaxy lensing), we need to define these distortions with respect to the lens position. To this end, we can define the following linear combinations, where ϕ is the angle between the x-axis and the source galaxy

$$\gamma_t = -\gamma_1 \cos(2\phi) - \gamma_2 \sin(2\phi), \quad (2.7)$$

$$\gamma_\times = \gamma_1 \cos(2\phi) - \gamma_2 \sin(2\phi). \quad (2.8)$$

γ_t is then a measure of tangential shear, or how closely the major axes of the sources tangentially trace out the foreground matter distortion. Important for void lensing, a negative tangential shear would indicate radial alignment of the sources with the foreground lens. $\gamma_\times$ then represents a similar alignment pattern across the sources, but with respect to an axis system tilted by 45 degrees. Foreground lenses are not expected to produce these types of shear patterns, so $\gamma_\times$ becomes a useful null diagnostic test.

We can then relate these lensing distortions to the surface mass density Σ of the foreground lens. Plugging in Eq. 2.3 into Eq. 2.6 for convergence leads to

$$\kappa(R) = \frac{\Sigma(R)}{\Sigma_{\text{crit}}}, \quad (2.9)$$

where R is the distance from the lens to the source projected to the plane of the lens. This then gives us an easy method to quantify the limit of the weak lensing regime relative to the critical surface density. The weak lensing regime occurs when $\Sigma \ll \Sigma_{\text{crit}}$, or $\kappa \ll 1$. We can also define γ_1 and γ_2 in terms of the surface density by plugging into the formula for γ_t . The end result is

$$\gamma_t(R) = \overline{\kappa(R)} - \kappa(R) = \frac{\overline{\Sigma(R)} - \Sigma(R)}{\Sigma_{\text{crit}}} = \frac{\Delta\Sigma(R)}{\Sigma_{\text{crit}}}, \quad (2.10)$$

where the bar indicates an average within the radius R . This then gives the definition of $\Delta\Sigma$, which is the primary weak lensing statistic that is used in this thesis.

2.2 Galaxy Shape Measurements

Now we turn to the question of how to extract the shear components from a galaxy image. We will first introduce the basic ideas of surface brightness moments and point spread function correction,

then describe more modern implementations through codes like NGMIX, which is used in one of the data sets for this thesis. In the simplest case scenario, one can obtain a measure of the galaxy ellipticity by considering the second moment of the weighted surface brightness profile of the image

$$Q_{ij} = \frac{\int d\theta q_I[I(\theta)](\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int d^2\theta q_I[I(\theta)]}, \quad (2.11)$$

where q_I is the weighting function applied to the surface brightness profile $I(\theta)$. We note that the observed surface brightness profile has two separate components: surface brightness due to an intrinsic ellipticity ϵ^S , and a lensing distortion component. To define the ellipticity in terms of these moments, we note that there are two common definitions used for ellipticity. We follow the definition used in [Bartelmann & Schneider \(2001\)](#) as it agrees with the definition used in the data throughout the thesis. Therefore, the components of the ellipticity are defined as

$$\epsilon = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}. \quad (2.12)$$

We then define the reduced shear parameter

$$g(\theta) = \frac{\gamma(\theta)}{1 - \kappa(\theta)}, \quad (2.13)$$

to relate the observed and intrinsic ellipticity components in the weak lensing regime (i.e. $g \ll 1$)

$$\epsilon = \epsilon^S + g \quad (2.14)$$

We further note that in the weak lensing regime, $g \approx \gamma$, as $\kappa \ll 1$. Given the assumption of isotropy in the Universe, we also have that $\langle \epsilon^S \rangle = 0$, so $\langle \epsilon \rangle = g$. This points to the statistical nature of weak lensing detections, where the more source galaxies that are measured, the smaller the contaminant of the intrinsic ellipticity gets, in this idealized scenario. This also means that one of the dominant sources of noise in weak lensing analyses is “galaxy shape noise,” or the spread in measured shapes caused by intrinsic ellipticity. This form of noise variance is Poisson distributed, and therefore decreases as $1/N_g$ where N_g is the total number of galaxy shapes. This has been the driving factor in constructing larger and larger galaxy imaging surveys to drive down shape noise for lensing analyses.

One assumption this idealization makes is that the measurement is devoid of observational systematics, such as aberrations or diffraction caused by, for example, variable weather conditions and airmass. Some of these systematics are more prevalent in ground-based telescopes that observe through an atmosphere than space-based ones, but both need to correct for these effects.

These observational systematics primarily act to smear and distort the shapes of light sources. Fortunately, nearby stars effectively act as point sources of light. Therefore, any corrections that return an observed star to a point source can be used to model and subtract out these observational systematics. We therefore measure the extent of these stellar ellipticities to model the Point Spread Function (PSF). We can then interpolate the PSF effects from nearby stars to the location of a galaxy image to approximate the contributions of the observed ellipticity due to observational systematics. This PSF contribution is then modelled as a convolution to the measured surface brightness profiles of the galaxy.

Modern implementations of shape measurement algorithms (such as NGMIX or *lensfit*) use a process called “forward modelling.” Instead of convolving the PSF directly with the noisy pixel data from observations, these algorithms instead assume a model of the surface brightness profiles that depend on a number of free parameters. They then try to find the set of parameters that minimizes (Farrens et al., 2022)

$$\chi^2 = \sum_{i=0}^{n_{\text{pixel}}} (I_i - G_i * P_i)^2 W_i, \quad (2.15)$$

where I is the surface brightness profile at a given pixel, G is the value of the model, P is the value of the PSF interpolation, and W is the pixel weight. This methodology instead convolves the PSF with the proposed model instead of the noisy observation, and tries to find the most likely model of the original surface brightness profile to explain the observation. There are still lingering biases and systematics in these shapes, so the shape measurement algorithms are calibrated using a metacalibration procedure (Huff & Mandelbaum, 2017). This procedure produces shape corrections by applying an artificial, known shear to an image and measuring the response of the shape algorithm when run on the sheared image. After correcting for these biases, each galaxy shape is then assigned a weight based on measurement uncertainties

$$w = \frac{1}{\sigma_{\text{SN}}^2 + \sigma_{\epsilon_1}^2 + \sigma_{\epsilon_2}^2} \quad (2.16)$$

where σ_{SN}^2 is the raw shape noise due to the spread in intrinsic ellipticity, and $\sigma_{\epsilon_1}^2$ and $\sigma_{\epsilon_2}^2$ are the uncertainties in the model fits for each individual galaxy.

2.3 Redshift Distributions

Lensing inherently assumes that we know every source is located behind a foreground lens. Given that when we take an image of the night sky, we observe a two-dimensional projection, how can we

be certain of this? As mentioned previously, we know that the expansion of the Universe stretches out the wavelengths of light in a process known as redshifting. We therefore need a sense of what the spectrum of light was like at the time of emission to gain an understanding of how far away a light source is. We can use templates of emission and absorption spectra where specific emission and absorption lines have known wavelength values in the laboratory rest frame. We can then take measured spectra from a galaxy and fit it to these various templates. This results in spectroscopic redshifts, which are accurate measurements of typical galaxies. These redshifts will be used to define the galaxy samples used to identify voids in this thesis. However, this process requires measuring the spectra of each individual targeted object, which is more observationally expensive than taking galaxy images. Therefore, for weak lensing imaging surveys which can target tens to hundreds of millions of galaxies, a faster method is preferred. Photometric redshifts require observing a target galaxy multiple times through different broadband filters. The differences in the brightness of the image through these different filters then gives the galaxy colours, which is the primary statistic used to determine the redshift, either through a template fitting method, or through machine learning methods. Imaging all galaxies in a patch of sky multiple times is much faster than taking spectra of each galaxy, at the cost of redshift accuracy, so photometric redshifts are preferred for modern surveys.

When applying weak lensing methodologies to photometric imaging surveys, we note that we cannot heavily rely on individual source galaxy positions. Instead, we can rely on their redshift distribution, $n(z_s)$. Since the source distance only enters into the formula for Σ_{crit} , we can then integrate over this redshift distribution to give

$$\langle \Sigma_{\text{crit}} \rangle = \frac{c^2}{4\pi G D_L} \int_{z_L}^{\infty} dz_s n(z_s) \frac{D_S}{D_{LS}}. \quad (2.17)$$

However, since the integral is not well defined for sources at the plane of the lens (i.e. $D_{LS} = 0$), we instead typically compute

$$\langle \Sigma_{\text{crit}}^{-1} \rangle = \frac{4\pi G D_L}{c^2} \int_{z_L}^{\infty} dz_s n(z_s) \frac{D_{LS}}{D_S}, \quad (2.18)$$

to avoid those numerical instabilities. If we have individual photometric redshift measurements for each galaxy, we can additionally construct a set of redshift bins for the sources, called tomographic bins. These smaller bins restrict the range of $n(z_s)$, allowing for a more careful selection of lens and source bin pairings such that the sources have a high likelihood of existing behind each lens. Further, it is common to use an inverse variance weighting scheme in lensing estimators like those considered in this thesis, where the lenses are weighted by $\langle \Sigma_{\text{crit}}^{-1} \rangle^2$. Applying tomography allows for individual source bins to be upweighted or downweighted relative to the given lens bin. Sources furthest from the lens bin are more strongly lensed, and are given more weight than sources right next to the lenses, which increases the Signal-to-Noise Ratio (SNR).

Chapter 3

Cosmic Voids

In this chapter, we will turn to the main object of study: cosmic voids. We will first begin by discussing voids as idealized spherical underdensities in the total matter distribution, and discuss models of their formation and evolution. We will also motivate the existence of the void lensing signal. We will then turn to identifying voids in observations, wherein we are only sensitive to underdensities in a set of luminous tracers, such as galaxies. We will then discuss previous attempts to measure the void lensing signal.

3.1 Idealized Voids

Cosmic voids are regions of the Universe that are underdense, meaning that they have less density than average. This density can be defined relative to the underlying matter distribution, or defined relative to a set of luminous tracers of the matter distribution, such as galaxies or quasars. They were first found from early studies of the distribution of galaxies around individual superclusters ([Gregory & Thompson, 1978](#)), where it was found that regions around the cluster were often devoid of galaxies, as well as early studies of galaxy distributions ([Jöeveer et al., 1978](#)), who found “holes” of 20 Mpc diameter with fewer than average galaxies.

Voids form when matter from the homogeneous Universe is pulled into the overdense structures of the cosmic web, comprising the near-empty regions in between the filaments, clusters, and walls. In an idealized scenario, a void grows by evacuating matter radially away from its central underdensity and towards its boundary, where the matter collects into what is called a compensating wall. This compensating wall is then itself overdense, which can further pull in external matter. In fact, the formation of overdense filaments and walls can be thought of in terms

of void growth. The expansion of multiple nearby voids can “compress” surrounding matter into sheets and filaments (Sheth & van de Weygaert, 2004). Further, voids are expected to become more spherical with time (Icke, 1984), and stacks of voids can be used as a standard shape to test cosmology.

The evacuation of matter from the centre causes initial underdensities to expand faster than the Hubble expansion at early times. Since matter is pushed to the boundary in shells, the shells nearest the central underdensity will be moving faster than those near the boundary, which are slowed by the increased density of the surrounding environment. Therefore, at some point the shells will cross. When this occurs, the void expansion decouples from linear growth. This means that the void stops expanding faster than the Hubble expansion, but the innermost density shells continue to evacuate matter towards the boundary. As was demonstrated in Bromley & Geller (2025), the evolution of the density profile is well-approximated by the Zel’dovich approximation after linear theory breaks down. This means that a non-linear density can be well approximated by linear theory down to the smallest scales even for voids identified in observational data (Hamaus et al., 2014b; Schuster et al., 2023), a unique feature of voids that is the result of shell-crossing occurring in the boundaries, not at the centre like in overdense structures.

When considering voids in a realistic density background, two broad classes of voids emerge: void-in-voids and void-in-clouds (Sheth & van de Weygaert, 2004). Void-in-clouds are underdensities that are embedded in a larger-scale overdense environment. As demonstrated by excursion set theories of halos (Bond et al., 1991) when solving the cloud-in-cloud problem of halo mass functions, the largest scale that satisfies the condition of structure formation (i.e. a value of linear density) subsumes all smaller-scale fluctuations. Thus, a void-in-cloud is an underdensity that will be squashed by the surrounding collapsing overdensity in some future time. A void-in-void on the other hand is simply a void that exists within some broader underdense region. These voids will typically exhibit weak overdense structure in their walls. Often, these walls will contain less matter density than is required to fill the void back in, a property referred to as undercompensated. Void-in-clouds are mostly smaller voids, and void-in-voids tend to be larger voids, therefore the void radius can be a useful, but not perfect (as some small voids can be void-in-void systems as seen in Nadathur & Hotchkiss (2015a)), metric to separate these systems.

How matter is distributed within voids is of cosmological interest. The expected depth of voids is a potential probe into neutrino masses and warm dark matter models, as these will not only allow mass to free stream back into the void, producing shallower typical voids, but also suppress structure growth around the scale of the void, making voids evolve like younger or smaller counterparts in massless neutrino cosmologies (Massara et al., 2015; Schuster et al., 2019; Kreisch et al., 2019; Vielzeuf et al., 2023). On the other hand, modified gravity models will act to empty matter out more efficiently, producing deeper voids on average (Cai et al., 2015; Barreira et al., 2015, 2017; Cautun et al., 2018; Paillas et al., 2019; Su et al., 2023). In particular

for modified gravity, voids provide a unique environment where the modified gravity effects are not screened (i.e. appear as standard General Relativity). This means that void matter profiles can carry modified gravity signals at the smallest scales, something that would normally be suppressed by screening mechanisms in overdensities. The rate of outflow of matter from the void centres are also sensitive to the cosmology (Bromley & Geller, 2025). Further, different combinations of overdense and underdense cosmological probes have been found to be particularly beneficial to breaking parameter degeneracies, as well as providing tighter constraints overall (Bayer et al., 2021; Kreisch et al., 2022; Pelliciari et al., 2023; Thiele et al., 2024; Contarini et al., 2024), including void matter profiles (Salcedo et al., 2025).

Since the majority of the matter in the Universe is not visible, we rely on weak lensing to measure its distribution. However, lensing has typically required the presence of an overdensity to create the gravitational potential required for lensing. However, we can imagine the underdensity of a void as a negative overdensity. This then allows for an intuitive understanding of the lensing effect that a void has on a given background source. Instead of an image deflecting away from a lens, stretching into tangential alignment, we instead reverse the direction of the distortion. The image is now deflected towards the underdensity, pulling a circular shape into a radially-aligned ellipse. This radial alignment will then appear as a negative value of the tangential shear, which then carries into a negative value of the $\Delta\Sigma$ profile. This effect is illustrated in Fig. 3.1. Since voids, which cannot be emptier than a pure vacuum, have an underdensity bounded by -1, they are expected to produce a much weaker lensing effect than even a typical galaxy lens, as virialized masses have overdensities on the order of magnitude of 200 or higher. Further, since galaxy-galaxy lensing produces percent-level distortions, this means that void lensing is expected to be a weak effect. This is one of the main reasons why void lensing has needed modern surveys in order to measure this signal.

3.2 Voids in Observations

Given most of the matter is dark and not easily accessible to the observer, how do we identify voids in observations? In an observational framework, we are restricted to relying on identifying underdensities in luminous, biased tracers. These are commonly voids in galaxy populations, but can be others, such as voids in clusters (Sutter et al., 2014; Pollina et al., 2019), or voids in quasars used to probe high-redshift underdensities (Hawken et al., 2020; Aubert et al., 2022; Kovács et al., 2022; Arsenov et al., 2025). There exist many different methods to identify the extents of these underdensities, and each method has a significant impact on the resulting void statistics, including the void profiles. As such, there exists a number of void finding algorithms in the literature (see, e.g. Colberg et al., 2008; Cautun et al., 2018; Paillas et al., 2019; Massara et al.,

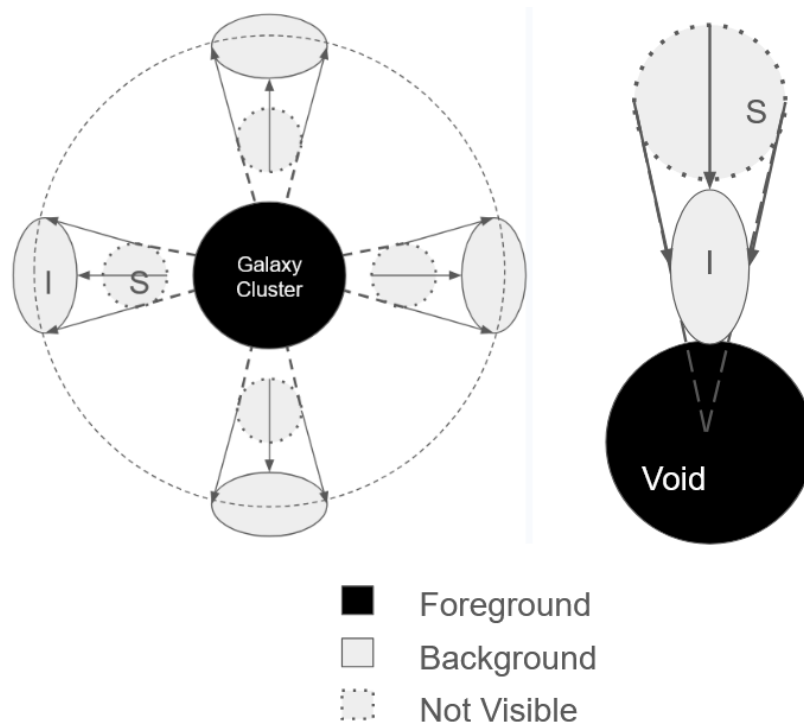


Figure 3.1: Diagram of the lensing effect on a circular source from both an overdensity and a void. For the void, light rays along the sides of the source are pulled radially in, compressing the sides of the image into a radially aligned ellipse.

2022, for comparisons of different void finders). For the purposes of this thesis, we will focus on two: `VIDE` and `Revolver`. Both algorithms are based on the ZOnes Bordering on Voidness (ZOBOV) algorithm (Neyrinck, 2008) and differ in two main regards. We will first describe the ZOBOV algorithm, then the differences.

ZOBOV algorithms work by first performing a Voronoi tessellation in the tracers. Each Voronoi cell contains exactly one tracer and all spatial points closer to that tracer than any other tracer in the sample. This tessellation comes with the beneficial property that regions devoid of many tracer points contain predominantly large Voronoi cell volumes, while regions with many tracer points contain predominantly small cell volumes. Cells are then merged together into zones using a watershed algorithm. Neighbour cells with decreasing volumes are merged until encountering a neighbour cell that has a larger volume than the previously merged cell. Each zone is then identified as a void. Zones can in principle be merged together, particularly if the density wall separating them is small, but using the individual zones as voids has been favoured by recent literature (Nadathur & Hotchkiss, 2015a). For this thesis, no merging is applied to any of the void catalogues. ZOBOV is modified for use in surveys by applying a set of boundary particles around the edges of the survey footprint, and around any holes present within the survey to prevent voids from leaking into arbitrarily low-density regions. Any void that touches a boundary particle is additionally removed. Each void will be highly non-spherical, but the average shape across voids should be spherical under isotropy. Therefore, a spherical definition of the void radius is applied by assigning the radius of a sphere of equivalent volume

$$R_V = \left(\frac{3}{4\pi} V_{\text{void}} \right)^{-1/3}. \quad (3.1)$$

See Fig. 3.2 for an example of three different void finders, including ZOBOV, applied to the same N-body simulation slice.

Both `Revolver` and `VIDE` implement ZOBOV. However, `Revolver` includes an additional reconstruction step (detailed in Sect. 4.3.1) before ZOBOV to remove approximate redshifting effects caused by peculiar velocities from the tracer sample before finding voids. However, as lensing is only sensitive to the projected gravitational potential, it is expected that this will have little effect on the resulting void lensing profiles and is more relevant to studies of the void-galaxy cross-correlation, which relies on the spherical nature of averaged voids as a standard shape. The relevant difference of note for this thesis is the definition of void centre. `VIDE` uses a barycentre definition, where the centre is defined as the average position of tracers weighted by their corresponding Voronoi volumes. `Revolver` uses the circumcentre definition, which is found by taking the average position between the tracer with the largest cell and three neighbouring tracers. It is found that the circumcentre definition more accurately finds the minimum density point of the void to be the centre (Nadathur & Hotchkiss, 2015a), but given the non-spherical

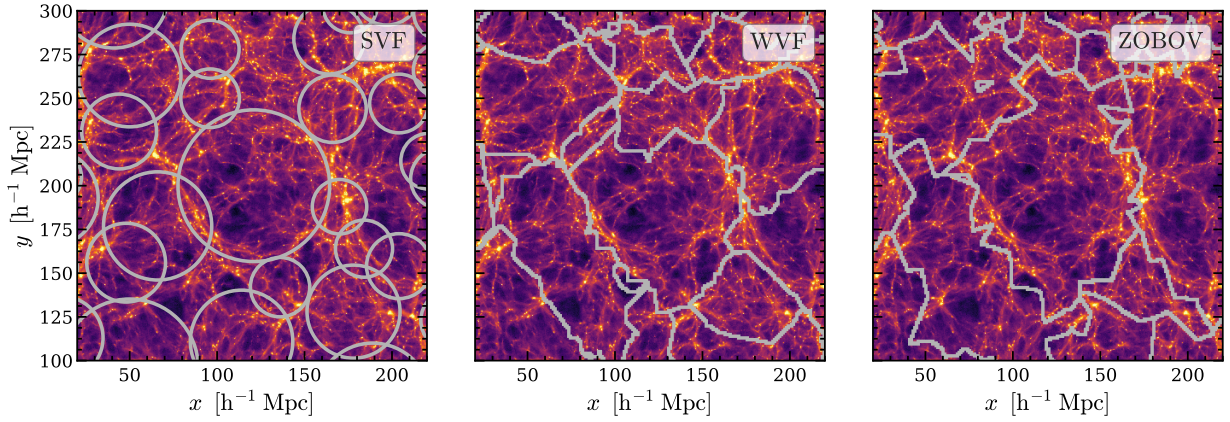


Figure 3.2: Examples of voids found by three different void finders in the same $50h^{-1}$ Mpc slice of a N -body simulation. The three void finders are the Spherical Void Finder ([Padilla et al., 2005](#)), which assumes a standard spherical shape, the Watershed Void Finder ([Platen et al., 2007](#)), which is a variant of ZOBOV with a different tessellation scheme, and the ZOBOV finder described in this chapter. Image credit: [Paillas et al. \(2019\)](#)

shapes of ZOBOV voids, this centre can be defined in different places relative to the wall. Therefore, in spherical averages, circumcentre voids present with washed out overdense signals in the walls. Barycentre voids, on the other hand, are defined more strongly relative to the walls (where there are more tracers), so in the spherical average, the wall is more defined and the central underdensity is washed out. In some cases, the centre of the profile of barycentre voids can even turn upwards, resulting in centres that are more dense than the immediate surrounding area ([Massara et al., 2022](#)). Much like different void finders, both definitions of void centres have their merits and drawbacks, and as such there is no accepted standard.

Observational voids are also dependent on properties of the tracers they are found in. For example, a highly biased tracer will more strongly cluster around peaks in the matter density field. Therefore, a catalogue of voids found in these tracers will present with much larger radii as tracers will not appear around weaker density peaks ([Pollina et al., 2019](#)). The more sparse a sample, the larger the voids will be as well ([Sutter et al., 2014](#)). Further, it was demonstrated in [Fang et al. \(2019\)](#) that if the line of sight distance to the tracers is estimated from photometric redshifts instead of spectroscopic redshifts, the increased uncertainty in the photometric redshifts produces a selection bias in the void finding step. This selection bias prefers voids that are stretched along the line of sight. This has important ramifications for void lensing, where we are sensitive to the projected mass density. A sample of voids aligned along the line of sight produces deeper projected underdensities, increasing the lensing signal at the cost of statistical

isotropy. Effects like these are important considerations to make when comparing results from different void papers, and typically precludes direct comparisons.

Because of the very weak signal that voids impart on background galaxies, there are only a small number of published detections of void lensing in data. The first such detection was made by [Gillis et al. \(2013\)](#), who were studying environmental dependence of satellite galaxy masses measured from weak lensing. As part of that analysis, they measure negative $\Delta\Sigma$ signals, detecting underdense environments. [Melchior et al. \(2014\)](#) is the first dedicated void lensing study, and is the work most relevant for comparisons for this thesis. They found voids using a ZOBOV void finder in the spectroscopic sample of the Sloan Digital Sky Survey (SDSS) Data Release 7 (DR7). They fit the measured lensing signal with a parameter-free model, finding a significance of 2.4σ . [Clampitt & Jain \(2015\)](#) used their own void finder, which found projected two-dimensional underdensities in tomographic redshift slices. These voids would therefore appear more like truncated tunnels. They measured lensing around voids found on just the LRG portion of the DR7 data, finding a SNR of 7. A similar version of this void finder was applied to photometric data of the Dark Energy Survey, first in [Sánchez et al. \(2016\)](#), and then [Fang et al. \(2019\)](#) for Year 1, finding a SNR of 4.4 in the former, and 10.7 in the latter. [Fang et al. \(2019\)](#) additionally applied ZOBOV to the photometric redshift sample, finding a SNR of 14.0 in their lensing profile. Void profiles were also found in Year 3 in [Jeffrey et al. \(2021\)](#), but no significance measurements were reported. The effect of void lensing has also been measured in the CMB in a number of different studies ([Cai et al., 2016](#); [Chantavat et al., 2017](#); [Raghunathan et al., 2020](#); [Hang et al., 2021](#); [Vielzeuf et al., 2021](#); [Kovács et al., 2022](#); [Camacho-Ciurana et al., 2024](#); [Sartori et al., 2025](#)). Further, lensing relative to regions of the lowest percentile of galaxy density has been measured as part of density-split statistics ([Gruen et al., 2015](#); [Brouwer et al., 2018](#); [Gruen et al., 2018](#); [Burger et al., 2023](#)).

Chapter 4

Data

For the purposes of this thesis, we consider two different groups of data sets, one to make a measurement with currently existing data to showcase what significance of the void lensing signal is currently detectable at, and a simulated data set to forecast how well expected future data can improve significance of this detection. In this chapter, we will introduce these data sets, starting with the SDSS BOSS and UNIONS surveys, then the simulation data used to forecast *Euclid* results. We use the void catalogues of [Woodfinden et al. \(2022\)](#) in BOSS data and [Hamaus et al. \(2022\)](#) in the Flagship simulation of *Euclid*, who both constructed the void catalogues for use in measuring the void-galaxy cross-correlations. We will additionally detail efforts that were made to assist in the production of the UNIONS data at the photometric redshift level in Sect. 4.2.

4.1 UNIONS Galaxy Shape Catalogue

The catalogue of background source galaxy shapes that are used in Chapter 5 are from the Ultraviolet Near-Infrared Optical Northern Survey (UNIONS; [Gwyn et al., 2025](#)). UNIONS is a planned 6250 deg² multi-band survey in the northern hemisphere, with about 3200 deg² completed for use in this thesis. This is the largest lensing survey of the northern hemisphere to date, and offers significant overlap with spectroscopic surveys like SDSS and DESI. This overlap makes UNIONS an ideal choice for the analysis of Chapter 5. UNIONS is comprised of the CFIS, Pan-STARRS, and WISHES surveys, each providing data in different broad band filters. In the future, these data sets will be combined for full *ugriz* photometry, but individual photometric redshifts and lensing tomography are in preparation and therefore currently unavailable for this work. The source galaxy shapes are measured using the *r*-band data provided by CFIS using the

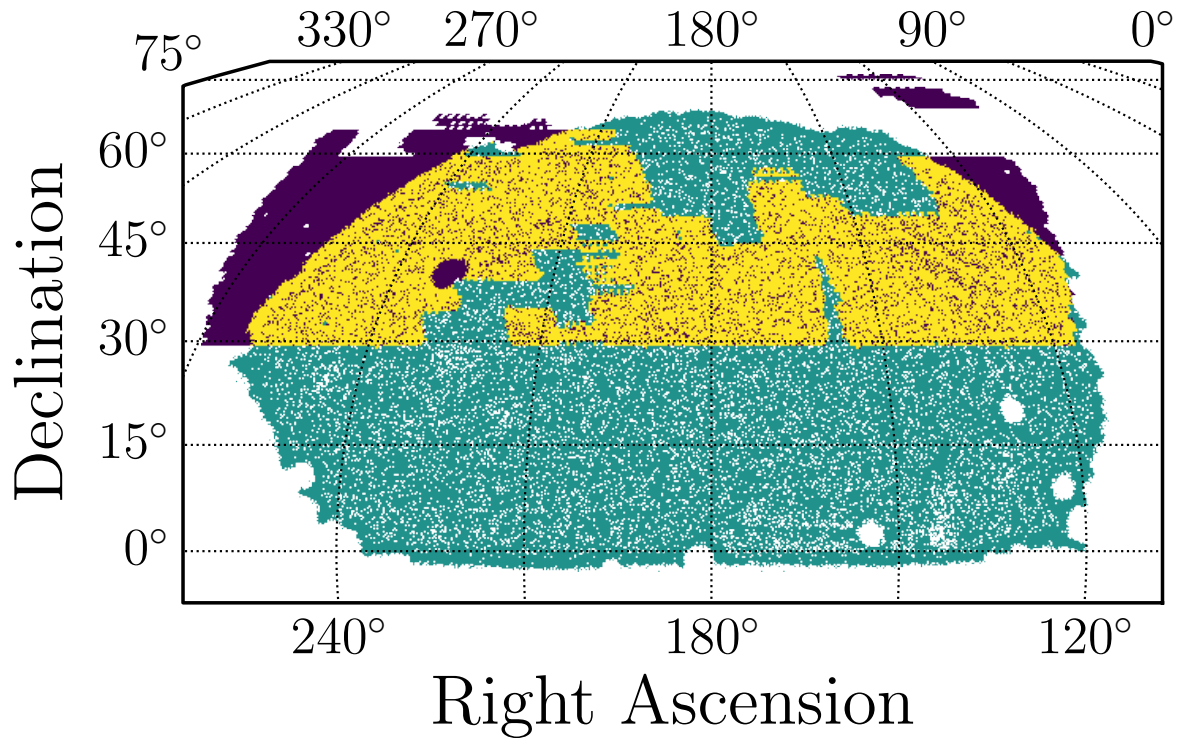


Figure 4.1: Footprints of the two surveys used in this analysis and their overlapping region. The UNIONS footprint at the time of writing is in purple, and the CMASS footprint is shown in light blue. The overlapping region of the two surveys is shown in yellow.

Canada-France Hawai’i Telescope (CFHT) on Mauna Kea. The r -band data is planned to reach a depth of $r = 24.2$ at $\text{SNR} = 10$ with ≈ 0.7 arcsec seeing (Gwyn et al., 2025). This is ideal for measuring galaxy shapes for weak lensing measurements. The footprint of the UNIONS survey showing the overlap with the foreground SDSS BOSS CMASS survey is shown in Fig. 4.1

As described in Li et al. (2024), in order to build the redshift distribution, the UNIONS survey uses the overlap between the UNIONS r -band footprint and the CFHTLS-W3 field (Heymans et al., 2012). UNIONS galaxies in the overlap are then matched to their corresponding galaxy image in the public W3 field to assign full $ugriz$ photometry found in the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS; Hildebrandt et al., 2012). This is then used to populate a self-organizing map (SOM) trained on publicly available overlapping spectroscopic surveys (Wright et al., 2020). The spectroscopic redshift distribution is then reweighted by the ratio of

UNIONS galaxies to spectroscopic galaxies in each SOM cell, resulting in a calibrated source redshift distribution. Then, given that UNIONS has a constant depth across the survey area, we are justified in using this redshift distribution as representative across the entire survey. This distribution is shown along with the redshift distribution of foreground lenses in Fig. 4.11.

The shapes of these source galaxies were measured using version 1.3 of the ShapePipe pipeline, the details of which are outlined in Guinot et al. (2022) and Farrens et al. (2022). Version 1.3 of the pipeline uses the MCCD (Liaudat, T. et al., 2021) software package to model the PSF (Li et al., 2024; Guerrini et al., 2025; Zhang et al., 2024), and continues to use NGMIX to perform METACALIBRATION (Huff & Mandelbaum, 2017). The result is a catalogue that includes the sky position, the two ellipticity components ϵ_1 and ϵ_2 , as well as an inverse variance weight that accounts for the quality of the shape measurement:

$$w_s = \frac{1}{2\sigma_\epsilon^2 + \sigma_{\epsilon_1}^2 + \sigma_{\epsilon_2}^2}, \quad (4.1)$$

where σ_ϵ represents the intrinsic dispersion of one component of the ellipticity, set to the same value of 0.34 as used in Guinot et al. (2022), and the other two terms are the measurement errors in each corresponding ellipticity component from a model fit. The 1.3 version of this catalog includes 83 812 739 total galaxy shapes and represents the status of the data as of 2022.

4.2 UNIONS Multiband Photometry and Quality Control Tests

The development and release of photometric redshift estimates are a key development step for the UNIONS survey. While the calibrated redshift distribution allows for accurate weak lensing measurements, the full photo- z data will allow for the construction of tomographic bins, which as mentioned in Sect. 2.3, improves signal-to-noise ratios by not only isolating the sources that are behind the lens, but also by implementing a weighting scheme that upweights the furthest sources with the highest lensing signals.

In general, there exist two popular methods to obtain photometric redshift information: template-fitting and machine learning methods. In UNIONS, we use the former. Both methods fit redshifts to measurements of galaxy colours, which in turn require galaxy photometry to measure the flux relative to the broad band filter. In UNIONS, the GAaP photometry code is used (Kuijken et al., 2015) to measure the fluxes in the five *ugriz* bands. Starting with the pre-seeing image $I_{\text{pre}}(\mathbf{x})$, the expected output of GAaP is the Gaussian-weighted flux through an aperture represented by a weight matrix $\mathbf{W}$

$$F_{\mathbf{W}} = \int d\mathbf{x} I_{\text{pre}}(\mathbf{x}) e^{-\frac{1}{2}\mathbf{x}^T \mathbf{W}^{-1} \mathbf{x}}. \quad (4.2)$$

The pre-seeing image is not directly measurable, but the equation can be rewritten in terms of the observed image convolved with a Gaussian PSF (see Eq. 16 in [Kuijken et al. \(2015\)](#)). Therefore, the GAaP code attempts to construct a kernel that transforms the measured PSF into one that is approximately Gaussian first. Then, the fluxes of each luminous object are measured in each of the five bands in this Gaussianized space using an elliptical Gaussian aperture. Since this aperture has a finite size, these fluxes and their corresponding magnitudes are not total magnitude measurements, as only the part of the extended object inside the aperture is used. However, for stars, which can be treated as point sources, these measurements do represent total magnitudes. The size of the aperture is fit to each object to maximize SNR, but the aperture size does not go below a chosen minimum size. Two runs were performed on the UNIONS catalogues with two different values for the minimum aperture size, 0.7 arcsec and 1.0 arcsec. The best result is chosen of the two for the photometry, but both values are recorded in the catalogues. Since stars are point sources, the resultant magnitude measurements do not change between shifts in minimum aperture size ([Kuijken et al., 2019](#)). This property will be useful later when performing star-galaxy separation.

Once the fluxes and their corresponding magnitudes are measured for each of the five bands, they are input into the redshift finding code. UNIONS will use the BPZ code ([Benítez, 2000](#)), which performs likelihood fits the observed galaxy colours obtained by subtracting the magnitudes from neighbouring filters (e.g. $u - g$) to a series of templates constructed by modelling galaxy spectral energy distributions (SEDs). The magnitudes are also used directly in BPZ to build Bayesian priors for the expected distribution in redshift of a galaxy with a given magnitude. Using the broad band filters, the SEDs and their redshifted values can be translated to template colours that can fit the observed colours using a Bayesian fitting algorithm to incorporate known priors.

The current state of UNIONS photometric redshifts is that the GAaP and BPZ codes have had an initial run, and now the resulting photometry and redshift measurements need to be checked for quality before being used in cosmological analyses. These quality control tests will also help inform final decisions regarding which data sets and objects to keep and use in the final data set, and which to cut from the analysis entirely. To assist in this effort, we produced a series of plots requested by the collaboration that examined the statistics of the photometry and redshift on the sky. The UNIONS data is split into individual tiles, and each tile contains a multitude of luminous objects. Each object then has its corresponding photometry measured and a photometric redshift assigned by the corresponding codes. This includes a limiting magnitude that was measured from the faintest detections in the tile for each band. Each object then received a limiting magnitude value that corresponds to the optimal aperture the object was measured with. This limiting magnitude measurement is then a good test for the depth of the survey at a given tile and a quality test for the raw observations. Regions with unexpectedly shallower observations (smaller

magnitude limits) are then flagged as regions of concern.

Examples of these tile statistics plots are shown in Figs. 4.2 through 4.8, showing the median limiting magnitude each of the five bands with the median redshift measurement per tile at the end. The z -band coverage was split between the Pan-STARRS and WISHES surveys. We note that these plots represent a 5000 deg^2 area, which is larger than the footprint seen in Fig. 4.1. This larger region at the time did not have shape measurements available. Therefore, the smaller 3200 deg^2 region will be used for the lensing analysis presented in Chapter 5, but the larger region seen here represents future data releases.

We can observe some of the anomalous regions and failure modes in these plots. The deep red tiles seen in, for example, the majority of the z -band plots indicate no photometry was taken in that band on those tiles. This is either due to no imaging observations in that band, or due to unexpected code failures. This was specifically noted as a concern in the arc-like streaks of red tiles seen in the i -band at high declination in Fig. 4.4. This was an unexpected result that prompted further investigation from the photo- z team. We can also observe other features, such as the deep field patch present in the top left corner of the g -band plot in Fig. 4.3, as well as particularly shallow regions, like the lower-right corner of the i -band plot. In a previous version of the data a localized region of the g -band observations was found to be unexpectedly shallow. This prompted further investigation and was improved in the current version. We can additionally identify anomalous regions in the redshift plot, such as the blue arc-like streaks bounding the lower-left corner in Fig. 4.8. Looking to the limiting magnitude plots, we find that multiple bands have deep red tiles in those arcs, like the u , g , and r bands. Tiles like these would likely be removed by the collaboration before performing any scientific analysis.

We can additionally separate out point sources from extended objects. Using the two different minimum aperture values, a point source will remain unchanged relative to the error in the measurement. Therefore, the difference in magnitude between the smaller aperture and larger aperture can help classify objects as stars or non-stars. This was shown in Kuijken et al. (2019), where they further demonstrated using a stellar locus plot that the objects identified as stars using the aperture difference method were in fact stars. We perform this separation in the r band, which are the most reliable measurements and show an example of the star-galaxy separation plot in a single tile shown in Fig. 4.9. The orange points are identified as stars, with the exception of the bright end with magnitudes less than 16. These objects are too bright to be galaxies, but are not point sources. They are likely stars that are too close to be considered point sources for these aperture sizes, and will likely be cut from the final analysis. The faint end of the separation is arbitrarily chosen to be at 21 magnitude. This results in a pure sample of stars, at the cost of labelling some faint stellar objects as galaxies, but this choice is kept free until the collaboration makes a final decision. Refining the tile statistic plots to include only stellar objects will allow for diagnostic testing of the photometric zero-points.

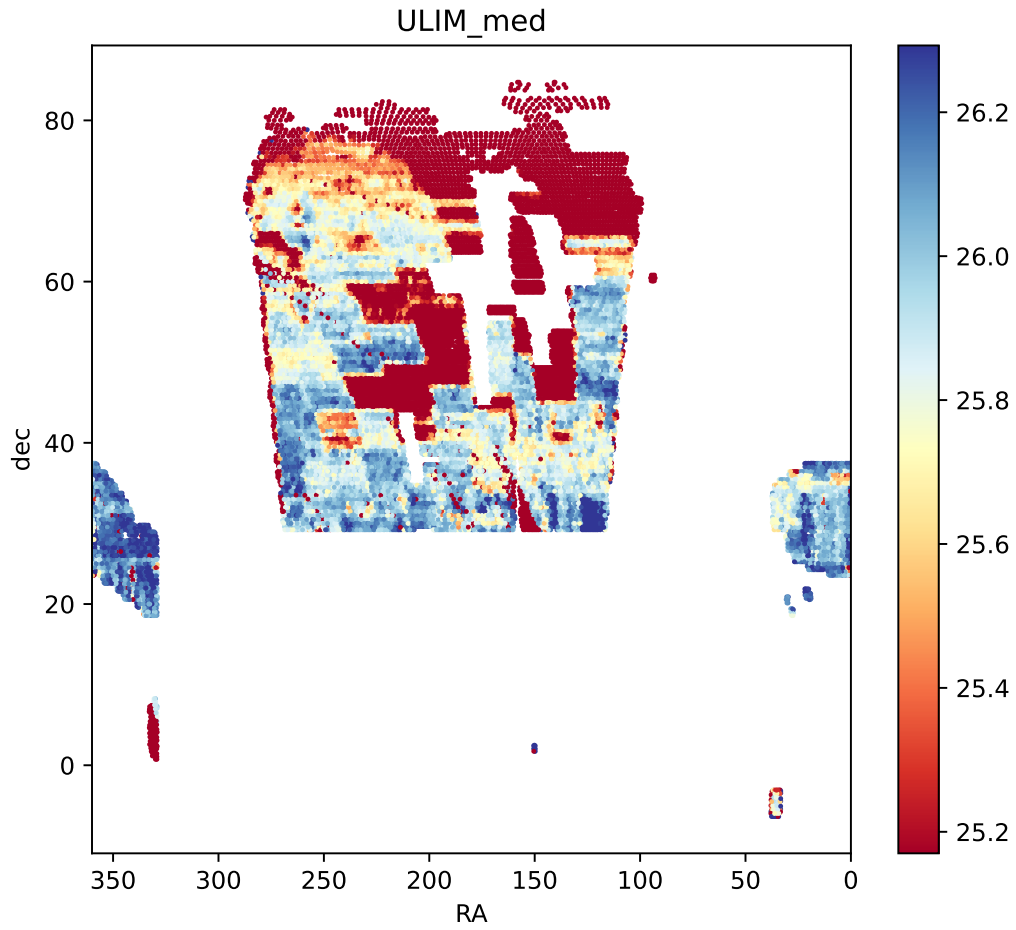


Figure 4.2: Plot of the median limiting u -band magnitude per individual tile of the UNIONS survey. Objects in dark red represent either statistical outliers, or regions where no data is found.

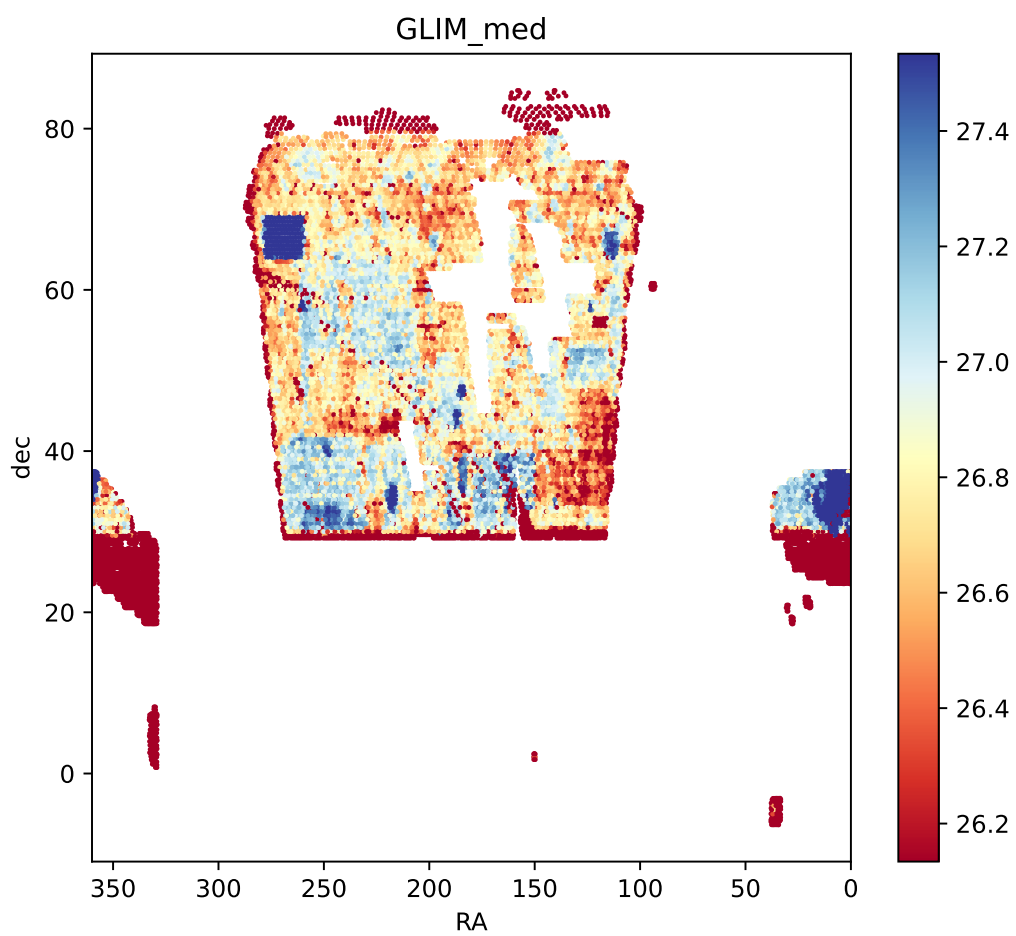


Figure 4.3: Same as Fig. 4.2, but for the limiting g -band magnitude.

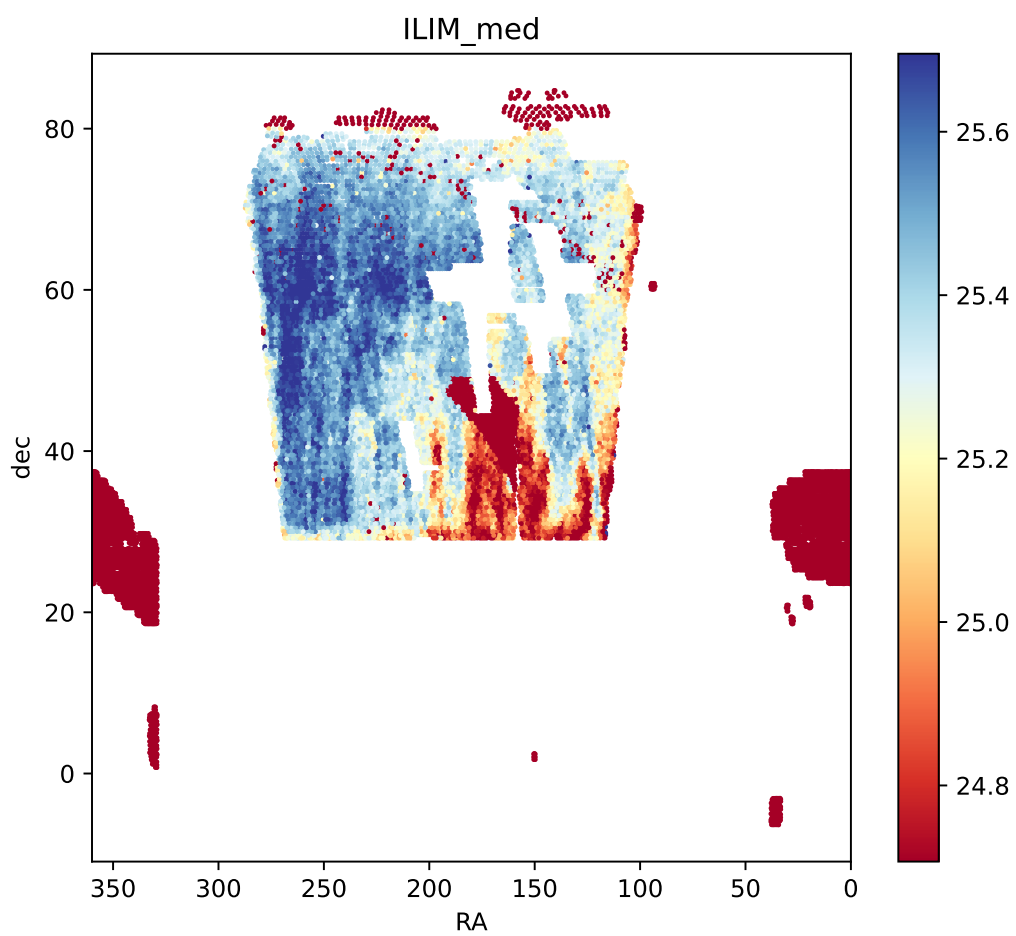


Figure 4.4: Same as Fig. 4.2, but for the limiting i -band magnitude.

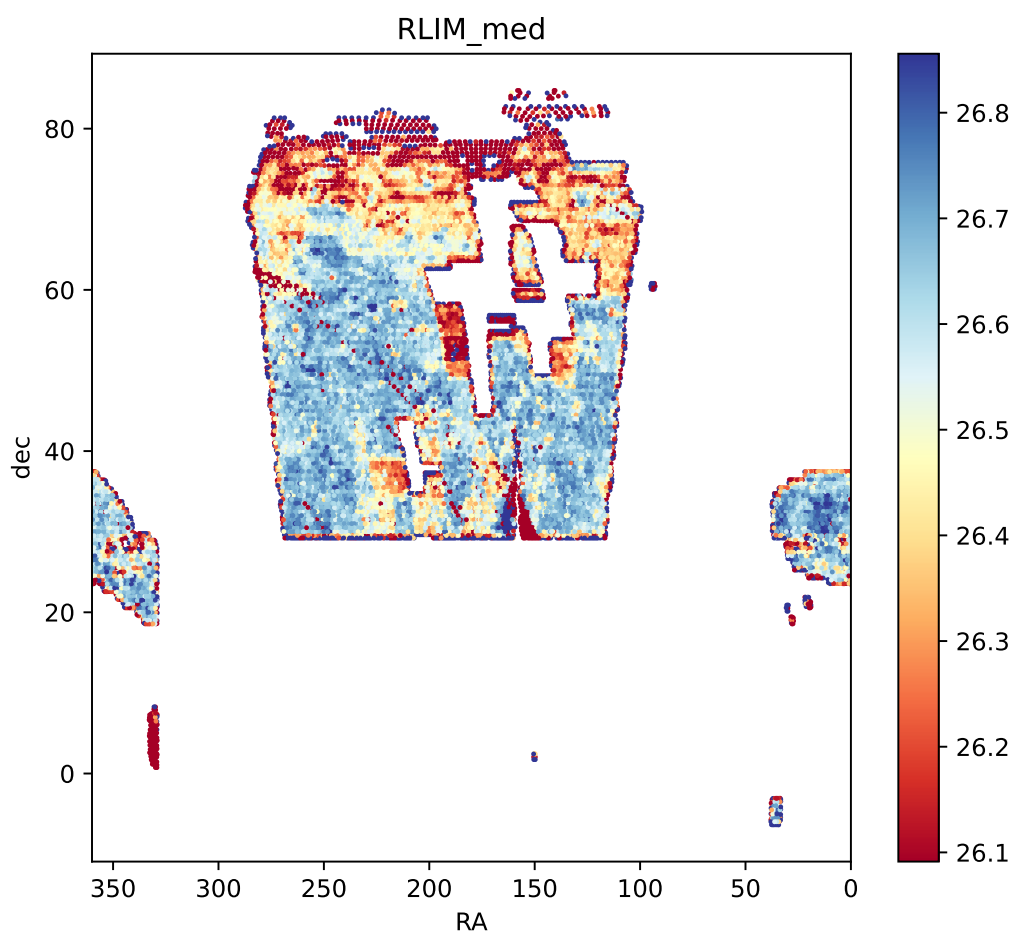


Figure 4.5: Same as Fig. 4.2, but for the limiting r -band magnitude.

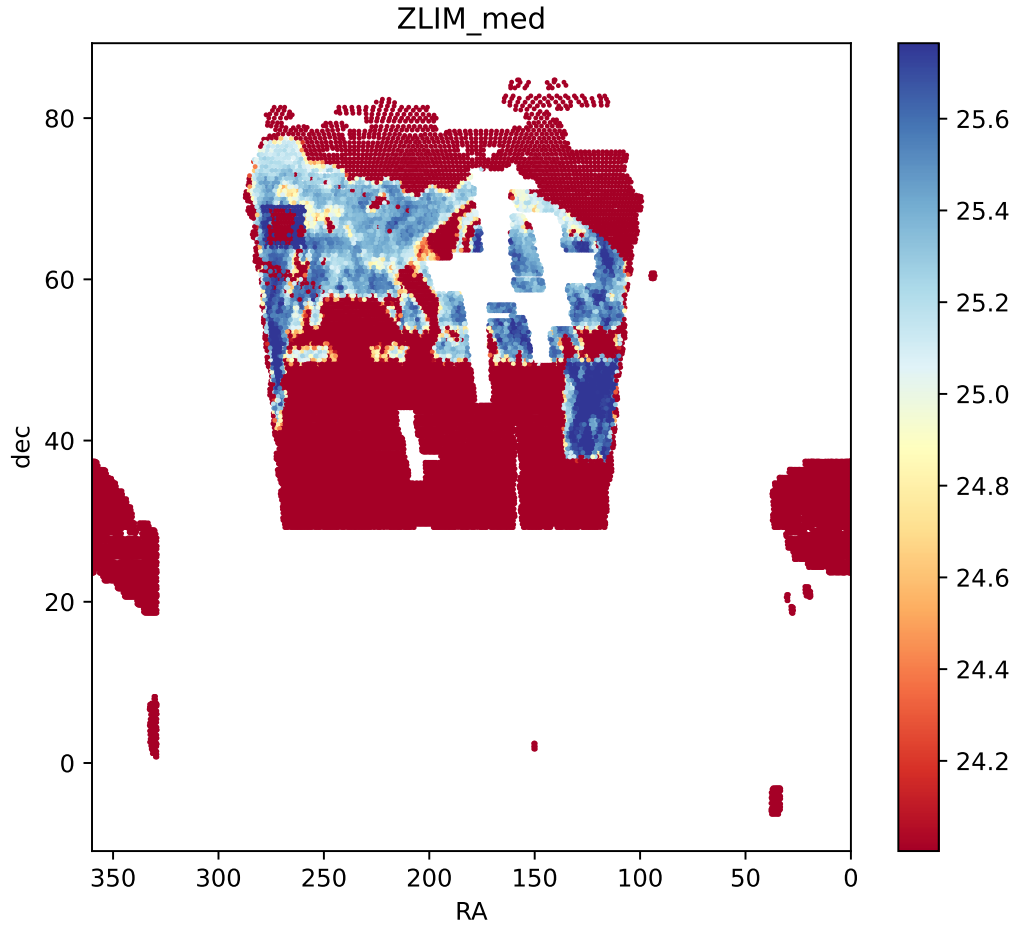


Figure 4.6: Same as Fig. 4.2, but for the limiting z -band magnitude. This plot shows the WISHES survey component of the z -band limiting magnitude that will be used in conjunction with the Pan-STARRS component.

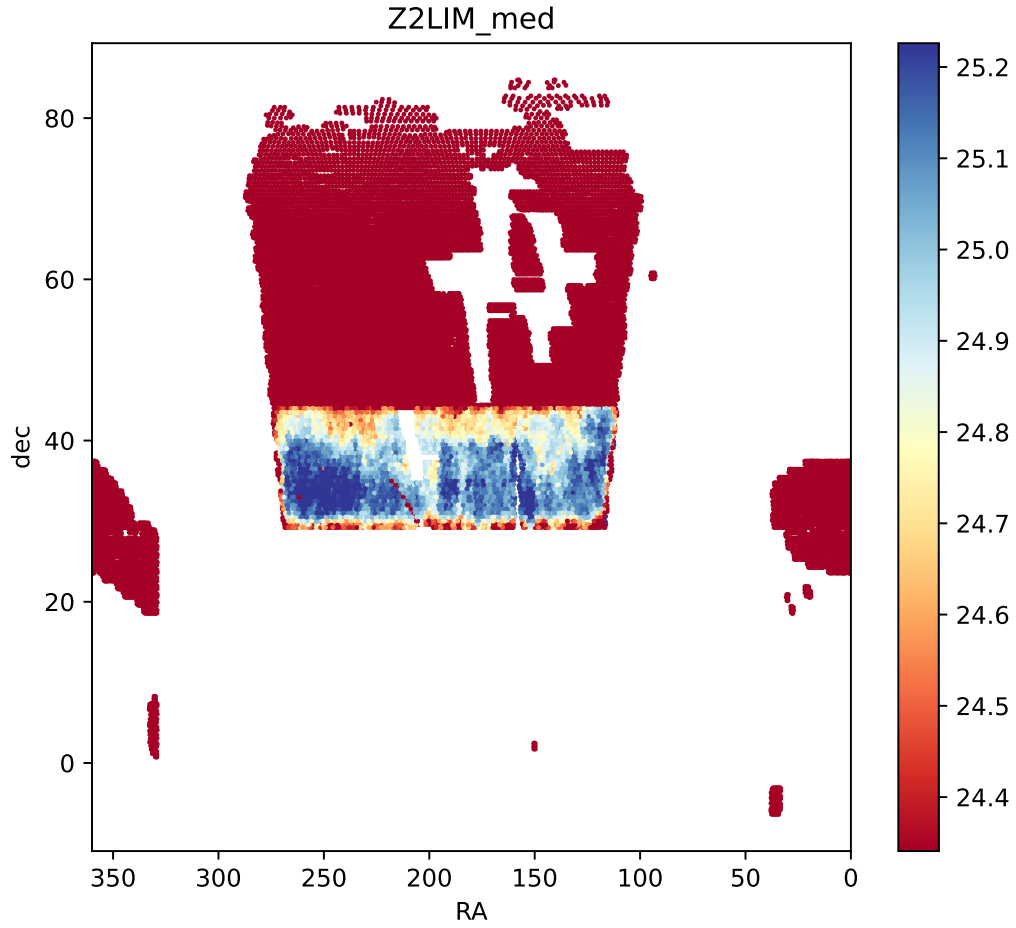


Figure 4.7: Same as Fig. 4.2, but for the limiting z -band magnitude. This plot shows the Pan-STARRS survey component of the z -band limiting magnitude that will be used in conjunction with the WISHES component.

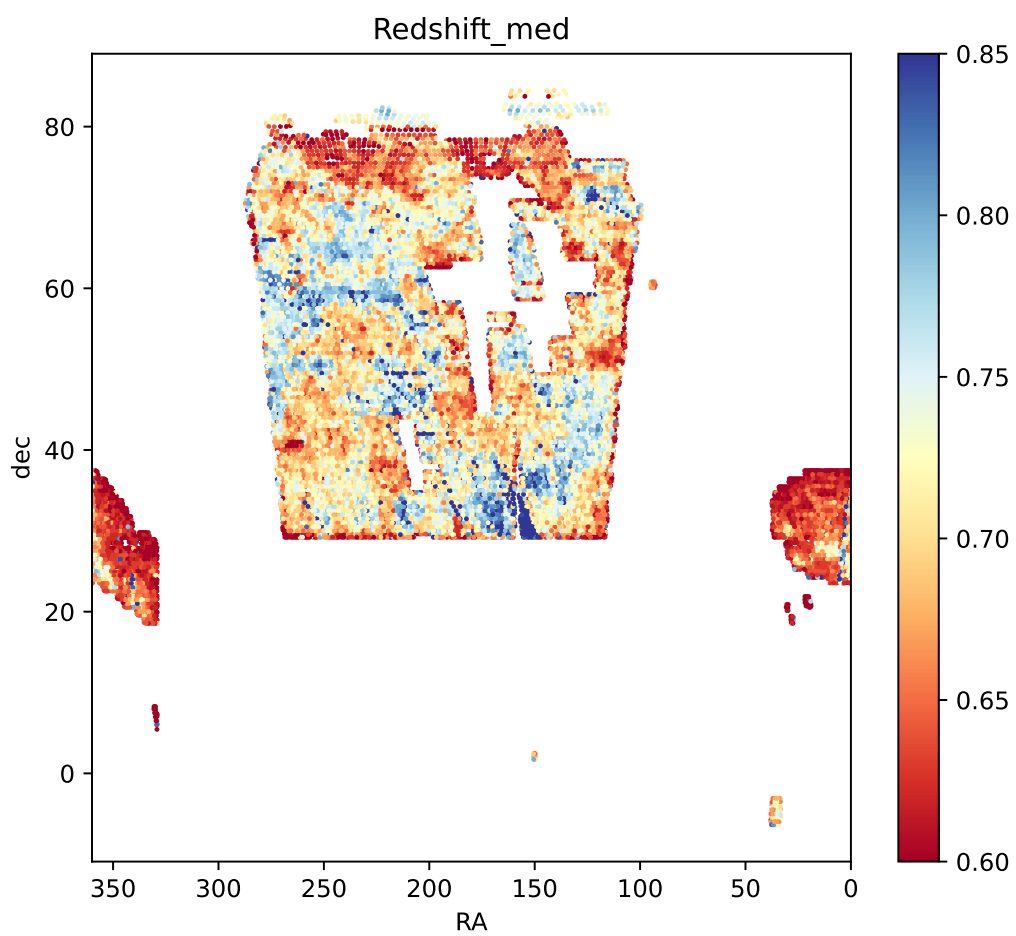


Figure 4.8: Same plot as Fig. 4.2, but for the median photometric redshift measurements per tile.

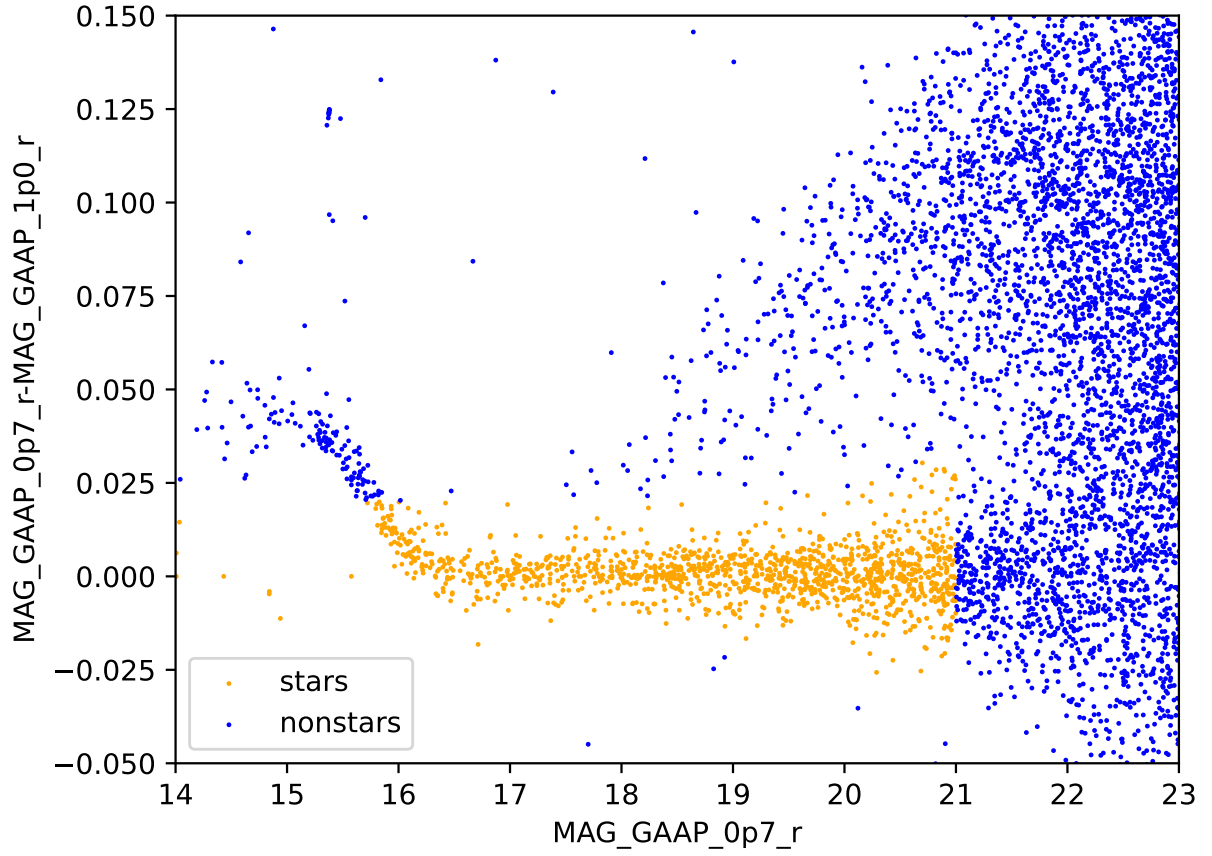


Figure 4.9: Plot of the difference of r -band magnitudes measured with two different minimum aperture sizes. Objects with little change (relative to measurement errors) are identified in orange as stars, and objects with significant changes are classified in blue as nonstars.

The code for these plots was handed off to the UNIONS collaboration for use in future versions of the photometric redshift data. Since the photometric redshift data is still in development at the time of writing, it will not be used in the analysis of Chapter 5. Instead, we will use the redshift distribution built from an overlapping region of UNIONS and CFHTLenS described in the previous section.

4.3 SDSS BOSS

For the foreground voids considered in Chapter 5, we use the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al., 2012) void catalogue of Woodfinden et al. (2022). The final BOSS survey is the 12th Data Release of the Sloan Digital Sky Survey (SDSS-III; Eisenstein et al., 2011), a spectroscopic survey targeting LRGs with two different targeting algorithms, LOWZ and CMASS (Reid et al., 2015). LOWZ targeted LRGs in the nearby universe ($0.2 < z < 0.43$) to produce a volume-limited survey. CMASS targeted LRGs just beyond the LOWZ sample ($0.43 < z < 0.75$) with the goal of targeting galaxies with constant stellar mass over the redshift range. Both surveys used the 2.5m Sloan telescope at the Apache Point Observatory, and covered roughly the same area, with LOWZ being smaller than CMASS. The total survey area is around $10,000 \text{ deg}^2$. To increase the statistical power of our measurement, we can combine voids found in both the LOWZ and CMASS LRG catalogues. This is reasonable because the galaxy clustering amplitudes of the two LRG samples are within 20 per cent of each other (Ross et al., 2016), meaning that the tracer populations used to identify the voids will exhibit similar properties, including similar galaxy bias values.

4.3.1 Void Finding

We will now summarize the steps of the void finding implementation. For more specific details, we refer readers to Woodfinden et al. (2022) for the implementation that was applied to our data set, and the references therein for more details on the void finding code Revolver (Nadathur et al., 2019a).

Reconstruction

Individual voids can be highly non-spherical, but stacked voids under the assumption of cosmological isotropy can be used as a standard spherical shape in cosmology, and are a candidate

for the Alcock-Paczyński (AP) test. Since the line-of-sight distance depends on the Hubble expansion, while the angular separation depends on the angular diameter distance, both distances have different dependencies on the underlying cosmology. If we have an a priori expectation of the shape, then we can constrain cosmology by varying the parameters until the measurement returns the expected shape. However, RSDs caused primarily by the peculiar velocities of tracer galaxies along the line of sight can artificially shift redshift estimations when compared to a tracer’s real-space position. Typically, void finding is done in redshift space, where observations are made, resulting in an artificial elongation along the line of sight, even for spectroscopic tracers where redshifts are well-estimated. To correct for this effect, the **RevoLver** void finder as described in [Nadathur et al. \(2019a\)](#) first attempts to reconstruct the real-space positions of tracers before running the void finder. RSDs are removed by solving the Zel’dovich equation in redshift space for the displacement field Ψ . Following the discussion of the Zel’dovich approximation in Sect. 1.4, the location of a particle at time t depends on its initial Lagrangian position $\mathbf{q}$ and the displacement field created by the matter field:

$$\mathbf{x} = \mathbf{q} + \Psi(\mathbf{q}, t) . \quad (4.3)$$

To first order, the displacement field and the matter field are related by

$$\nabla \cdot \Psi(\mathbf{q}, t) = -\delta_m(\mathbf{x}, t) . \quad (4.4)$$

To account for the fact that we are considering a set of linearly biased tracers of the matter distribution, and that these tracers are subject to an additional component of the displacement field caused by RSD, this equation is then modified to include an additional RSD term along the line of sight ([Nusser & Davis, 1994](#); [White, 2014](#); [Nadathur et al., 2019b](#)):

$$\nabla \cdot \Psi + \frac{f}{b} \nabla \cdot (\Psi \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}} = -\frac{\delta_g}{b} . \quad (4.5)$$

This then gives the Zel’dovich equation in redshift space. To approximately remove the RSD effect, this equation is solved for Ψ , and galaxies are then shifted to their real-space position using only the component of Ψ that contributes directly to RSD, $-\Psi_{\text{RSD}}$, defined as

$$\Psi_{\text{RSD}} = -f (\Psi \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}} . \quad (4.6)$$

This procedure introduces a dependence on $\beta \equiv f/b$ for our void catalogues, where f is the linear growth rate factor defined in Eq. 1.21, and b is the linear galaxy bias defined in Eq. 1.29. For the data set provided, β was set to 0.37, chosen to be the closest match to the fiducial cosmology of the PATCHY mocks in Sect. 4.3.1.

Void Finding Procedure

Once the set of tracer galaxies has been reconstructed to their real-space positions, they are passed to a modified ZOBOV void-finder (Neyrinck, 2008) provided by Revolver. Details of the ZOBOV void finding algorithm are detailed above in Sect. 3.2. To modify ZOBOV for use outside of simulated data, a set of barrier particles is added to the galaxy field, and any Voronoi cell or zone that touches a barrier particle is removed from consideration. These create a selection bias towards smaller voids near the mask edges, necessitating careful consideration when constructing catalogues of randomly positioned lenses for the void lensing measurement as described in Sect. 5.2.2. Voids are then identified as the full extent of a watershed basin around a local minima in the Voronoi tessellation. These basins are usually not spherical, so the definition of the void centre is not unique. As mentioned in Sect. 3.2, Revolver uses the circumcentre definition of the void centre, and the void radius is given as the radius of the sphere of equivalent volume given by Eq. 3.1.

Application to the BOSS Void Catalogues

As discussed in Woodfinden et al. (2022), for LOWZ and CMASS, the reconstruction step was applied to the combined sample of Alam et al. (2017), but the ZOBOV step was applied to each catalogue separately, due to differences in the survey masks. This was achieved by first adding a layer of barrier particles at $z=0.43$, the redshift boundary between the two surveys. This means that no voids are allowed to span between the LOWZ and CMASS surveys. The resulting LOWZ and CMASS void catalogues contain 2147 and 4781 voids in the North Galactic Cap (NGC), respectively. This total sample of voids will be used to measure the void-galaxy cross-correlation as discussed in Sect. 5.2.4.

It should be noted that not all of the CMASS LRGs were used in this void catalogue. Due to a partial overlap of the eBOSS+CMASS void cosmology analysis of Nadathur et al. (2020), LRGs at $z > 0.63$ were excluded from the void finding step as outlined in Woodfinden et al. (2022). Further, to avoid issues with edge effects in the void finding step, only voids with centres at $z < 0.6$ are included in the data set.

For void lensing, we are primarily interested in the subset of voids that have significant overlap with the UNIONS footprint, as these voids will be guaranteed to have source galaxies for the weak lensing calculation. In order to perform this cut, we use the HEALPIX (Górski et al., 2005) mask of the UNIONS survey with $N_{\text{side}} = 1024$, and we build a HEALPIX disk around each void centre with radius equal to the radius of the void.

Comparing the disk with the mask, we include the void in our lens sample if there are any pixels that overlap. Given that we are binning out to three times the void radius in our measurement, this

Table 4.1: Total number, average unweighted void radius, and average redshift of the 5 void catalogues used in the lensing analysis of Chapter 5, as well as for the void sample used in the full void-galaxy cross-correlation measurements.

Catalogue	N_{Void}	$\bar{r}_v [h^{-1} \text{Mpc}]$	$\bar{z}_v$
Full	2975	40.3	0.467
LOWZ	944	40.3	0.331
CMASS	2031	40.3	0.530
Small	2167	33.7	0.466
Large	808	57.9	0.469
Void-Galaxy	6928	39.5	0.468

condition means that we exclude most of the voids that would contribute at most a small number of lens-source pairs. After this cut, we are left with 944 and 2031 voids for LOWZ and CMASS void lensing, respectively. The radius and redshift distribution are shown in Figs. 4.10 and 4.11, respectively. We will measure the void lensing signal coming from each catalogue individually, as well as the combination of the two, which we will label Full.

We also construct two additional catalogues by splitting the Full catalogue into large and small voids. In order to split the lensing information evenly, this split is done using the average void radius, but rather than a simple average, it is weighted by the lensing weights, which will be described below in Sect. 5.2.1, summed over all lens-source pairs out to $3R_v$. The weighted average void radius used to perform this split is $48.5h^{-1} \text{Mpc}$. We note that this weighting upweights larger voids as they overlap more source galaxies than smaller voids. These void size catalogues will allow us to examine the void size evolution in the matter profiles. All five void catalogues and their statistics, as well as the catalogue used in the void-galaxy cross-correlation measurement, are compiled in Table 4.1.

PATCHY Mocks

In addition to the void catalogue, we also use the same set of voids found from 1000 mock galaxy catalogues as used in Woodfinden et al. (2022). These mocks were generated using the PATCHY algorithm (Kitaura et al., 2016b). Given a reference mock found from a halo occupation distribution (HOD) model applied to an N-body simulation, the PATCHY algorithm produces multiple mocks using augmented Lagrangian perturbation theory to model structure formation, and a bias model with deterministic and stochastic elements conditioned on reproducing the two- and three-point clustering functions. The end result is a mock galaxy catalogue that reproduces the monopole and quadrupole galaxy clustering signal of LOWZ and CMASS, as well as the

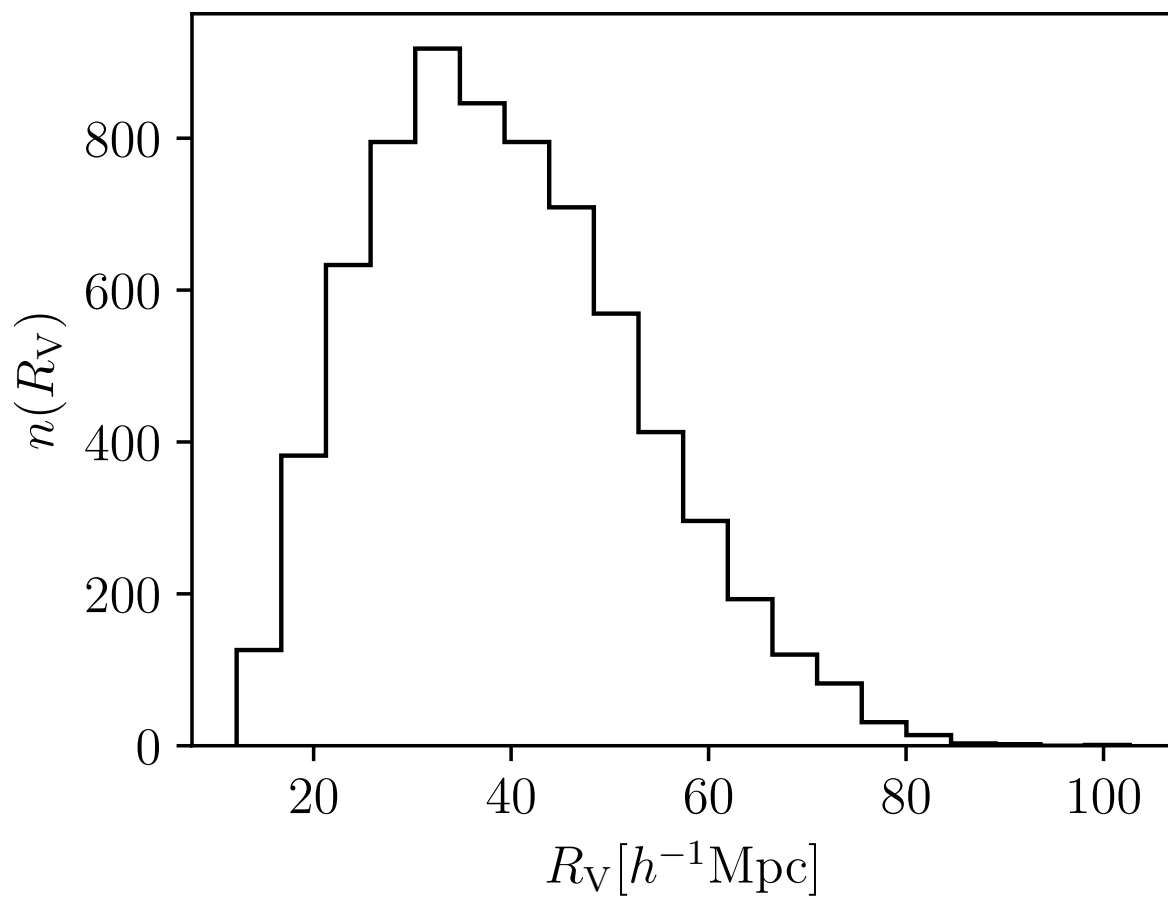


Figure 4.10: Void size distribution for the BOSS LOWZ+CMASS void catalogue from [Woodfinden et al. \(2022\)](#).

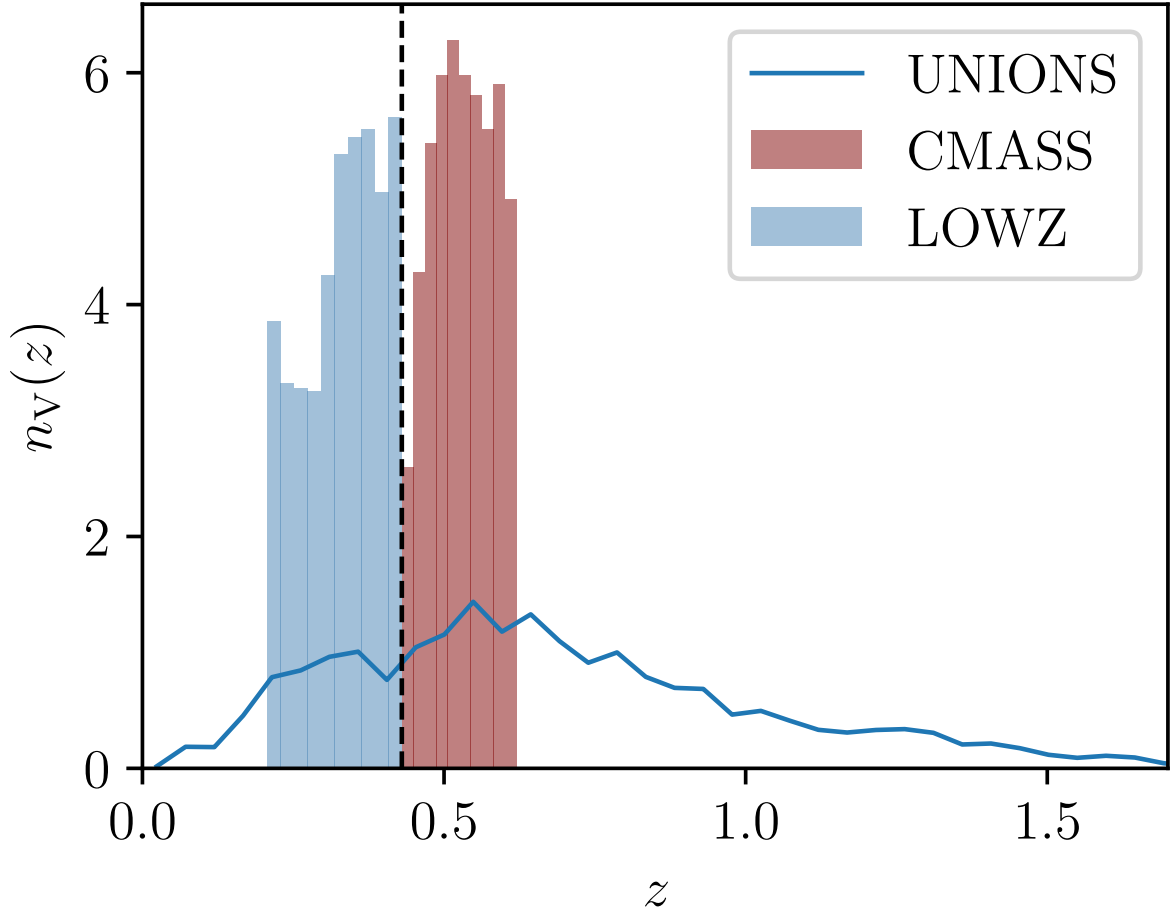


Figure 4.11: Void redshift distribution for the BOSS LOWZ and CMASS void catalogues from [Woodfinden et al. \(2022\)](#) with the UNIONS source redshift distribution. The UNIONS source distribution plotted in this figure is one of the blinded distributions. This is done in order to preserve the blinding of the cosmological analysis. It is close, but not identical, to the unblinded redshift distribution which is used for the analysis in this paper.

angular footprint, masking effects, and redshift distribution of the voids detected in the real data. We use the cosmology values in the PATCHY mocks as the fiducial cosmology for Chapter 5 for consistency.

The goal of void catalogues found from these mocks is to construct sets of random lenses that will be used to measure and subtract out lingering shear systematics in the UNIONS shape catalogue, as described in Mandelbaum et al. (2005). Further, as we will later choose to use an analytic model of our covariance, it was demonstrated in Singh et al. (2017) that subtracting the signal around random lenses simplifies the covariance model. Lastly, we use the covariance across these random lens catalogues to validate some of the simplifications outlined in Sect. 5.2.3 for the analytic covariance.

4.4 Euclid Flagship Simulation

In Chapter 6, we use the mock galaxy catalogue of the *Euclid* Flagship simulation (version 1.8.4). This version of the catalogue was chosen to match the version of Flagship that was used in the construction of the void catalogue of Hamaus et al. (2022), and is an older version than the Flagship 2 simulation presented in Euclid Collaboration: Castander et al. (2025). This choice was made because we will be using the void catalogue of Hamaus et al. (2022) to forecast the void lensing signal. Version 1.8.4 uses an N -body simulation inside a periodic box of side length $3780 h^{-1}$ Mpc with $12\,600^3$ particles (Potter et al., 2017). Dark matter haloes were then found using the ROCKSTAR halo finder (Behroozi et al., 2013), and subsequently populated with galaxies using a HOD. This distribution model (Crocce et al., 2015; Carretero et al., 2015) was calibrated to reproduce local observational constraints, including the luminosity function for the faintest galaxies (Blanton et al., 2003, 2005), and galaxy clustering measurements as a function of luminosity and colour (Zehavi et al., 2011). The cosmology of the simulation matches the 2016 *Planck* results, with $\Omega_m = 0.319$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.681$, $\sigma_8 = 0.83$, $n_s = 0.96$, and $h = 0.67$ (Planck Collaboration et al., 2016). The simulation box was then converted to a light cone, which covers 5157 deg^2 on the sky, or one octant. This light cone then extends to a redshift of $z = 2.3$ (compared to $z = 3$ with Flagship 2). From this, we construct two sub-catalogues corresponding to the two different surveys conducted by the two different instruments onboard *Euclid*, the VISible imager (Euclid Collaboration: Cropper et al., 2025, VIS), which will be used to provide the photometric catalogue for background galaxy shapes, and the Near-Infrared Spectrograph and Photometer (Euclid Collaboration: Jahnke et al., 2025, NISP), which is used for the spectroscopic tracer catalogue that was used for the void finding in Hamaus et al. (2022).

4.4.1 Euclid Flagship NISP Void Catalogue

The void catalogue used in this analysis comes from [Hamaus et al. \(2022\)](#), who used this catalogue for the purposes of forecasting the cosmological signal from jointly constraining the RSD signal and the AP test. This was chosen so that we can forecast the expected signal for the first data release of *Euclid*. The second and third data releases will offer additional opportunities, such as other overlapping spectroscopic catalogues like DESI or the aforementioned BOSS void catalogue, but these surveys will have little expected overlap with the first data release. To generate the NISP-like spectroscopic galaxy catalogue used to identify the voids, the authors applied a cut on the galaxy mock catalogue of Flagship v1.8.4 on the $H\alpha$ flux, $f_{H\alpha} > 2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$. This is expected to match the 3.5σ detection threshold for *Euclid* NISP objects between redshifts $0.9 < z < 1.8$ ([Laureijs et al., 2011](#); [Euclid Collaboration: Mellier et al., 2025](#)). To match the completeness of the expected survey, the authors then randomly sample the catalogue to 60% of objects, and apply a Gaussian redshift error with a RMS of $\sigma_z = 0.001$ to simulate the expected spectroscopic performance, resulting in a catalogue of roughly 6.5×10^6 spectroscopic galaxies across the octant on the sky.

The void finder algorithm used to build the catalogue is the public VIDE library ([Sutter et al., 2015](#)), which is based on a modified version of the ZOBOV algorithm ([Neyrinck, 2008](#)), and details of the ZOBOV algorithm were discussed in Chapter 3. This is a different void finder than the Revolver void finder that was used for the BOSS voids, and the differences are discussed in Sect. 3.2. Void centres of Flagship voids are chosen to be the volume-weighted barycentre, defined to be the averaged centre of the galaxies that constitute the void region. Since the outer walls will have many more galaxies than the interior, this definition identifies the centre with respect to the outer walls of the void over identifying the minimum density location of the void as the centre, which is done in Revolver and can exist anywhere inside of the void walls. This means that in a stack of individually non-spherical voids, a barycentre void profile will emphasize the overdense wall and potentially average out the interior, making it shallower. This definition is beneficial for lensing as the walls will capture more background sources due to having greater radial bin areas, while the innermost regions will be the least constrained. The radius of the Flagship voids are then defined to be the radius of a sphere with volume equivalent to the total volume of the void as shown in Eqn. 3.1. This procedure yields 58 601 voids when applied to the light cone of Flagship v1.8.4. Additionally, we follow [Hamaus et al. \(2022\)](#) and apply a purity cut based on the void radius. The reason for this cut is that a void finder may identify a Poisson fluctuation in the galaxy sample as a small void region that does not physically correspond to an underdensity in the matter distribution. Given that the galaxy number density of the Flagship NISP catalogue changes with redshift (see Fig. 1 of [Hamaus et al., 2022](#)), this radius cut must

scale with the galaxy number density $n_g(z_v)$ as

$$R > N_s \left[\frac{4\pi}{3} n_g(z_v) \right]^{-1/3} \quad (4.7)$$

where N_s is defined in units of the average galaxy separation to set where the minimum radius occurs. We use the same value for N_s as [Hamaus et al. \(2022\)](#), who used $N_s = 3$. This results in the catalogue used for this chapter containing 44 469 voids.

From this catalogue, we separate the voids in two different ways. The first method is used to maximize the strength of the void lensing detection, where we construct a set of tomographic bins in redshift. To this end, we construct a set of four equinumber bins, matching the binning of [Hamaus et al. \(2022\)](#). The second method is used to determine how well we can constrain the relationship of void lensing profiles with void size. For this, we construct three equinumber bins in void radius. The radius bins can be seen in [Fig. 4.12](#), which shows the total size distribution for the whole catalogue, with black dashed lines to show the bin edges for the radius-split catalogues. The tomographic void redshift bins can be seen below with the VIS source redshift distributions in [Fig. 4.13](#). Statistics for all seven catalogues are shown in [Table 4.2](#).

4.4.2 Euclid Flagship VIS Sources

We construct the catalogue of galaxy shapes from the Flagship mock galaxy sample using simulated VIS magnitudes and lensing information. As outlined in [Euclid Collaboration: Castander et al. \(2025\)](#), galaxy shear components γ_1 and γ_2 are simulated by first constructing the dark matter lensing mass maps by binning the dark matter light cone in concentric spherical shells. Each shell is then projected into 2D, where the combination of shells can be used to compute the galaxy shear components as described by [Fosalba et al. \(2008\)](#). This produces a set of convergence maps, which are then converted into shears of individual galaxies by matching the pixel that a galaxy is located in to the same pixel in a corresponding convergence map. The accuracy of this methodology is therefore limited by the size of the pixel, which for the Flagship simulation corresponds to a HEALPix $N_{\text{side}} = 8192$, or about 0.5 scales ([Euclid Collaboration: Castander et al., 2025](#)). This procedure produces the true distorted galaxy ellipticity produced by the underlying matter distribution without any additional observational systematics like shape noise, which will be used to measure the underlying true lensing measurement of the Flagship simulation.

In order to build the catalogue of expected *Euclid* galaxy shapes, we take the mock galaxy catalogue and consider all galaxies with VIS magnitude below 24.5. This magnitude limit corresponds to the expected depth of Data Release 3 (DR3), or the full 15 000 deg² survey. This

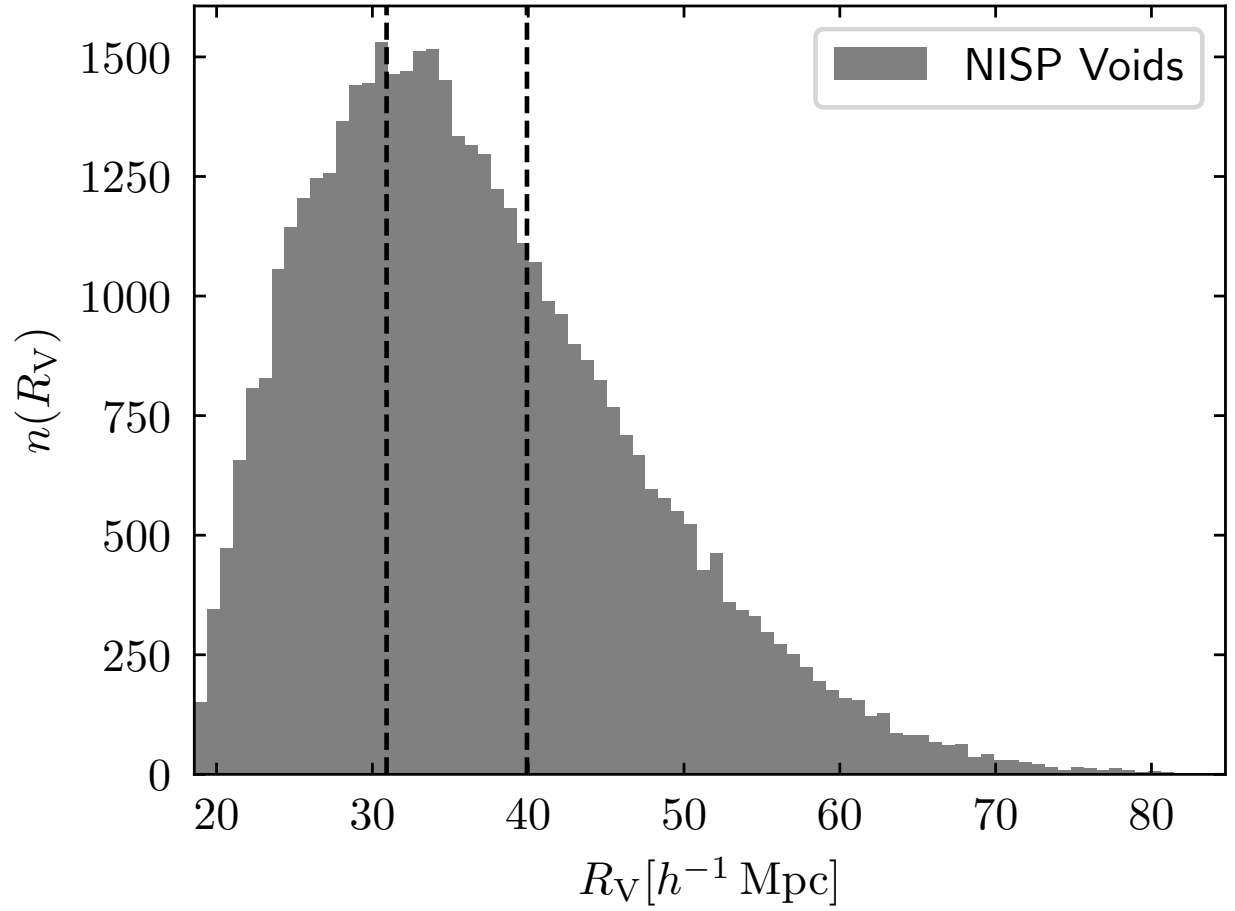


Figure 4.12: Void size distribution for the NISP-like void catalogue of [Hamaus et al. \(2022\)](#) used in Chapter 6. The vertical dashed lines indicate cuts made for the void size catalogues in equal-number bins.

produces a catalogue with an area number density of $n_g = 28.8 \text{ arcmin}^{-2}$ with position, true redshift, and lensing shear data for each galaxy across 5157 deg^2 of area. Individual simulated photometric redshifts were not included in Flagship v1.8.4 beyond a small patch measured by [Euclid Collaboration: Pocino et al. \(2021\)](#). Instead, we generate our own individual photo- z measurements such that they match the expected distributions from that work.

For these source redshift distributions, we rely on the measured photo- z distributions from [Euclid Collaboration: Pocino et al. \(2021\)](#). They measured realistic photo- z distributions in a 402 deg^2 patch of the simulation using a Directional Neighbourhood Fitting machine learning algorithm ([De Vicente et al., 2016](#)). They performed the measurement on galaxies binned into 13 bins of equal width in true redshift, releasing the realistic individual photo- z measurements for each galaxy in the small patch internally to the collaboration. It was found in [Hoyle et al. \(2018\)](#) that distributions in photometric redshifts found by the Directional Neighbourhood Fitting algorithm did not have an impact on cosmological results compared to those found using the expected BPZ code in UNIONS. Therefore, it is expected that the differences in photometric redshift codes will not have a significant impact on the void lensing results. To apply these individual measurements our simulated VIS catalogue, we first need to build the corresponding distributions of the scatter in the photo- z from the small patch results. We do this within each of the 13 bins by subtracting each small-patch true redshift from its corresponding photo- z measurement. This then gives a mapping between a bin in true redshift to a distribution of expected photo- z scatter. We then take the true redshifts from the total VIS catalogue, which are available within Flagship v.1.8.4, and apply the same binning scheme in true redshift, allowing us to map each VIS galaxy to a corresponding photo- z scatter distribution. We then sample the corresponding distribution and add the sample to the true redshift value, resulting in an estimated individual photo- z measurement for that galaxy. The combination of individual photo- z measurements for a given true redshift bin then give a realistic photo- z distribution of a given tomographic source bin.

Given that we are only interested in VIS galaxies that exist comfortably behind our voids, we compare these source tomographic bins to the redshift distributions of our NISP void sample and exclude the first seven photo- z bins from our analysis, leaving us with six tomographic source bins to pair with our four tomographic void bins. Since this excludes the peak of the VIS source distribution, this also cuts the effective galaxy number density to 4.2 arcmin^{-2} . These distributions can be seen in [Fig. 4.13](#), where grey denotes the void redshift distribution, with void bin edges denoted by the black dashed lines, and the photo- z distributions denoted by the different coloured histograms. From this plot, we can then choose which combinations of void and source bins we want to take measurements from. For the case of the size-cut void samples, we opt to combine all six of the shown photo- z distributions together into a single distribution, as we do not define a clear separation in redshift between voids of various sizes. This comes with

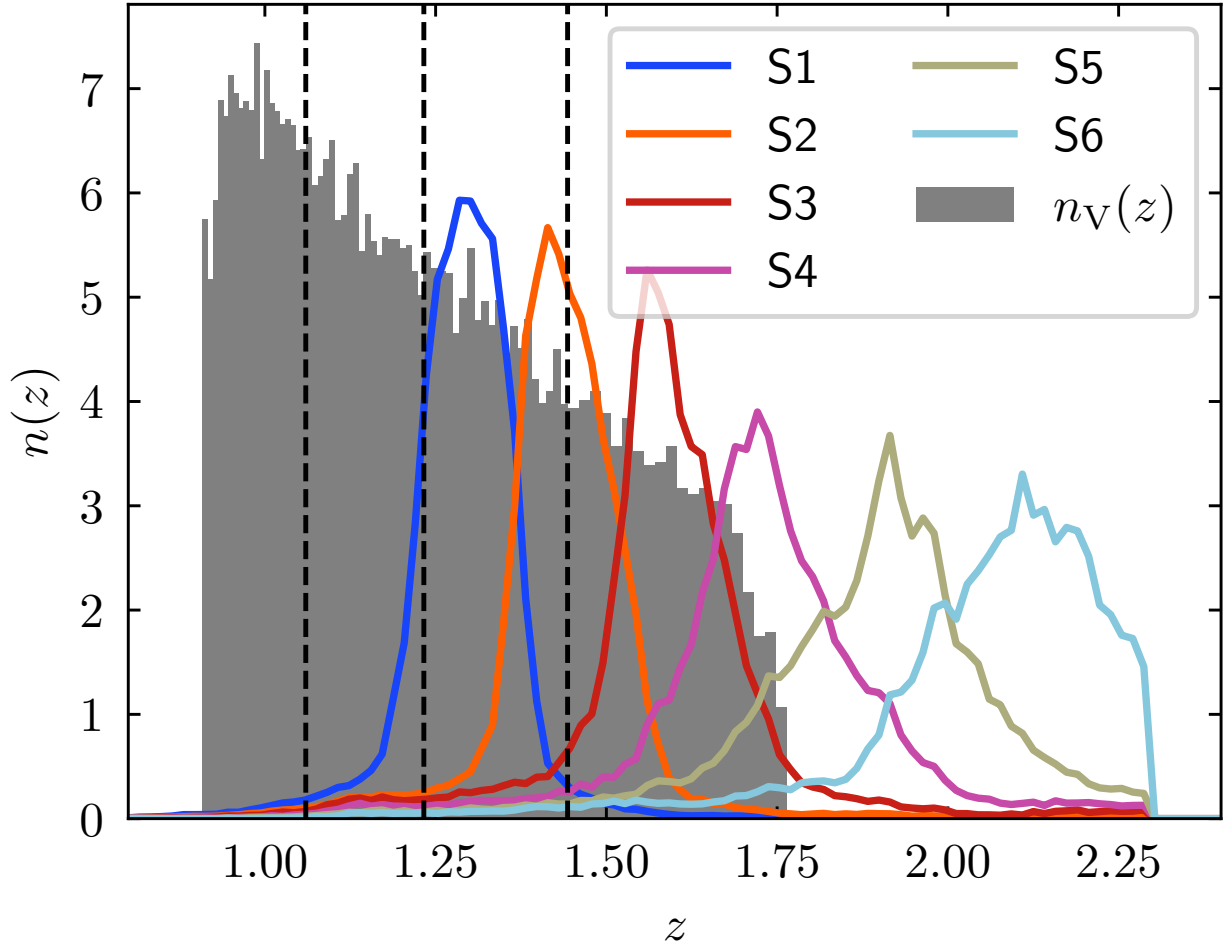


Figure 4.13: Redshift distributions for the NISP-like void catalogue (in grey) and the VIS-like source catalogue (in colour). The vertical dashed lines indicate the boundaries for the equal-number void tomographic bins, and the different coloured bins indicate the tomographic bins for sources.

Table 4.2: Void catalogues and statistics for the seven catalogues considered in Chapter 6. The last column indicates which source bins are used in the lensing analysis for the tomographic catalogues, as well as the void size range for the void size catalogues.

Catalogue	$\overline{R_v} [h^{-1} \text{ Mpc}]$	$\bar{z}$	N_v	Sources/Size
L1	35	0.99	11117	S1-S6
L2	39	1.1	11117	S2-S6
L3	43	1.3	11117	S3-S6
L4	48	1.6	11118	S5-S6
Small	26	1.0	14823	$19 < R_v < 31$
Medium	35	1.1	14823	$31 < R_v < 40$
Large	48	1.1	14823	$40 < R_v < 85$

the unavoidable contaminant that some galaxies will exist in front of our size-cut voids, which are not lensed and therefore reduce the signal.

Chapter 5

Measuring Lensing from BOSS Voids with UNIONS Sources

5.1 Introduction

In this chapter, we are interested in developing the methodology for and performing a void lensing measurement using currently available data. Given the significant overlap of the foreground spectroscopic BOSS survey and background imaging UNIONS survey, we want to measure the lensing signal around three-dimensional voids identified in spectroscopic tracers. By using the spectroscopic void catalogue of [Woodfinden et al. \(2022\)](#), who built the catalogue to measure the void-galaxy cross-correlations, we are then able to test the linear bias relationship between galaxies located inside and around voids to the underlying matter distribution. This is possible by comparing the void lensing profiles we obtain to the void-galaxy cross-correlation measurements, which should have different amplitudes represented by a single multiplicative bias factor. We additionally aim to test the findings of [Pollina et al., 2017](#) who demonstrated through comparisons in hydrodynamical simulations of void tracer profiles and void matter profiles that a linear biasing model is sufficient to relate the void-galaxy profile to the void-matter profile, but that the bias changes with void radius. They found that the void-galaxy bias (what we will call b_{vg}) is larger than the large-scale linear galaxy bias (b_g) for small voids, then asymptotically approaches b_g with increasing void size. This effect was tested using DES Y1 data in [Fang et al. \(2019\)](#), who found consistency with the large-scale bias irrespective of void size, but with large enough scatter to mask any trend. In contrast, [Nadathur & Percival \(2018\)](#), using N -body simulations with HOD modelling, argue from a stochastic basis that clustering of a tracer sample around voids identified within that tracer sample produce a selection effect. That is to say that measuring

the clustering of tracers around regions pre-selected to be underdense in the tracers necessarily produces a clustering measurement that prefers regions in density that are scattered low. This then biases the stacked void-galaxy profile such that the galaxy bias depends on the distance from the centre of the void, with a fractional difference up to 25 per cent. One of the goals of this chapter is therefore to test the linear bias model using voids identified in spectroscopic data by comparing our void lensing measurement to the void-galaxy correlation measurement from [Woodfinden et al. \(2022\)](#).

The aim of this chapter is to measure and determine the significance of the void lensing signal of 3D spectroscopic voids coming from current generation surveys with the intention to demonstrate the potential for next-generation surveys such as *Euclid* ([Euclid Collaboration et al., 2025](#)) and Roman ([Eifler et al., 2021](#)), both of which promise larger and higher-quality data sets than currently available. We are additionally interested in how well we can constrain the large-scale structure information contained within this signal. As such, the chapter is organized as follows. In Section 5.2, we detail the methodology for obtaining the lensing measurement for these voids, using the BOSS and UNIONS data sets introduced in Chapter 4. This includes methodology for obtaining the lensing signal around random voids and the novel derivation of an analytic covariance model. Additionally, we describe both models used in this analysis: an analytic model used to assess the significance of the detection, and a void-galaxy measurement to constrain linear void-galaxy bias. In Section 5.3, we present our measurements. In Section 5.4, we compare and discuss our measurements in comparison with other works and discuss future work in light of upcoming galaxy imaging and clustering surveys. The fiducial cosmology we assume for this chapter is a flat Λ CDM cosmology with $\Omega_m = 0.307$, $\Omega_b = 0.0482$, $h = 0.6777$, $\sigma_8 = 0.8225$, and $n_s = 0.96$, and all distances are in comoving units.

5.2 Methodology

5.2.1 Weak Lensing Formalism

In a weak lensing framework, the correlated shapes of background galaxies are directly related to the excess surface mass density (ESMD) $\Delta\Sigma$. We can recover the ESMD profile from source shapes by taking a weighted average over the shapes in each radial bin, R , around a void:

$$\Delta\Sigma(R) = \frac{\langle\gamma_t\rangle(R)}{\langle\Sigma_c^{-1}\rangle}, \quad (5.1)$$

where $\langle \Sigma_c^{-1} \rangle$ is the inverse of the critical density averaged over the source distribution

$$\langle \Sigma_c^{-1} \rangle (z_L) = \frac{4\pi G \chi(z_L) (1 + z_L)}{c^2} \int_{z_L}^{\infty} dz_s n(z_s) \frac{\chi(z_L, z_s)}{\chi(z_s)}, \quad (5.2)$$

where $n(z_s)$ is the redshift distribution of the sources, $\chi(z_L)$ is the comoving distance to the lens at redshift z_L , $\chi(z_s)$ is the distance to the source at redshift z_s , $\chi(z_L, z_s)$ is the distance between the lens and source. This equation is the comoving counterpart to Eqn. 2.18. All distances are measured in comoving units, chosen to facilitate easier modelling of the covariance, as described later. In the version of the UNIONS data used here, photometric redshifts are not available, so we integrate over the entire redshift distribution.

In the case of many weak lensing applications, such as galaxy-galaxy lensing, a single lens does not yield a significant detection. This is also true of cosmic voids (Amendola et al., 1999) and so, as suggested in Krause et al. (2013) it is necessary to stack the voids by combining the source shapes behind all of the voids into a single set of radial bins and making a single measurement, effectively treating all of the sources as though they were lensed by a single, average void.

However, when stacking voids, an additional issue arises that is not typically present in standard galaxy-galaxy or cluster lensing analyses. Voids have a wide range of radii, and so if we were to apply a universal set of radial bins specified in Mpc, for example, to each void, then the resulting average void density profile would be washed out, as the overdense walls of the smaller voids would be entirely contained within of the underdensities of the larger voids. Instead, because the matter distributions inside of voids have been found to be self-similar in simulations (Hamaus et al., 2014b; Nadathur et al., 2015), we can therefore scale the radial bins by the radius of each void. Doing so allows the distinct features of the profile to occur in roughly the same bins. The resulting measurement then becomes the weighted average of the shapes in these scaled bins:

$$\Delta \Sigma (R/R_V) = \frac{\sum_L \sum_S w_{LS} \gamma_{t,LS} (R/R_V) \langle \Sigma_c^{-1} \rangle^{-1}}{\sum_L \sum_S w_{LS}}, \quad (5.3)$$

where the sum over S is the sum over all sources in a single radial bin across all lenses, and w_{LS} is the weight assigned to each lens-source pair. The optimal weight for a shape-noise dominated regime is given by (Sheldon et al., 2004; Shirasaki & Takada, 2018)

$$w_{LS} = \frac{\langle \Sigma_c^{-1} \rangle^2}{2\sigma_\epsilon^2 + \sigma_{\epsilon_1}^2 + \sigma_{\epsilon_2}^2} = \langle \Sigma_c^{-1} \rangle^2 w_s, \quad (5.4)$$

using the shape weights w_s as in Eq. (4.1), with an additional factor of $\langle \Sigma_c^{-1} \rangle^2$. This additional factor downweights lens-source pairs that are physically close to each other, where source galaxy shapes will not experience a strong lensing effect (Shirasaki & Takada, 2018).

5.2.2 Random Lensing Voids from PATCHY Mocks

It is common practice to subtract the lensing signal around randomly distributed points in the lens plane. This allows for the estimation and subsequent subtraction of lingering shear systematics that can spuriously correlate the shapes of the source galaxies (Mandelbaum et al., 2005). However, due to the fact that our radial bins scale with the physical size of the void, we must assign a radius and redshift to our random points. Because voids that touch the boundary of the BOSS survey are removed, we must also consider a selection bias towards smaller voids near the BOSS survey boundaries. This bias would not be present if we built a catalogue from a random distribution of points across the survey footprint and assigned each point a radius and redshift by sampling the corresponding data distributions. Furthermore, building a void catalogue by running *Revolver* on a random set of tracer points would be insufficient, as this catalogue would fail to reproduce the void size distribution in the data.

Fortunately, void catalogues produced by running the same *Revolver* void finder on the PATCHY mocks will reproduce these edge effects, so they can be used to generate a set of realistic randomly distributed lenses. This produces a set of 1000 mock void catalogues. However, if we were to sample random voids entirely within a single mock catalogue, since each mock attempts to reproduce the clustering signal in the galaxy surveys, it would necessarily approximate the clustering of the voids present in the data. This clustering signal needs to be removed so that the mocks can mirror a uniform random distribution in position while maintaining the selection biases at the boundaries.

To remove this clustering signal and build our random lens catalogues, we randomly pick out individual voids from across all 1000 mocks until we construct a catalogue that matches the number of voids in each data set. We then do this 30 times to construct 30 separate random catalogues per data set. Then, we measure the stacked $\Delta\Sigma$ signal from each of the catalogues. This results in 30 independent measurements of systematic shape correlations in each radial bin. Then we take an average across all of the random measurements in each bin, as well as compute the 14×14 covariance matrix across the 14 radial bins using the 30 measurements. This covariance, which represents an empirical estimate of the covariance expected in the data, will be used to validate our covariance model in Sect. 5.2.3.

The measured $\Delta\Sigma$ profiles around randomly distributed lenses for all catalogues can be seen in Fig. 5.1. The error bars shown in this figure represent the 1σ scatter in each bin across all 30 catalogues, reduced by the factor $1/\sqrt{30}$ to represent the error in the mean. Since the lens centres for these data sets are distributed randomly, we expect that there should not be any coherent lensing signal present in these results. However, we do find a consistent trend across all but the smallest scales where the measured values are larger than zero, pointing towards low-level systematics that are nevertheless large enough (compared to the weak void measurement, roughly

10 to 20 per cent of the signal) that they will need to be subtracted off. These systematics can arise as a result of additive shear biases that are caused by, for example, residual PSF effects, bias due to noise in the original image when correcting for PSF, or a shear selection bias that prefers highly elongated galaxy images as a result of improved resolution metrics when convolved with a PSF (Mandelbaum et al., 2005). We do note however, that these systematics are of the order 4×10^{-5} when converted to tangential shear, and would represent a small contribution to more standard applications of lensing.

To test the effects of this methodology over a simple random sample, we measured the signal around a uniform random sample of voids, with void radii sampled from the distribution of void radii in the actual LOWZ voids. Additionally, to see if this signal is due to the bin scaling required by voids, we measured the signal around a uniform random sample of galaxies, chosen in the same way as the uniform random voids, just without specifying a radius. Both uniform samples were chosen with properties of the LOWZ catalogue as it has the largest amplitude random “signal”. We found that there still exists a non-zero signal, but at a roughly constant value of around $0.05 M_{\odot} h^{-1} \text{pc}^{-2}$, consistent between random voids and random galaxies, indicating the bin scaling does not have a strong effect on the random “signal”. Comparing this low and constant value from a uniform random sample to the non-constant and larger amplitude signal seen in LOWZ in Fig. 5.1, which comes from mock voids, we conclude that the selection effects at the boundary captured by the mocks has an effect on the measured systematics.

To obtain our final measurements, we subtract the randoms (shown in Fig. 5.1) from the data. We further add the covariance matrix obtained from the randomly distributed lenses, divided by 30 (representing the covariance in the mean), to the analytic covariance discussed in Sect. 5.2.3 to propagate these errors to our final measurement.

5.2.3 Analytic Covariance Modelling

Obtaining an accurate estimate for the void lensing covariance is essential because the signal itself is weak, as demonstrated by other analyses (e.g. Melchior et al., 2014; Fang et al., 2019), and also due to the nature that the scales considered are large. The total number of lenses is also relatively low, indicating that large-scale structure (LSS) effects are expected to have a stronger presence in the error budget of this form of lensing analysis. This differs from the more familiar case of galaxy-galaxy lensing, where the small scales probed by galaxy lenses, coupled with the large number statistics of lens galaxies, usually means that covariance contributions from LSS are subdominant to the purely diagonal shape noise component. On the other hand, the large scales of cosmic voids mean that the covariance is insensitive to non-linear scales, simplifying the modelling. To this end, we opt to analytically model the covariance for void lensing. This is

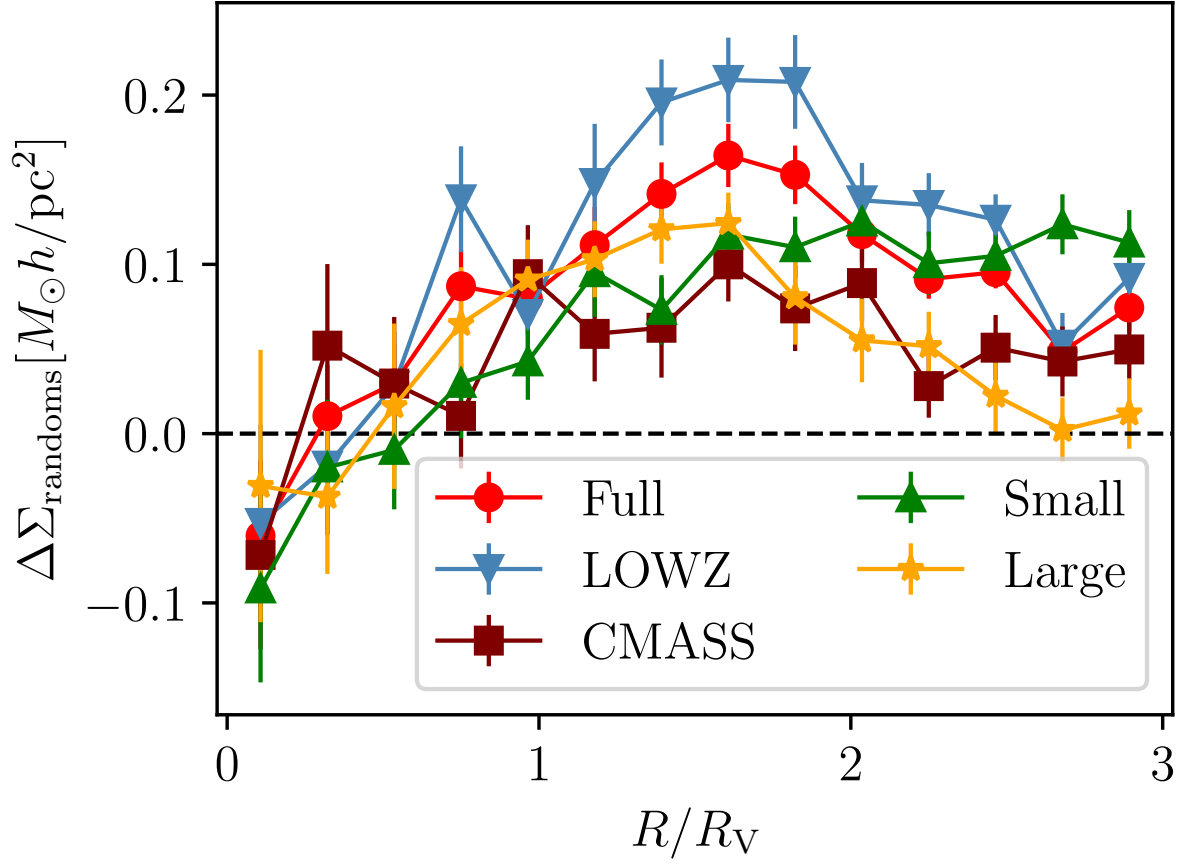


Figure 5.1: Results from measuring the $\Delta\Sigma$ profile around 30 sets of random points from mocks. Each data point is an average of the 30 random catalogues for each radial bin, and the error bars are the square root of the diagonal of the covariance matrix found from the randoms divided by the 30 catalogues that went into the measurement.

the first time an analytic covariance matrix has been used for a void lensing analysis with data. We start with the analytic covariance formula from [Krause et al. \(2013\)](#)

$$\text{Cov} [\gamma_t (\theta_1), \gamma_t (\theta_2)] = \frac{1}{4\pi f_{\text{sky}}} \int \frac{\ell d\ell}{2\pi} \overline{J_2 (\ell \theta_1)} \overline{J_2 (\ell \theta_2)} \left\{ \left[C_{\text{VV}} (\ell) + \frac{1}{n_{\text{V}}} \right] \left[C_{\text{KK}} (\ell) + \frac{\sigma_{\epsilon}^2}{n_{\text{g}}} \right] + [C_{\text{VK}} (\ell)]^2 \right\}, \quad (5.5)$$

where f_{sky} is the sky fraction of the survey, $\widehat{J}_2$ is the 2nd order Bessel function of the first kind averaged over each radial bin, n_{V} and n_{g} are the angular number density of voids and source galaxies, respectively. θ_1 and θ_2 are defined to be the angular extent of the radial bins, which we calculate by multiplying the corresponding R_1/R_{V} and R_2/R_{V} values by the angular radius of the void θ_{V} to convert to angular units. The terms C_{VV} , C_{KK} , and C_{VK} are the angular power spectra representing the clustering of voids, source galaxy shapes, and their cross-correlation, respectively. This covariance formula is similar to the Gaussian component of covariances used in standard lensing analyses, only using voids as a tracer population (e.g. [Jeong et al., 2009](#); [Singh et al., 2017](#); [Krause & Eifler, 2017](#)). Per [Singh et al. \(2017\)](#), this form of the covariance matrix only applies once the tangential shear around random points is subtracted from the overall signal. Otherwise, additional terms would need to be added in to account for the fact that we would be correlating a void field (which does not have a mean of zero) with a mean-zero shear field. Subtracting the random signal turns the correlation into a correlation between mean-zero quantities.

We demonstrate in Appendix [A.1](#) that terms involving C_{VK} and C_{VV} are subdominant to the $C_{\text{KK}}/n_{\text{V}}$ term across all scales, and can therefore be dropped. To measure the variance due to shape noise (which arises from the product of $1/n_{\text{V}}$ and $\sigma_{\epsilon}^2/n_{\text{g}}$), we compute the error in the weighted mean:

$$\sigma_{\text{SN}, \Delta\Sigma}^2 = \frac{\sum_{\text{L}} \sum_{\text{S}} w_{\text{LS}}^2 \langle \Sigma_{\text{c}}^{-1} \rangle^{-2} \sigma_{\epsilon}^2}{(\sum_{\text{L}} \sum_{\text{S}} w_{\text{LS}})^2}, \quad (5.6)$$

where the $\Delta\Sigma$ subscript indicates that we include a factor $\langle \Sigma_{\text{c}}^{-1} \rangle^{-2}$ to convert the shear to $\Delta\Sigma^2$ units, and $\sigma_{\epsilon} = 0.34$ as measured in the UNIONS survey. These assumptions yield the much simpler formula:

$$\text{Cov} [\gamma_t (\theta_1), \gamma_t (\theta_2)] = \frac{1}{4\pi f_{\text{sky}}} \int \frac{\ell d\ell}{2\pi} \overline{J_2 (\ell \theta_1)} \overline{J_2 (\ell \theta_2)} \left[\frac{C_{\text{KK}} (\ell)}{n_{\text{V}}} \right] + \delta_{\theta_1, \theta_2}^{\text{K}} \sigma_{\text{SN}}^2, \quad (5.7)$$

where δ^{K} is the Kronecker delta function, to represent that the measured shape noise only adds a diagonal contribution to the covariance. Using the Limber approximation ([Limber, 1953](#);

LoVerde & Afshordi, 2008)

$$C_{\kappa\kappa}(\ell) = \int d\chi \left[\frac{\bar{\rho}}{\chi} \int_{\chi}^{\infty} dz_s \frac{c}{H(z_s)} \frac{p(z_s)}{\Sigma_c(z_s, z)} \right]^2 P_{\text{mm}} \left(k = \frac{\ell + 1/2}{\chi} \right), \quad (5.8)$$

with $1/\Sigma_c$ is defined as in Eq. (5.2) without the integral over the source distribution. As mentioned earlier, because the scales we probe with voids are quite large, we only need to use a linear matter power spectrum model.

The remaining $C_{\kappa\kappa}/n_V$ term in the covariance formula physically represents covariance between radial bins due to lensing from intervening large-scale structure along the line of sight. It is then intuitive why this is the dominant term in the LSS component of the covariance: the combination of the sparseness of the LRG tracers resulting in very large voids, coupled with the low redshifts of the BOSS sample, creates bins that can extend upwards of eight degrees (using the average void radius) in radius on the sky. The chances of large-scale structure existing within these bins along the line of sight in a single pointing are then reasonably high. Furthermore, the relatively low number density of foreground voids results in higher shot noise compared to analogous applications of this covariance formula, like galaxy-galaxy lensing.

However, we note that Eq. (5.7) is the covariance of the lensing shear γ_t , and not the covariance of the excess surface density $\Delta\Sigma$. Following the examples of Shirasaki & Takada (2018) and Wu et al. (2019), who modified this covariance matrix for cluster lensing, we need to include an additional factor of $\Sigma_c^2(z_s, z_L)$ into Eq. (5.8). Since there is a remaining integral over source redshift present in Eq. (5.8), normally, this extra factor would need to be included in this integrand. However, this integral is not well-behaved for sources in front of or at the same redshift as the lenses, z_L . As there is no method to identify which sources these are without some estimate of individual redshifts, we are not computing the same measure of the excess surface density as Wu et al. (2019) or Shirasaki & Takada (2018), as their lensing estimator has Σ_c is an explicit function of the source redshift, while ours is expressed as an average over the source distribution, as seen in Eqs. (5.1) and (5.2). Since $\langle \Sigma_c^{-1} \rangle$ no longer depends on the source redshift, it can factor out of both integrals in Eq. (5.8). It can also factor out of the integral of Eq. (5.7), leaving

$$\text{Cov} [\Delta\Sigma(\theta_1, z_L) \Delta\Sigma(\theta_2, z_L)] = \frac{1}{4\pi f_{\text{sky}} \langle \Sigma_c^{-1} \rangle^2(z_L)} \int \frac{\ell d\ell}{2\pi} \widehat{J_2(\ell\theta_1)} \widehat{J_2(\ell\theta_2)} \left[\frac{C_{\kappa\kappa}(\ell)}{n_V} \right] + \delta_{\theta_1, \theta_2}^K \sigma_{\text{SN}, \Delta\Sigma}^2. \quad (5.9)$$

Finally, because our measurement uses radial bins that depend on each individual void's radius, there is an ambiguity in terms of what angular scales to consider for the Bessel window

functions, as each void effectively probes different scales of the LSS covariance term. This is in contrast to traditional weak lensing setups, where each lens uses the same set of bins, so θ_1 and θ_2 are consistent between all lenses. There are two main ways to account for this. One method would be to bin our void catalogue by void radius, reducing this ambiguity by reducing the spread in void size. However, the goal of this work is to maximize the level of significance of our detection, so we want to avoid splitting the void lensing information too finely. The other method which we chose to perform is to compute the covariance for each void individually. We first note that in the reduced form of the covariance equation in Eq. (5.9), we can bring $(4\pi f_{\text{sky}})^{-1}$, which is the inverse of the survey area, into the integral and multiply it by the inverse area density of voids, n_V^{-1} . This results in the inverse total number of voids, N_V^{-1} . For a single void, we set this to one, and in this way compute the covariance per void. We can then sum the covariances and weight them in the exact same way we did for the shape noise

$$\text{Cov} [\Delta\Sigma (\theta_1, z_L) \Delta\Sigma (\theta_2, z_L)] = \frac{\sum_L w_L^2 f(z_L, \theta_1, \theta_2)}{(\sum_L w_L)^2} + \delta_{\theta_1, \theta_2}^K \sigma_{\text{SN}, \Delta\Sigma}^2, \quad (5.10)$$

where

$$f(z_L, \theta_1, \theta_2) = \frac{1}{\langle \Sigma_c^{-1} \rangle^2(z_L)} \int \frac{\ell d\ell}{2\pi} \overline{J_2(\ell\theta_1)} \overline{J_2(\ell\theta_2)} \left[\frac{C_{KK}}{1} \right], \quad (5.11)$$

and w_L is the lens-source weights defined in Eq. (5.4) summed over all sources belonging to each lens.

In principle, one can use the covariance of the signal found using the random catalogues as an empirical estimate of the covariance matrix. This matrix would only be sensitive to the shape noise and C_{KK} terms due to the use of random unclustered lenses. However, we find that the covariance from randoms in the off-diagonal elements is too noisy to be used in lieu of this model for the analysis. This is largely due to correlating random voids with sources from data. Given that voids are very large objects, we quickly begin to oversample the footprint of the survey overlap with increasing numbers of randoms. We therefore find that increasing the number of random catalogues from 30 to 50 does not significantly dampen the noise.

However, the signal in the diagonal elements is larger than the typical fluctuations in the off-diagonal elements (< 0.01), so the diagonal components of the random covariance matrix can be used as a comparison test to help validate our model, as our model is similarly insensitive to the C_{VK} and C_{VV} terms. Since our model does not include any terms that only contribute to the off-diagonal elements, testing the diagonal alone would be sufficient. Fig. 5.2 shows the comparison of the diagonals of the matrices from Eq. (5.10) and from the randoms. We can see good agreement across samples and across bins, with the empirical estimate only having a large disagreement with the model at the smallest bin, which is likely due to scatter from a large shot

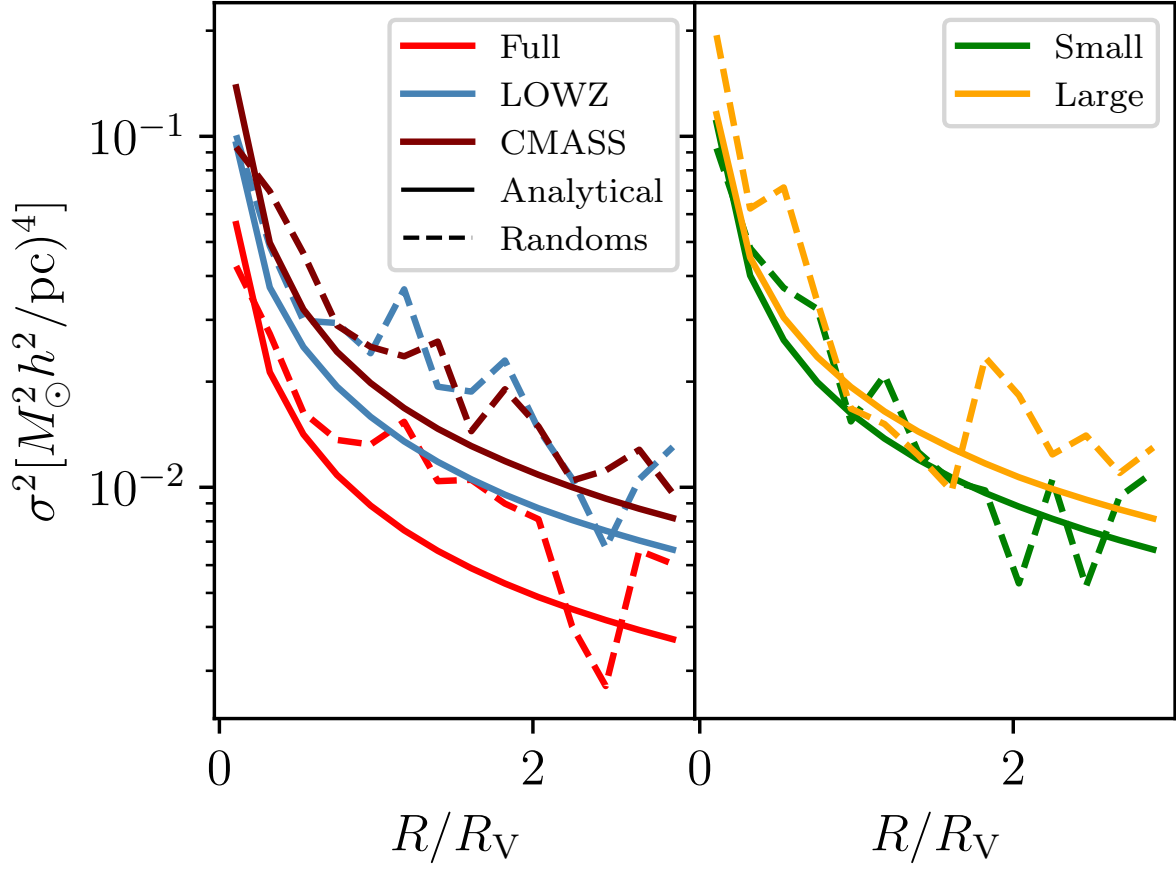


Figure 5.2: Comparison between the diagonals of the analytic covariance calculation and the empirical estimate of the covariance around mock-random lenses for the Full void catalogue and the four additional subsets used in this chapter.

noise due to the smallest bin having the fewest galaxies. This result, coupled with the result in Appendix A.1, then validates the approximations that go into our covariance model.

5.2.4 Void Density Profile Models

Analytic Void Profile

The primary result from this measurement will be an estimate of the excess surface mass density, defined in terms of the surface mass density Σ as

$$\Delta\Sigma(R/R_V) = \overline{\Sigma(<R/R_V)} - \Sigma(R/R_V), \quad (5.12)$$

where $\overline{\Sigma(<R/R_V)}$ is the average of the surface mass density within a circle of radius R/R_V .

To obtain Σ , the three-dimensional mass density profile is integrated along the line of sight. For spherically symmetric profiles, such as the case of a void profile averaged over many voids, this can be simplified to an Abel transform

$$\Sigma(R) = 2 \int_R^\infty \frac{\rho(r) r}{\sqrt{R^2 - r^2}} dr. \quad (5.13)$$

For a particular model $\rho(r)$, we can calculate its corresponding $\Delta\Sigma$ profile, and then fit the model to the data. In the literature, there is one popular choice of empirical void profile presented in [Hamaus et al. \(2014b\)](#), hereafter HSW)

$$\frac{\rho_V(r)}{\bar{\rho}} - 1 = \delta(r) = \delta_c \frac{1 - (r/r_{s1})^\alpha}{1 + (r/r_{s2})^\beta}. \quad (5.14)$$

This model as originally proposed (seen in Eqn. 6.11) has four free parameters: the central underdensity δ_c , the zero-crossing scale r_{s1} , and the power-law parameters α and β , with the r_{s2} parameter fixed to the void radius R_V . We follow the modification proposed by [Barreira et al. \(2015\)](#), hereafter B+15) and free the r_{s2} parameter. This parameter then controls the scale at which the profile asymptotically approaches zero. The HSW model was found to fit the void profile of simulated voids found by the VIDE void finder, and has been found to fit void profiles from both data and simulations with a variety of void finding setups ([Chan et al., 2014](#); [Barreira et al., 2015](#); [Sánchez et al., 2016](#); [Pollina et al., 2017](#); [Falck et al., 2018](#); [Baker et al., 2018](#); [Pollina et al., 2019](#); [Fang et al., 2019](#); [Shim et al., 2021](#); [Tavasoli, 2021](#); [Williams et al., 2025](#); [Davies et al., 2025](#)).

We make note that the voids we consider in this analysis are circumcentre voids, which differ from barycentre voids found by VIDE. This is notable because circumcentres will by definition

find the centre of the largest empty sphere, which can be located anywhere inside of the void. Since the sphere is empty, it is expected that the corresponding matter profile will have a steeper slope inside this sphere around the centre. Barycentre voids on the other hand will not have this feature. This methodological difference may require a modification to the HSW model. There have been some studies that have modified the HSW profile to account for differences in void finding methodology and modified gravity (Barreira et al., 2015; Chantavat et al., 2017; Falck et al., 2018; Baker et al., 2018; Perico et al., 2019; Williams et al., 2025; Davies et al., 2025). Additionally, for *Revolver* specifically, Nadathur & Hotchkiss (2015a) qualitatively find that the five-parameter modification by B+15 produces better visual fits to simulated *Revolver* void profiles.

We investigate the difference in fits between the four- and five-parameter HSW model applied to our voids further in Appendix B. There, we find that these two models are relatively interchangeable (with at most a 1.9σ significance improvement of the fit quality for the Full catalogue down to $< 1\sigma$ in the CMASS and Small catalogues). However, we find that the original four-parameter HSW model produces best-fit values of $\alpha < 1$ and $\delta_c = -1$ across all catalogues. This produces void profiles with concave behaviour in the innermost regions. This is a property that has not been directly observed in the literature, as void profiles are typically flat to convex near their centre. This is likely due to a more fundamental limitation in the void-galaxy lensing methodology generally. Since the lensing profile relies on the number of source galaxies in a given bin, the innermost bins will have the fewest galaxies, therefore the innermost regions will not be well constrained. Since we are directly measuring $\Delta\Sigma$, large error bars in the smallest bins correspond to a large spread in possible inner slopes. We find that the five-parameter HSW model following the modification in B+15 produces best-fit void profiles that are somewhat more agreeable with expectations with $\alpha \geq 1$ and $\delta_c > -1$ at the cost of a more complex posterior region as investigated in Appendix B. We therefore proceed with presenting results from the five-parameter HSW model, but note that the unmodified four-parameter HSW fits are still acceptable.

Void-Galaxy Cross Correlation Measurement

Assuming a scale-independent linear bias relation, we can write the void-matter cross-correlation as a function of the void-galaxy cross-correlation divided by a bias factor b_{Vg} . How this b_{Vg} factor relates to the global linear b_g can then be used to better understand the dependence of galaxy bias on environment by examining its relationship to the least dense parts of the Universe. To this end, we use the results and methodology of Woodfinden et al. (2022), with slight modifications, to measure the void-galaxy cross-correlation with the intent to construct a model with a free parameter b_{Vg} that will be constrained by fitting this model to our void lensing measurement. We refer readers to Woodfinden et al. (2022) for additional details and systematics tests for the

void-galaxy cross-correlation.

The measurement of the void-galaxy cross correlation is performed entirely within the NGC component of the LOWZ and CMASS surveys, in contrast with the void lensing measurement, where we consider the overlap of UNIONS and either LOWZ or CMASS. The galaxy data set used is the same LRG data set used in the void-finding procedure. We note that this means this galaxy data set is also in reconstructed space with approximate RSDs removed, and is therefore dependent on the same fiducial choice of $\beta = f/b = 0.37$ as the location of the void centres. The cross correlation is measured using the [Landy & Szalay \(1993\)](#) estimator.

To account for the same selection bias as considered in the random catalogues used in the void lensing measurement, the void randoms for the void-galaxy cross-correlation are constructed from a subset of 250 of the 1000 PATCHY mocks. The void randoms are then randomly chosen until a catalogue exists with 50 times more randoms than data points for the purposes of the void-galaxy cross correlation measurement. Furthermore, we opt to use the void-galaxy cross correlation measurement in reconstructed space rather than the redshift space correlation that was the primary result of [Woodfinden et al. \(2022\)](#). Correlating with the galaxy field in reconstructed space will in principle produce a more spherically-symmetric profile that will better represent the isotropy expected from the average void lensing profile.

We note that in the case of both void lensing and void-galaxy measurements, we are computing a weighted average across all voids in the sample. We therefore want to ensure that we choose a similar weighting scheme between both measurements. To do this, we slightly modify the setup of [Woodfinden et al. \(2022\)](#). We do not perform a minimum radius cut on the void catalogue, and instead bin voids by void radius, in steps of $10 h^{-1}\text{Mpc}$. This radius cut in [Woodfinden et al. \(2022\)](#) was done for the purposes of analysing the RSD signal, and as such is not needed for our analysis. We then perform the void-galaxy measurement on each of these radius bins, the results of which are plotted in Fig. 5.3. We then build an effective void size distribution from the lensing catalogues (which are subsets of the total NGC void sample), weighted by the lensing weights. We choose the bins of this distribution to match the $10 h^{-1}\text{Mpc}$ bins used in the void-galaxy measurement. We then take the average of the void-galaxy profiles shown in Fig. 5.3, using this void size distribution as weights. This procedure will approximate the weighting scheme used in the lensing analysis, since not all of the NGC void catalogue are in the void lensing catalogues. The resulting void-galaxy profiles binned by void radii are shown in Fig. 5.3, where this average profile used for the Full catalogue is shown in the dashed line. It should be noted that these void-galaxy cross-correlation measurements were performed in unscaled bins in the same manner as [Woodfinden et al. \(2022\)](#) and, as such, will not span the entire range of radial bins used in the void lensing measurement.

For the void-galaxy measurements shown in Fig. 5.3, we note a few particular properties of

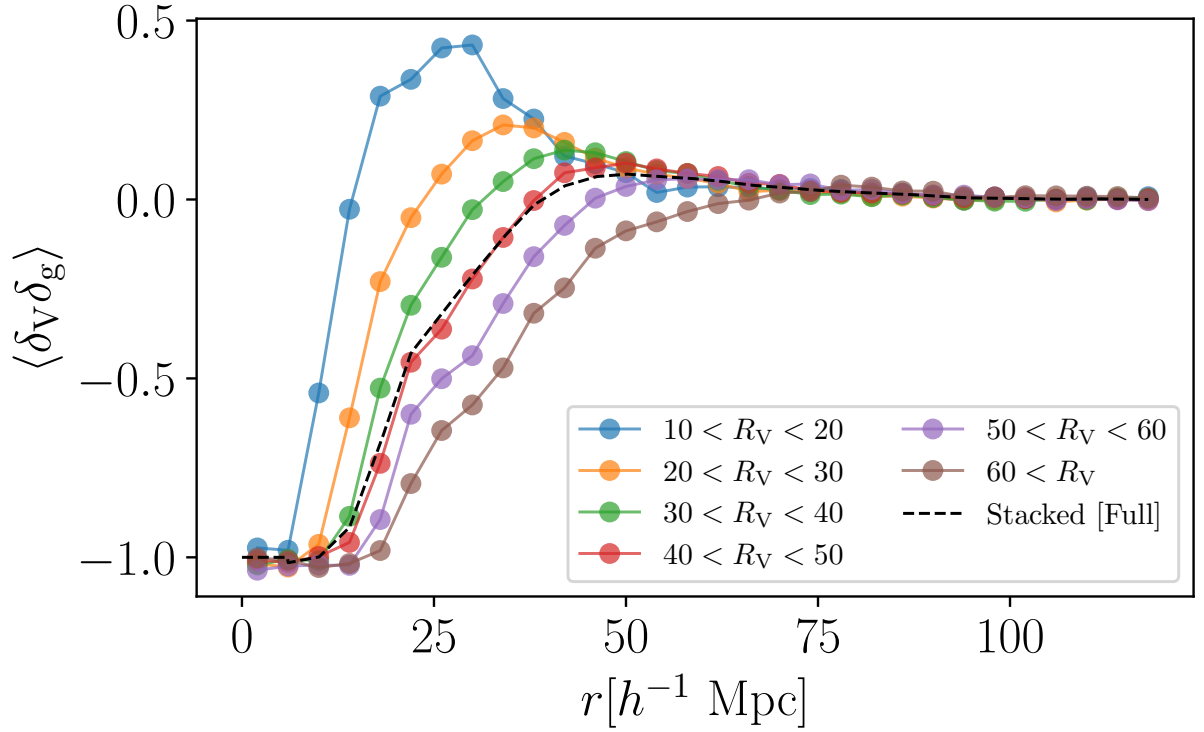


Figure 5.3: BOSS void-galaxy cross-correlation measurements binned by void radius. The black dashed line represents the weighted average across all binned profiles for the Full catalogue. This average is then used as the void-galaxy model for the void lensing measurements to estimate galaxy bias. These measurements use the entirety of the BOSS void catalogue in the NGC as opposed to the region that overlaps UNIONS considered in the void lensing analysis.

the interior of the profiles that can be observed. Since we are considering circumcentre voids, where the centre is maximally located away from all galaxies within the void, it must be the case that the central galaxy underdensity minimizes to -1 due to a constructed lack of galaxies. Furthermore, there exists a kink in the profiles at around 20 to 25 h^{-1} Mpc that exists because of a sharp density transition that occurs between the galaxies used for determining the circumcentre and the location of the centre (Nadathur & Percival, 2018). Since the linear bias modelling of Pollina et al. (2017) considered barycentre voids, which have a more smoothed profile, it may lead to the linear bias modelling failing on these smaller scales. This is largely due to the statistical effect of relying on discrete galaxy counts to measure the profile and not encountering any pairs in some bins. This can partly explain the amplitude of the galaxy bias residual found in Nadathur & Percival (2018), however they additionally find that this scale-dependent bias residual also exists for barycentre voids. In contrast, Curtis et al. (2025), using the TNG300 simulation find galaxy bias terms from circumcentre voids that are fit well by the linear bias model when compared to dark matter profiles of voids in the range of 0.1 to 1.2 R_V . Therefore, we do not yet exclude the innermost region in this model.

In order to convert this measurement into a model that can fit our void lensing measurement, we will assume the linear bias relation

$$\langle \delta_V \delta_g \rangle (r) = \xi_{Vg}^{3D} (r) = b_{Vg} \langle \delta_V \delta_m \rangle (r) = b_{Vg} \xi_{Vm}^{3D} (r) , \quad (5.15)$$

and so

$$\xi_{Vm}^{3D} (r) = \frac{\xi_{Vg}^{3D} (r)}{b_{Vg}} . \quad (5.16)$$

We can thereby treat $1/b_{Vg}$ as a free parameter to scale the amplitude of the void-galaxy cross-correlation measurement to fit our measured void lensing signal. We differentiate this value from b_g , which is the large-scale linear galaxy bias, as the values are expected to differ (at least for smaller voids). However, like the analytic models, the void-galaxy measurements are done in 3D space, while void lensing is done in units of excess surface mass density. Starting with Eq. (5.15), we can take the projection of both sides to arrive at

$$\Sigma (R) = \frac{2\bar{\rho}}{b_{Vg}} \int_R^\infty \frac{\xi_{Vg} (r) r}{\sqrt{R^2 - r^2}} dr , \quad (5.17)$$

where $\bar{\rho}$ denotes the average density of the Universe, which converts the dimensionless void-galaxy cross correlation into the same units as the ESMD measurement. Then, plugging this equation into Eq. (5.12), we can relate ξ_{Vg} to $\Delta\Sigma$ from our void lensing measurement with a free parameter of b_{Vg} .

Model Fitting Methodology

For both the HSW and void-galaxy models, we perform a least-squares fit by minimizing χ^2

$$\chi^2 = \sum_{i,j} (\Delta\Sigma_{i,\text{data}} - \Delta\Sigma_{i,\text{model}}) C_{ij}^{-1} (\Delta\Sigma_{j,\text{data}} - \Delta\Sigma_{j,\text{model}}) . \quad (5.18)$$

The resulting parameter fits for the HSW model are then used to determine the significance of the measurement. This is done by comparing the χ^2 of the best-fit model to the χ_0^2 of a model set to zero in every bin, which represents a null detection

$$\Delta\chi^2 = \chi_0^2 - \chi^2 . \quad (5.19)$$

Using Wilks' Theorem ([Wilks, 1938](#)), the distribution of this $\Delta\chi^2$ statistic is itself a χ^2 distribution, with degrees of freedom equal to the difference in degrees of freedom between the two nested models, which in our case for the modified HSW model is five. The area under the tail of this distribution then carries information about the significance of the detection, and serves as one of the primary results of this chapter. For the void-galaxy model fits, the resulting parameter fit is a measurement and uncertainty of b_{vg} , the galaxy bias parameter in underdense environments, another primary result of this chapter.

We additionally perform a MCMC analysis to the HSW model to explore the posterior around the best-fitting parameters. This was performed using the `emcee` ([Foreman-Mackey et al., 2013](#)) library minimizing the log likelihood, converted from Eq. (5.18) by a factor of -0.5 . For both the MCMC and least-squares fits of the HSW model, we apply a set of priors on the five different parameters. For δ_c , r_{s1} , and r_{s2} , these are uniform flat priors. For α , we enforce $\alpha \geq 1$, as the model goes as $\delta_c (1 - r^\alpha)$ on small scales. As mentioned previously, if $\alpha < 1$, then the matter density profile near $r = 0$ is concave, a behaviour not seen in previous works, which have all shown void profiles with convex behaviour, i.e. whose initial central slope increases with radius.

Additionally, looking at the large-scale behaviour of the HSW model, the density scales as $r^{\alpha-\beta}$, and so this combination controls the rate at which the 3D matter profile drops to 0. In performing initial fits to the Full, LOWZ, and Large void catalogues, we initially tested a prior of $\alpha \leq \beta$, as this prevents models with infinite density at large scales. However, it was found that these catalogues in particular prefer fits with $\alpha \approx \beta$, which results in nonphysical 3D matter profiles with constant density out to the largest scales. This prompted the use of a stronger prior where we rely on the fact that small voids tend to have overcompensated profiles, where the matter in the overdense wall is more than is required to “fill in” the underdensity, while large voids result in compensated or undercompensated profiles. This phenomenon occurs because voids are expected to grow in size with time until shell-crossing occurs, wherein faster-evacuating inner

matter density shells cross with slower matter density shells near the boundary. However, for voids in overdense environments, the collapse of the environment dictates the void size evolution, shrinking the size of the void. While some small voids can exist in neutral or underdense environments, on average smaller voids will tend to be the result of this ‘void-in-cloud’ effect, while few average-to-large voids will exist in overdense environments (see, for example Figure 4 in [Nadathur & Hotchkiss \(2015b\)](#)). This implies that smaller voids on average will have a stronger overdensity signal at larger r/R_V bins than bigger voids. Therefore, using the best-fit model of the Small void catalogue, we construct a prior where all model parameters that result in a density profile with a value larger than the Small catalogue at a sufficiently large radius (the value at $5 r/R_V$ was adopted) are considered to be outside of the prior bounds and therefore rejected, effectively setting the scale where all profiles should be asymptotically close to 0. To test how reasonable this prior is, we can fit the void-galaxy measurements with the HSW profile to quantify their large-scale trends. Since we expect the void-galaxy and void lensing profiles to be similar, this gives us an order of magnitude test for the expected values of $\alpha - \beta$. We find that this choice of prior results in void lensing profiles with slightly lower values of $\alpha - \beta$ than values found in the void-galaxy profiles. This indicates that the void lensing profiles return to zero at large scales slightly faster, but at a relatively similar rate as the void-galaxy profiles, supporting the use of this prior.

5.3 Results

5.3.1 Void profiles

In Figs. 5.4 and 5.5, we show the measured excess surface density profiles in the first column, and the corresponding three-dimensional void matter profiles in the second column for all five catalogues. The ESMD measurements are shown in black points, with the corresponding least-squares fits from the five-parameter HSW model in the black solid line and the void-galaxy correlation in the red dashed line. The blue lines represent 200 random samples of the MCMC posterior of the HSW fit to show the range of acceptable model fits given the priors discussed in Sect. 5.2.4. The error bars on the data points come from the diagonal of the covariance matrix in Eq. (5.10) added with the covariance matrix measured from the signal around random lenses. The orange dot-dashed lines indicate published simulation results that roughly correspond to our catalogues. These will be discussed in Sect. 5.4.1. We additionally find that the cross-component of shear, $\gamma_\times$, is consistent with zero for all five catalogues and therefore do not plot them.

The contour plot for the five free parameters of the modified HSW fit is shown, for the Full sample, in Fig. 5.6. We note that the best-fit parameters, represented by the grey dashed lines, are

located significantly far away from the mean of the marginalized posteriors in some cases. This is further explored in Appendix B and may be due to the union of two distinct posterior regions of the five-parameter HSW model that produce similar quality fits to the data. We also find that this behaviour is not seen in four-parameter HSW model fits to our data. Fit statistics and parameter values are compiled for all five void catalogues in Table 5.1 for the HSW model and Table 5.2 for the void-galaxy cross-correlation model. For the parameter values, we report both the mean and 1σ spread of the posterior, as well as the best-fit value in parentheses. Due to the complex nature of the posterior observed in Fig. 5.6 (e.g. the marginalized α PDF), in cases where the posterior is significantly skewed, we report the 95 per cent upper or 5 per cent lower limit. For the significance measure of each profile, we report the area under the tail of the χ^2 distribution used in Wilks' Theorem as described in Sect. 5.2.4 converted to the equivalent area under both tails of a Gaussian distribution, i.e. in units of σ , for convenience in interpretation.

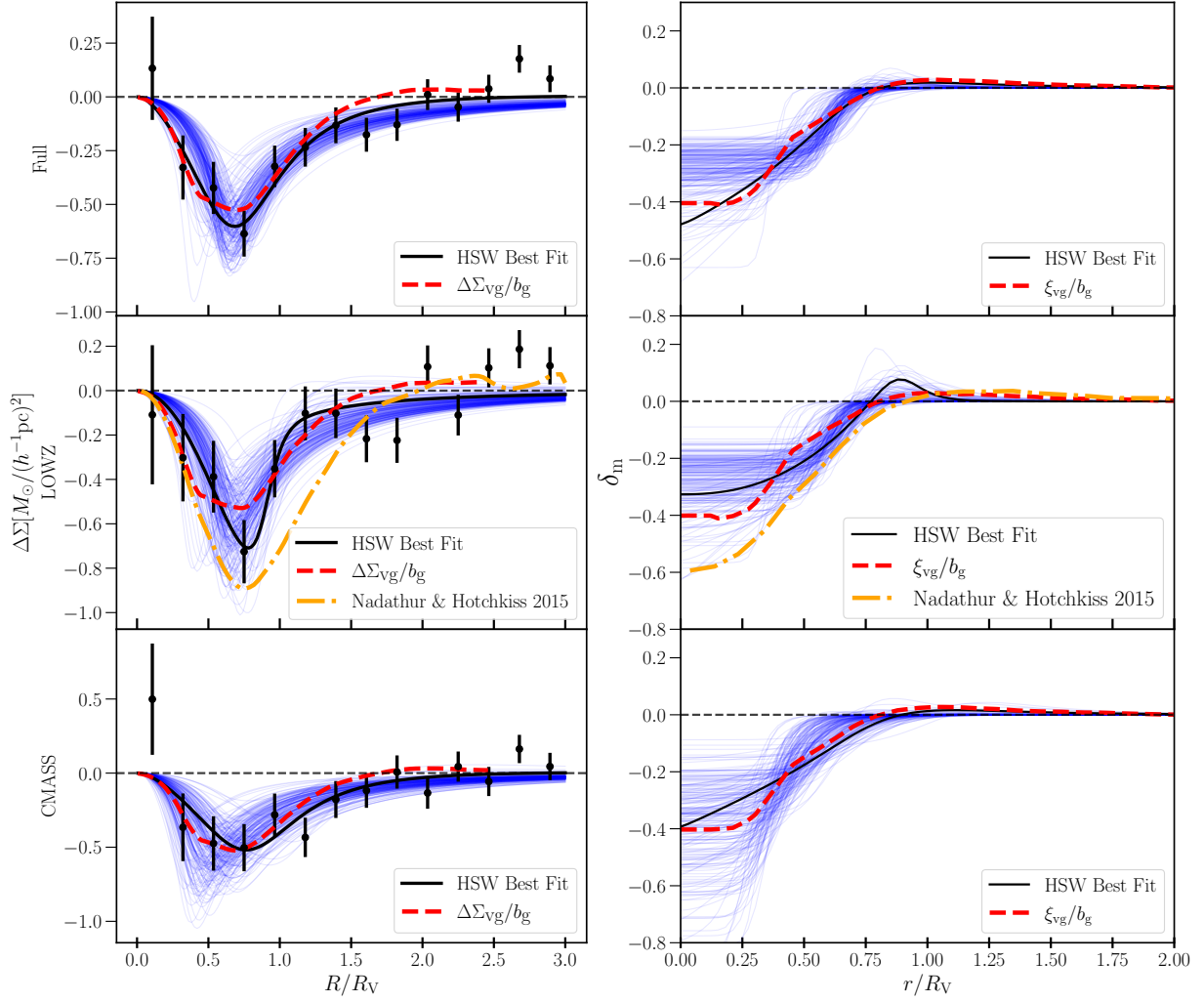


Figure 5.4: Void lensing measurements in $\Delta\Sigma$ units as a function of projected R/R_V on the left, and the corresponding 3D density profiles as a function of three-dimensional r/R_V on the right. These measurements are fit with two models: the analytic five-parameter HSW model in the black solid line, and the best fit bias scaling of the void-galaxy cross-correlation measurement. The thin blue lines show 200 draws from the HSW model MCMC chains. Orange dot-dashed lines show comparisons to applicable published simulation results. We note that the right column is plotted from 0 to 2 instead of 0 to 3 in r/R_V to highlight the non-zero portion of the plot.

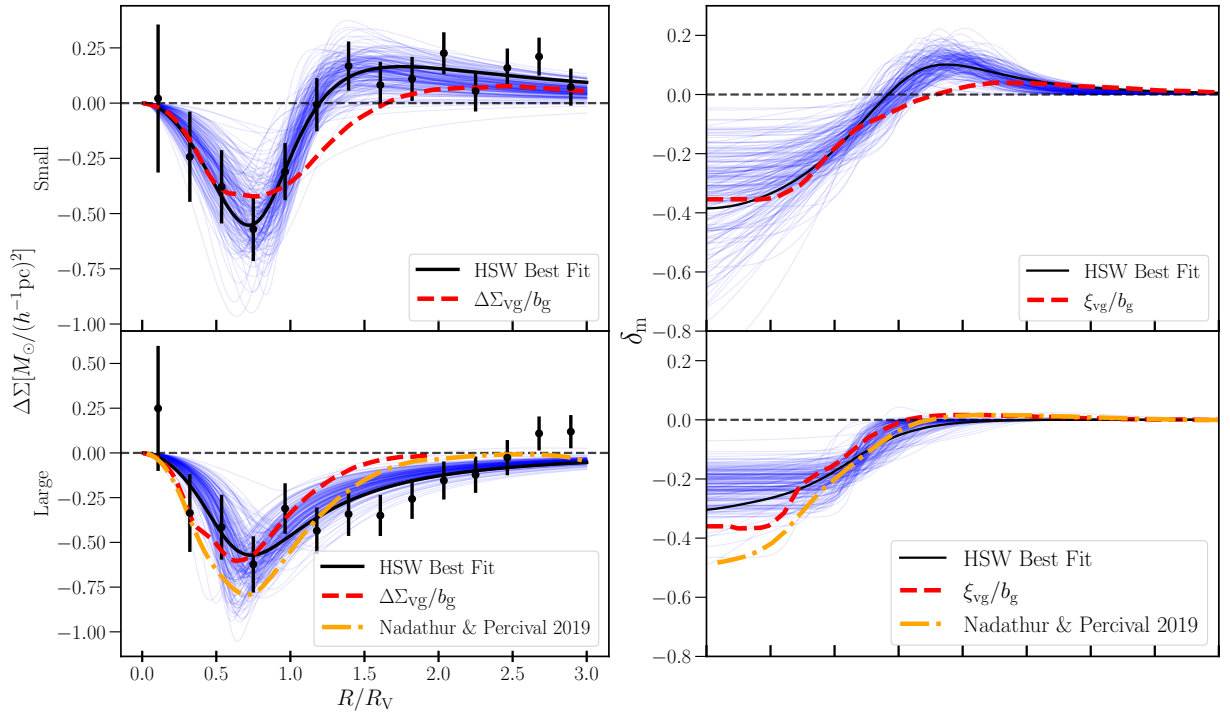


Figure 5.5: Same as Fig. 5.4, but for the catalogues split by void size.

Table 5.1: Fit statistics and significance measurements for the 5 void catalogues used in this chapter. The model considered here is the five-parameter HSW analytical formula, also used to determine the significance (Sig.) of the measurement. Quantities with error bars represent the mean of the projected MCMC posteriors with 1σ error bars. Quantities in parentheses represent the best-fitting model in the multidimensional posterior. Skewed posterior distributions are presented with upper or lower bounds, and represent the 95 per cent or 5 per cent upper or lower bound, respectively.

Five-Parameter HSW Model							
Sample	χ^2/dof	Sig.	δ_c	r_{s_1}/R_V	r_{s_2}/R_V	α	β
Full	15.6/9	6.2σ	$-0.30^{+0.09}_{-0.04}, (-0.48)$	$1.21^{+0.62}_{-0.54}, (0.81)$	$0.63^{+0.05}_{-0.08}, (0.74)$	$<14.8, (1.2)$	$>7.3, (6.0)$
LOWZ	20.2/9	4.4σ	$-0.28^{+0.10}_{-0.05}, (-0.33)$	$1.19^{+0.80}_{-0.53}, (0.77)$	$0.65^{+0.07}_{-0.10}, (0.91)$	$<15.1, (2.4)$	$>8.03, (20.0)$
CMASS	12.3/9	3.4σ	$-0.39^{+0.19}_{-0.10}, (-0.39)$	$>0.73, (0.90)$	$0.52^{+0.08}_{-0.16}, (0.85)$	$<7.4, (1.1)$	$7.9^{+2.1}_{-0.8}, (6.2)$
Small	5.7/9	4.5σ	$-0.33^{+0.16}_{-0.09}, (-0.39)$	$0.75^{+0.04}_{-0.06}, (0.71)$	$0.93^{+0.13}_{-0.13}, (0.87)$	$<8.8, (2.0)$	$12.1^{+2.0}_{-4.8}, (7.3)$
Large	11.8/9	4.4σ	$-0.24^{+0.08}_{-0.05}, (-0.30)$	$>0.81, (1.37)$	$0.67^{+0.06}_{-0.11}, (0.66)$	$<15.0, (1.2)$	$>6.9, (5.3)$

Table 5.2: Model fits and fit statistics for the void-galaxy cross-correlation model to the void-lensing data for all 5 void catalogues in this chapter.

VG Model		
Sample	$\chi^2_{\text{vg}}/\text{dof}$	b_{vg}
Full	10.2/10	2.47 ± 0.36
LOWZ	17.4/10	2.49 ± 0.49
CMASS	9.5/10	2.48 ± 0.55
Small	15.1/13	2.82 ± 0.60
Large	14.3/8	2.77 ± 0.57

From the $\Delta\Sigma$ plots of all 5 catalogues, we can clearly observe underdense structure, with minima at just short of the void radius, in agreement with the void lensing profiles of other works. The HSW fits from the MCMC chain visually trace a scatter in the ESMD profile predicted by the error bars, short of the last two or three bins. Fitting these last bins results in $\Delta\Sigma$ profiles that trend upwards at large scales, which result in model fits where $\alpha \approx \beta$. As discussed in Sect. 5.2.4, these model fits produce 3D density profiles that are near-constant and non-zero at large scales, which are unphysical. We therefore have chosen a set of priors that reject these types of fits, which has resulted in $\Delta\Sigma$ profiles that asymptotically approach 0 at large scales. This is with the exception of the Small voids, whose $\Delta\Sigma$ measurement turns positive at comparatively much smaller radial bins, avoiding this issue. However, looking at the χ^2 values, these restricted fits are still acceptable for all catalogues except for LOWZ, which is marginally rejected ($p \lesssim 0.05$). However, this marginal rejection still holds even when the prior bounds are relaxed. For the LOWZ profile in particular, the least-squares fit pushes for abnormally large values of β , where we report the best fit here as the maximal value of the prior bound set on β for all catalogues (20). If this prior bound is relaxed, the LOWZ profile fits values of $\beta \approx 250$ for a marginal change in χ^2 , which still results in marginally rejected fits. This indicates that the LOWZ catalogue does not place useful constraints on β .

Comparing the profiles of the Small and Large samples, we can observe that smaller voids and larger voids seem to have substantially different $\Delta\Sigma$ profiles, with Small voids crossing zero just past the void radius, while Large voids cross zero at around 2.5 times the void radius. This is in line with findings that smaller voids tend to have more pronounced overdense walls than larger voids (Hamaus et al., 2014b; Nadathur & Hotchkiss, 2015a). As these profiles represent the average of a stack of voids, this supports the notion that smaller voids are more likely than large voids to be “void-in-clouds,” where the smaller-scale underdensity may be situated in a larger-scale overdense environment. Looking at the corner plot in Fig. 5.7, we can see that these differences in the model fits may not be as substantial, and may be consistent with each other at the

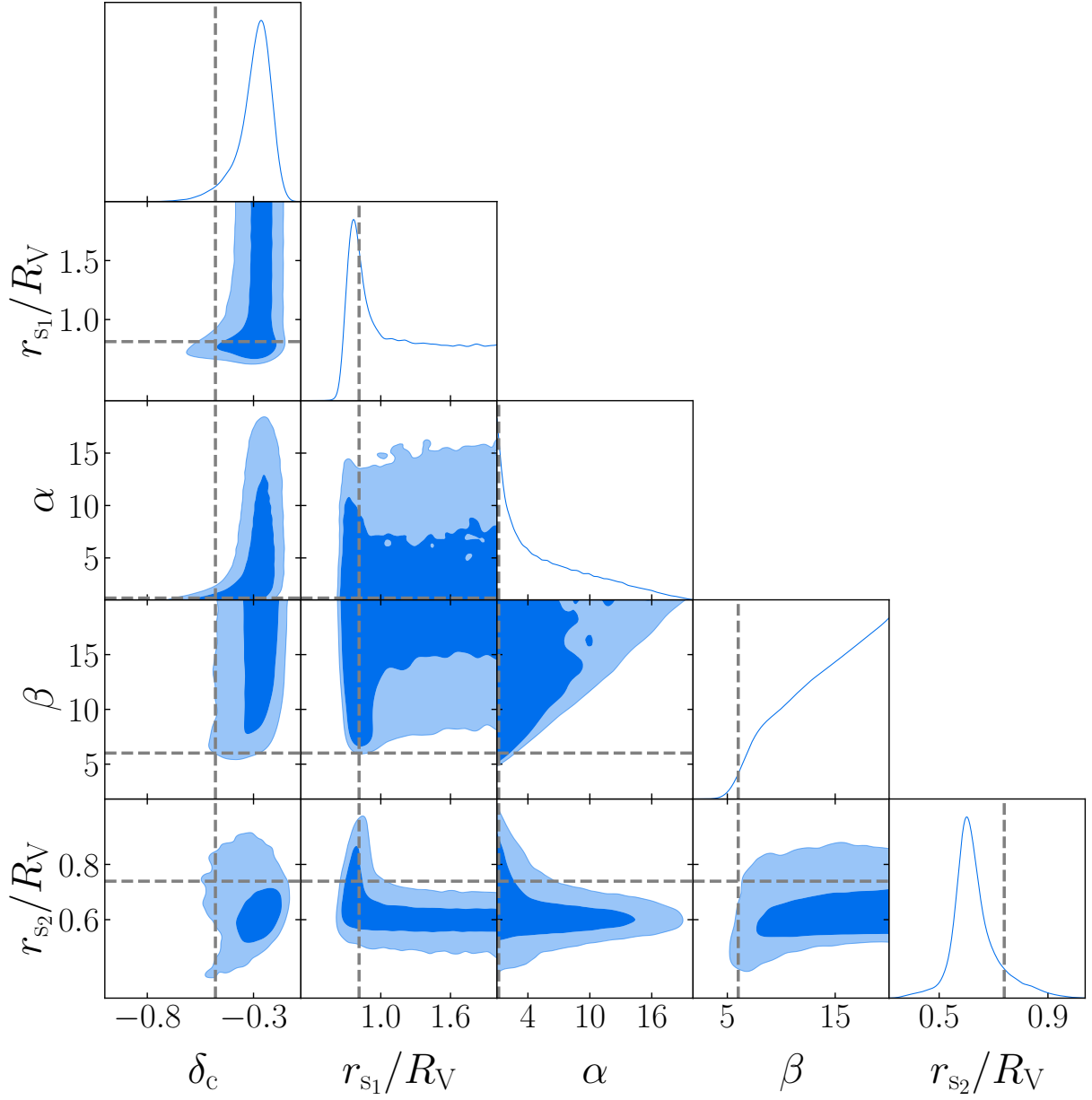


Figure 5.6: Contour plot of the five-parameter HSW model fit to the Full void catalogue. The least-squares fit is shown in the intersection of the dashed lines. Dark shaded regions indicate 1σ contours, while light shaded regions indicate 2σ contours.

1 to 2σ level, with any difference being driven primarily by r_{s1} , which affects the zero-crossing scale, and r_{s2} , which primarily affects the large-scale behaviour of the voids. This consistency is driven mainly by the uncertainty in the Large model fits. However, if we subtract the void lensing measurements from each other directly, we find that the void lensing profiles of the Small and Large catalogues differ significantly at the 2.3σ level.

To evaluate the linear biasing model, we can compare the predictions (red dashed curves) to the data. We note that the void-galaxy cross-correlation measurements that went into these models were measured to a maximum radius of $118 h^{-1}$ Mpc and as such will not extend across all radial bins of all catalogues. We find good visual agreement between the best-fitting void-galaxy model and the data for the Full and CMASS catalogues, which is supported by the χ^2_{vg} values reported in Table 5.2. Additionally, from the table, the void-galaxy fits to the Small catalogue, whose fits visually appear in the left column of Fig. 5.5 of similar quality to perhaps the LOWZ fits in Fig. 5.4, are actually a good fit to the data. The fits of the void-galaxy cross-correlation model to the Large and LOWZ catalogues are marginal, with p -values of 0.07 and 0.06, respectively. This may be because the model does not span the full extent of the data. To test, we can extrapolate the void-galaxy $\Delta\Sigma$ model by setting all values at radii larger than $118 h^{-1}$ Mpc to zero. We then find that the p -values for the Large and LOWZ catalogues become 0.13 ($\chi^2/\text{dof}=18.7/13$) and 0.04 ($\chi^2/\text{dof}=22.9/13$), respectively. While the extrapolated LOWZ catalogue produces a slightly poorer fit, the extrapolated fit has a p -value of the same order-of-magnitude as the HSW fit. In summary, these results suggest that the linear biasing model is not rejected by these measurements.

For the 3D overdensity plots in the second column of Figs. 5.4 and 5.5, we can compare the HSW fits and their scatter, extracted from the chains, with the void-galaxy tracer profile divided by the best-fitting b_{vg} parameter. We note that for the Full and CMASS catalogues, the best-fitting matter profiles trace the void-galaxy profile and the general scatter fairly closely, except at small radii. There, the best fit (black line) has a linear slope on small scales while the chain samples tend towards constant values at small scales. This small-scale discrepancy is due to behaviour seen in the corner plot of Fig. 5.6, which exhibits similar trends as in the CMASS catalogue. The resulting posterior contours are fairly complex, which results in the best fits for some parameters, notably δ_c , β , and r_{s2}/R_v , represented by the grey lines, being found away from the mean of the posteriors. We do however find that, when compared visually, the best fitting models are in better agreement with the void-galaxy profiles than a model constructed from the mean of the posterior. This is why we additionally include the parameter values of the least-squares fit in Table 5.1 in parentheses. We examine the properties of this complex posterior, including its mean, and how it may be due to the union of two distinct parameter regions of the five-parameter HSW model in Appendix B.

The 3D matter density profiles of the Full, CMASS, and Large samples also indicate that these

voids do not have a significant overdense wall, which is consistent with the void-galaxy profiles. This follows expectations from prior void analyses in simulations where larger voids tend to be compensated or undercompensated and thereby have less pronounced walls, as well as the use of the circumcentre definition for the void centres, which primarily presents with dampened walls in the density profiles in comparison to void profiles with the barycentre definition [Nadathur & Hotchkiss \(2015b\)](#). This is due to the minimum density centre not necessarily being located at the centre of the void, so the walls of the void become averaged out in the stacked profile. The LOWZ sample best fit HSW profile stands out with a comparably sharply peaked overdense wall, something that is not observed in the void-galaxy profile. This feature may be a result of the aforementioned weak β constraints offered by this measurement. A similar overdense feature is prominent within the matter profile of the Small void catalogue as well, which is expected given that smaller voids tend to be overcompensated or “void-in-clouds”, but the corresponding void-galaxy profile exhibits a weaker overdense feature. Although, as seen in Table 5.2, the void-galaxy model is still an acceptable fit to the Small catalogue.

5.3.2 Integrated Density Contrast

In order to assess the depth of the matter underdensities measured with this weak lensing analysis, in principle one can simply read off the best-fit values of the δ_c parameter of the HSW model, as this represents the value of the profile at the center. However, this value represents an extrapolation of the HSW model fit to $R = 0$. Since the innermost bins are the noisiest because they are based on the fewest galaxy shapes, δ_c will not be well-constrained by the data. Instead, to get a more robust measure of the depth of the void, we can calculate the integrated density contrast, Δ , defined as the volume-normalized integral to a radius r of the HSW model in Eq. (5.14)

$$\Delta(r) = \frac{3}{r^3} \int_0^r \delta(y) y^2 dy. \quad (5.20)$$

Then, choosing different values for the integrated radius can provide details about both the depth of the void stacks and how well each void stack is compensated (i.e. how closely the matter in the overdense wall can re-fill the void). To this end, we present the integrated density contrasts of the best-fit HSW models at three different values: $0.5R_V$, R_V , and $3R_V$ in Table 5.3. We find that, unlike the distribution of δ_c , the distribution of Δ from the MCMC samples at all three radius values are Gaussian and centred near the value at the best fit. Therefore, we can use the standard deviation in the MCMC samples as the error bars for the Δ measurements.

We observe that the integrated density contrasts out to half of the void radius are all consistent with each other for all 5 catalogues. This is somewhat expected for the Full, LOWZ, and CMASS

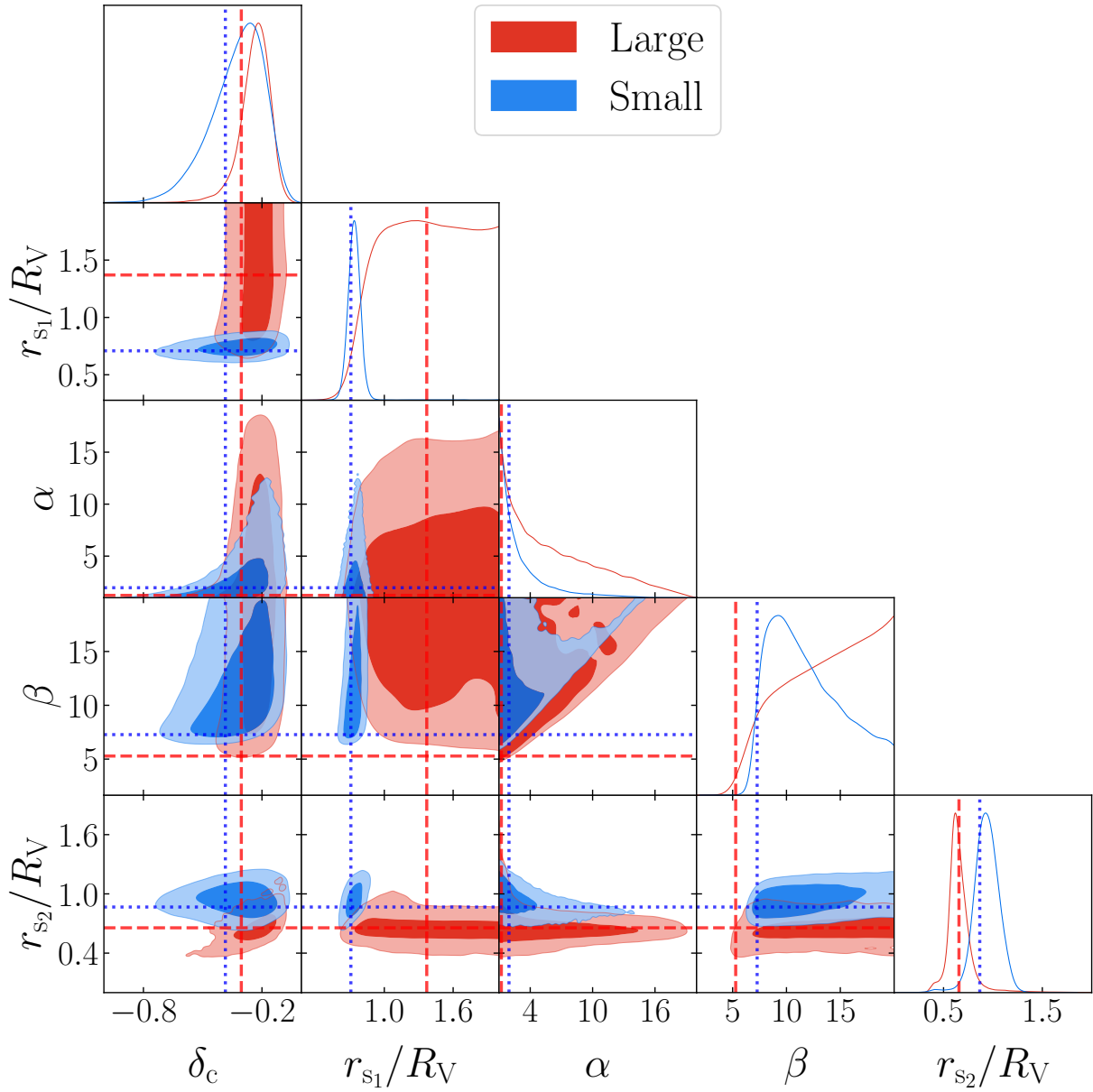


Figure 5.7: Corner plot of the HSW MCMC analysis for both the Large and Small void catalogues.

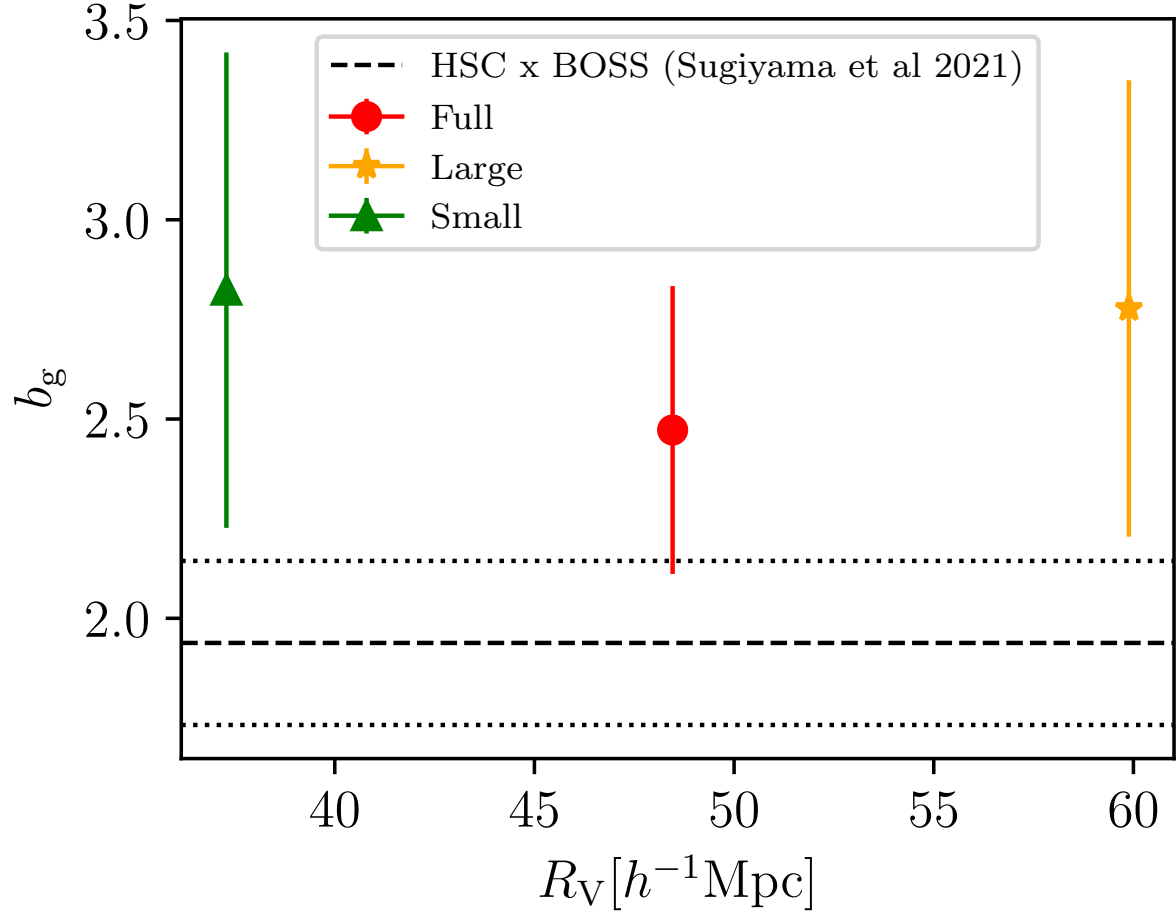


Figure 5.8: Galaxy bias estimates as a function of void size from comparing the void-galaxy cross-correlation to the void lensing measurements, as well as the large-scale galaxy bias measurement from [Sugiyama et al. \(2022\)](#) for comparison. The dotted lines represent the 1σ error on the averaged [Sugiyama et al. \(2022\)](#) galaxy bias, represented by the dashed line. We note that the Full measurement does not constitute an independent measurement, but it is intended to show the full constraining power in this analysis.

Table 5.3: Integrated density contrasts computed using the best-fit HSW profile shown in Figs. 5.4 and 5.5 at three different radius values. Error bars represent the standard deviation in the MCMC samples.

	$\Delta (r = 0.5R_V)$	$\Delta (r = R_V)$	$\Delta (r = 3R_V)$
Full	-0.28 ± 0.04	-0.055 ± 0.011	0.0003 ± 0.0006
LOWZ	-0.26 ± 0.06	-0.037 ± 0.016	-0.0010 ± 0.0008
CMASS	-0.24 ± 0.07	-0.063 ± 0.017	0.0003 ± 0.0008
Small	-0.27 ± 0.07	-0.001 ± 0.025	0.0077 ± 0.0023
Large	-0.22 ± 0.06	-0.072 ± 0.015	-0.0025 ± 0.0007

catalogues as void depths are not expected to vary significantly over this range of redshift. What is of note is that the Small and Large catalogues are consistent with each other at this radius, which could be due to the lack of constraining power of the measurement as seen in Fig. 5.5. This is somewhat unexpected, as typically smaller *Revolver* voids tend to exhibit shallower central underdensities than larger voids (Nadathur & Hotchkiss, 2015b).

Looking to integrated density contrasts at the void radius, the Small and Large catalogues have a significant difference, with the Small catalogue Δ consistent with being compensated ($\Delta = 0$) at this scale, while the Large catalogue features the deepest density contrast of the five catalogues. This, coupled with the consistency at $0.5R_V$ indicate that the differences in the density distributions comes largely from the placement and slope of the void boundary relative to the central underdensity. The Small catalogue’s density profile, seen in Fig. 5.5, features a relatively close zero-crossing scale, with a prominent overdense wall, so it is reasonable that the integrated density returns to zero at the smallest radii, while the Large profile is the only one across all five profiles to not have crossed zero by R_V , and as such is the deepest integrated density of the five.

The final column of Table 5.3 provides information on the overall compensation of the different void catalogues. At a distance of $3R_V$, one would expect to start probing the large-scale effect of the void on its surroundings. An overcompensated void would have a positive integrated density on these scales, while an undercompensated void would have a negative integrated density, indicating a much larger void-like environment. The Full, LOWZ, and CMASS are consistent with a compensated void, in agreement with the results of the simulated LOWZ voids of Nadathur & Hotchkiss (2015a). The Small voids, which were consistent with compensated at R_V , are now slightly overcompensated, in agreement with the relationship shown in Fig. 6 of Nadathur & Hotchkiss (2015a). We additionally find the Large catalogue remains undercompensated at this scale.

5.3.3 Galaxy Bias in Underdense Environments

One of the primary purposes of measuring the galaxy bias in this data set was to test the findings of [Pollina et al. \(2017\)](#), who found from simulated voids that b_{vg} depends on void size. One of the primary results of that work is that galaxies within smaller voids exhibit $b_{\text{vg}} > b_{\text{g}}$ by roughly 17 per cent, but b_{vg} asymptotically approaches b_{g} , the large-scale galaxy bias, for the largest voids. In Fig. 5.8, we plot the galaxy bias constraints from our Small, Large, and Full catalogues, noting that since Full is the combination of the Small and Large void catalogues, it is not an independent measurement. Instead, it is meant to show the tightest possible constraint we can make with this data. We can compare our bias measurements to the results of [Sugiyama et al. \(2022\)](#), who constrained b_{g} for LOWZ and CMASS independently, using a combination of LOWZ, CMASS and Hyper Suprime Cam (HSC) data to jointly measure $\Delta\Sigma$ and projected galaxy clustering. We find that all three points are consistent with the averaged large-scale bias from [Sugiyama et al. \(2022\)](#), though notably for all measurements, $b_{\text{vg}} > b_{\text{g}}$, and the error bars are too large compared to the fractional change observed in [Pollina et al. \(2017\)](#) to make definitive claims regarding any possible trend in the results. We do note, however, that the consistently larger b_{vg} measurement – compared to b_{g} – is consistent with the predictions of [Nadathur & Percival \(2018\)](#), who find a discrepancy between the void-galaxy measurement and the void density profile of simulated voids as a function of the distance from the centre of the void of up to 25 per cent when using the b_{g} value in the linear bias model. This discrepancy would show up as an inflated value of b_{vg} compared to b_{g} . Comparing our result to [Sugiyama et al. \(2022\)](#), we find $b_{\text{vg}}/b_{\text{g}} = 1.36 \pm 0.27$, consistent with a higher value of b_{vg} , but with uncertainties sufficiently large that we cannot exclude equality.

Given the nature that circumcentre voids by construction will not contain void-galaxy pairs within the innermost bins, we then question whether these innermost bins are suitable for use in the linear bias model. We attempt to exclude bins at less than $0.5 R/R_{\text{V}}$ from our fits to test. However, we find that doing so increases the values of the best-fit galaxy biases to scatter around a value of three in all cases while the error in the fits grow significantly, in some cases quadrupling in size. However, the larger error is expected, as stacked circumcentre voids typically do not have strong overdense walls, and thus there is not much signal to fit to once the central underdensity is excluded. We additionally find good visual agreement with the void-galaxy profile fit to the first three bins of the $\Delta\Sigma$ measurement. Therefore we do not wish to exclude these bins from the fits in order to get a realistic estimate of the error bars to determine how well the galaxy bias can be measured in data. We instead note that given these results, the current uncertainties in the lensing profile measurements are likely too large, particularly in the inner region, to determine if the lack of void-galaxy pairs near the centre of circumcentre voids causes the linear bias model to break down on those scales.

5.4 Discussion

5.4.1 Comparison with Previous Results

There have been few detections of the weak lensing signal from cosmic voids. The result here, a 6.2σ detection from 2975 voids, is only the third measurement of void-galaxy lensing using voids identified in spectroscopic tracers, and the second for spectroscopic ZOBOV-based voids. One goal of this analysis is to justify efforts in developing methods to interpret the resulting signal, whether it be for the purposes of cosmological parameter estimation or for understanding the formation and evolution of voids themselves, by showing that this signal is already detectable at high significance levels using currently available data. Given that void statistics, including the void matter profile, are sensitive to differences in void finders, survey geometry, and tracer properties (Nadathur & Hotchkiss, 2015b; Massara et al., 2022), it remains difficult to directly compare results from different works. As such, the most direct comparison we can make with this work is to the results of Melchior et al. (2014), as they also used a ZOBOV-based void finder on a spectroscopic tracer sample from the 7th data release of SDSS. However, the 7th data release contains a catalogue of main sample galaxies, as well as a separate catalogue of LRGs. Using the model of Lavaux & Wandelt (2012), their most significant measurement combining both catalogues was 2.4σ , while the measurement using only the LRG catalogue was 1.3σ . The second spectroscopic void lensing result comes from Clampitt & Jain (2015), however they use a void finder that identifies voids in projected slices which may result in significantly different void profile properties. It is therefore not clear how to directly compare our 3D void lensing profiles to their void lensing profiles.

We can compare the detection significance with that of other works as well, but we must note the caveats surrounding the methodologies. As an example, Fang et al. (2019) used the VIDE void finder on Year 1 of the DES survey, which is an imaging survey with photometric redshifts. For their 3D ZOBOV voids, they find a SNR of 14.0, using the equation from Becker et al. (2016):

$$S/N = \frac{\sum_{i,j} \Delta \Sigma_i^{\text{data}} C_{ij}^{-1} \Delta \Sigma_j^{\text{model}}}{\sqrt{\sum_{i,j} \Delta \Sigma_i^{\text{model}} C_{ij}^{-1} \Delta \Sigma_j^{\text{model}}}}. \quad (5.21)$$

Using our measurement, we find a SNR of 5.52 for the Full sample, with a maximum of 6.42 for the Small sample, with the increase due to the more pronounced overdense wall present in smaller voids. While our detection is formally less significant than that of Fang et al. (2019), it should be noted that, as demonstrated in simulations in that paper, 3D voids found from tracers with photometric redshifts experience a selection bias such that the resultant catalogue will be preferentially oriented along the line of sight. As shown in their Fig. 3, this results in a significant

anisotropic void-galaxy cross correlation with the line of sight axis being 8 to 10 times the length of the other axis. Consequently, the projected density, to which weak lensing is sensitive, will have a much deeper underdensity than for voids found using spectroscopic tracers, providing a much stronger void lensing signal but at the cost of breaking assumptions of isotropy in void models. As pointed out in their work, this results in their HSW fits to the DES Y1 void lensing measurements fitting to the boundary value of $\delta_c = -1$ in some cases.

There are two simulated void matter density profiles to which we can directly compare our results. They are both shown in Figs. 5.4 and 5.5 as orange lines. The first is the LOWZ void density profile from Appendix A of [Nadathur & Hotchkiss \(2015a\)](#) chosen to match the average radius of our void lensing profile, and the second is the void profile from in Fig. 1 of [Nadathur & Percival \(2018\)](#), although it should be noted that this profile is constructed of simulated CMASS voids, and includes all voids where $R_V > 43 h^{-1} \text{ Mpc}$, while our Large catalogue includes LOWZ and CMASS voids, with a radius cut at around $48 h^{-1} \text{ Mpc}$. Both of these profiles come from simulated catalogues of circumcentre voids. We find that compared to these simulations, our 3D void density profiles are significantly shallower both within the void at small scales and at the outside wall at large scales. This results in the simulated profiles exhibiting deeper $\Delta\Sigma$ profiles than our measurements. However, when treating these simulated profiles as fixed models and comparing them to the data, we find that the simulated matter profile of Large voids ([Nadathur & Percival, 2018](#)) is an acceptable fit to the data, with χ^2/dof of 19.4/14. As expected from a visual comparison, we find that the simulated profile of LOWZ voids ([Nadathur & Hotchkiss, 2015a](#)) is a poor fit to the data, with χ^2/dof of 38.6/15 ($p = 7.4 \times 10^{-4}$). The increased depth of the simulated voids, particularly in the case of LOWZ, is an interesting result that may warrant further investigation, as it could be due to differences in the void selection, untested systematics in the void lensing methodology, or more physical in nature, such as the effect of neutrinos which would free stream into the void centre.

5.4.2 Future Work

The significance of this detection represents a promising step for the future of weak lensing around voids as a viable probe. It shows that currently available ground-based survey data is of sufficient quantity and quality to detect this weak effect. With the vastly improved data quality expected from the rapidly approaching advent of Stage-IV surveys like *Euclid*, we expect to detect this signal with even greater precision (see Chapter 6). We also expect the detection from UNIONS to become stronger as more data becomes available from the survey. In particular, the inclusion of photometric redshift information would allow us to create narrower source bins in redshift and more accurately isolate and remove sources in front of our lenses, which act as a contaminating factor. Therefore, we would expect the constraints to improve significantly when

the full photometric UNIONS catalogue is available. Furthermore, current spectroscopic surveys, such as DESI (DESI Collaboration et al., 2016), will improve the coverage and quality of void catalogues in the NGC, and their expected overlap with UNIONS will provide another promising dataset with which to study void lensing.

This then raises the question of what additional analysis is needed to extract useful information from these void lensing profiles. One of the large uncertainties in extracting information from void statistics comes from the complex influence that the specific void finder has on a given statistic. As such, it may become necessary to simulate the specific void setup used in the data to properly test common assumptions and systematics, particularly for circumcentre voids, which are not as well studied as barycentre voids. For extracting the information itself, an emulator-based method may be ideal, as consistently running the same void finder across different cosmologies in a training set will allow an emulator model to learn how the specific void lensing profile changes with cosmology without the need to model out the relationship between the void finder and the void statistic. This is a procedure in development with void-galaxy cross-correlations and is seeing success (Fraser et al., 2025). Use of these emulators, depending on the variable parameters in the training simulation suite, should allow for constraints on cosmology, neutrino mass, and modified gravity.

5.5 Conclusions

This chapter represents a significant development in the detection of the void lensing signal from spectroscopic voids and the quality of available data for measuring it. To make this measurement, we used the substantial overlap of the UNIONS weak lensing survey with voids found in the foreground spectroscopic BOSS LOWZ and CMASS surveys. We also demonstrated that the use of mocks for selecting random lenses is necessary for voids to account for selection effects due to the survey boundaries. Additionally, we adapted and applied the standard Gaussian component of covariances of 2-point statistics to void lensing for the first time. By fitting the analytic matter density profile of Hamaus et al. (2014b) with the modification of Barreira et al. (2015), we determined the significance of our measurement to be 6.2σ , the most significant detection of 3D spectroscopic voids to date. We also compared our void lensing measurement to an adapted form of the void-galaxy cross correlation measurement from Woodfinden et al. (2022) in order to obtain a measure of the galaxy bias relative to underdense environments, at a value of $b_g = 2.45 \pm 0.36$ from the Full catalogue, consistent with the averaged large-scale galaxy bias value of Sugiyama et al. (2022). We find good agreement with a scale-independent linear bias model, but we do require more statistical precision in order to evaluate the findings of Pollina et al. (2017) and determine the trend of b_{Vg} with void size. Our galaxy bias findings are however consistent with

the findings of [Nadathur & Percival \(2018\)](#), who predict that the large-scale linear galaxy bias will underestimate the bias between the void-galaxy correlation and the void density profile. We believe that higher statistical precision can be achieved with the full tomographic sample of the UNIONS survey, as tomographic binning will allow for better control and removal of foreground sources that act as interlopers in our signal.

Chapter 6

Euclid Forecasts

6.1 Introduction

The statistical study of voids in data has only recently become possible in the era of “Stage-III” ([Albrecht et al., 2006](#)) wide-field surveys like UNIONS and SDSS (BOSS, eBOSS), that have provided data over tens of thousands of square degrees collectively. These surveys have led to the development of large enough void catalogues to develop statistical probes to extract cosmological information and to study large-scale structure. We find from the previous chapter that we are able to significantly measure the void lensing profile from these data sets. *Euclid*, as a Stage-IV survey, promises to usher in a brand new era in both depth and area, with over 15 000 square degrees of imaging from the VISible instrument (VIS), and spectroscopy from the Near-Infrared Spectrometer and Photometer (NISP) planned by the end of the survey. Previous void forecasts indicate that *Euclid* will provide a wealth of information for void studies ([Hamaus et al., 2022](#); [Contarini et al., 2022](#); [Bonici et al., 2023](#); [Radinović et al., 2023](#)). We are therefore interested in taking advantage of the dual nature of the spectroscopic NISP survey for void finding with the imaging of the portion of the VIS survey in the background to forecast the weak lensing signal of spectroscopic voids.

One feature of void analyses is the flexibility in the definition of a void. A void is simply defined as a region of the Universe with less matter density than average. Given that we cannot directly view the underlying matter distribution, we must rely on a biased bright tracer, such as galaxy populations with photometric or spectroscopic redshift estimates, or other objects like quasars ([Hawken et al., 2020](#); [Aubert et al., 2022](#); [Kovács et al., 2022](#); [Arsenov et al., 2025](#)). This choice in tracer has a significant impact on the resulting void statistics ([Sutter et al., 2014](#); [Pollina et al., 2019](#)). This allows for the selection of the void definition that is most useful to exploit a

given data set for a given scientific objective. To best utilize the spectroscopic overlap to recover the three-dimensional shape of the void, as well as to make our resulting profiles consistent with future *Euclid* void analyses, we are therefore interested in using the void catalogue of [Hamaus et al. \(2022\)](#) as described in Sect. 4.4.1, who forecasted the cosmological constraints of the void-galaxy cross-correlation measurement. The goal of this chapter is to forecast the improvement that *Euclid* can offer in constraining the void lensing profile as motivation to develop the methodologies required to utilize this novel signal to study cosmology or large-scale structure.

This chapter is organized as follows: Section 6.2 describes the weak lensing methodology that will be applied to the *Euclid* Flagship data described in Chapter 4 and is summarized in Table 4.2. This section will include extending the analytical covariance matrix model in Chapter 5 to lensing tomography. We additionally describe the methodology of stacking information across the different source bins in a given measurement, as well as the analytical model that we fit to the data. We present the results in Section 6.3, and discuss the future of void lensing in Section 6.4. Our fiducial cosmology in this chapter follows the cosmology of the Flagship simulation and is a flat Λ CDM cosmology with $\Omega_m = 0.319$, $\Omega_b = 0.049$, $h = 0.67$, $\sigma_8 = 0.82$, and $n_s = 0.96$. All distances are given in comoving units.

6.2 Methodology

6.2.1 Weak Lensing Methodology

The goal of this work is to forecast the expected detection significance of void lensing. As such, we opt to not modify the raw shape output of the Flagship simulations with an additional scatter due to shape noise. Performing the measurement in this manner will produce a close approximation to the true signal present in the simulation, and therefore a close approximation to the mean of the lensing signal within each radial bin. We can then use a realistic error budget that includes a shape noise component to represent its spread. This will allow us to study the distribution of possible realistic measurements and their corresponding significance. This methodology stands in contrast to performing a measurement on a specific realization of shape noise added to each galaxy shape individually, which would provide an accurate detection for only that single realization, which may or may not be a likely draw from the total distribution.

We perform the lensing measurement in a similar manner to what was presented in Chapter 5, namely Eqns. 5.3, 5.4, and 5.6. We do slightly modify the weights to treat each individual simulated source equally as the measurement uncertainties are not present in this setup. To calculate the shape noise in 5.6, we use the expected value of the shape noise per component from

the *Euclid* survey of 0.26. This shape noise measurement will be added as a diagonal component to the covariance matrix model discussed in the next section.

6.2.2 Analytical Covariance Model

The covariance due to cosmological effects is of particular importance for void lensing as opposed to similar weak lensing analyses. This is primarily due to the very large physical scale of voids as well as their relatively small population size. As such, we choose to analytically model the covariance as opposed to using jackknife or simulations. The covariance model used in this analysis follows the covariance model of Chapter 5, and we refer readers to that chapter for further details and validation tests surrounding the approximations made in this model. This model was applied to a data set without tomographic lens or source bins, and we will therefore expand the methodology to tomographic setups. To derive the model, we started with the standard Gaussian covariance term for general 2-point correlation functions, applied to the case of void positions (V) and galaxy shapes (κ) (Krause et al., 2013), converted into a covariance of $\Delta\Sigma$ by a factor of $\langle \Sigma_c^{-1} \rangle^{-2}$. We take the derivation a step further by considering the covariance between a $\Delta\Sigma$ measurement at radial bin θ_1 between a void bin i and a source bin j with a $\Delta\Sigma$ measurement at radial bin θ_2 between a void bin m and a source bin n

$$\begin{aligned} \text{Cov} [\Delta\Sigma^{ij}(\theta_1), \Delta\Sigma^{mn}(\theta_2)] &= \frac{1}{4\pi f_{\text{sky}} \langle \Sigma_c^{-1} \rangle_j \langle \Sigma_c^{-1} \rangle_m} \int \frac{\ell d\ell}{2\pi} \overline{J_2(\ell\theta_1)} \\ &\times \overline{J_2(\ell\theta_2)} \left\{ \left[C_{\text{VV}}^{im}(\ell) + \delta_{im}^{\text{K}} \frac{1}{n_{\text{V}}} \right] \left[C_{\text{KK}}^{jn}(\ell) + \delta_{jn}^{\text{K}} \frac{\sigma_{\epsilon}^2}{n_{\text{g}}} \right] \right. \\ &\quad \left. + C_{\text{VK}}^{in}(\ell) C_{\text{VK}}^{jm}(\ell) \right\}, \quad (6.1) \end{aligned}$$

where θ is defined to be the angle corresponding to a radial bin, found by multiplying R/R_{V} by the angular size of the void θ_{V} . Further, f_{sky} is the fraction of the sky area of the survey, $\widehat{J_2}$ is the 2nd order Bessel function of the first kind averaged across a radial bin, $n_{\text{V,g}}$ are the area number densities of voids and galaxies, respectively, σ_{ϵ}^2 is the value of the shape noise that we added to the simulated VIS catalogue described in Sect. 4.4.2 and δ^{K} is the Kronecker delta function. The product of the two noise terms (the terms multiplied by the Kronecker deltas) results in the measured shape noise component represented by Eqn. 5.6. The $C_{\text{AB}}(\ell)$ terms are the angular power spectra of voids, galaxy shapes, and their cross-correlation, in order. For any given tracer A and B, tomographic bins i, j and using the Limber approximation (Limber, 1953; LoVerde &

Afshordi, 2008), the angular power spectra is defined as

$$C_{AB}^{ij}(\ell) = \int d\chi \frac{W_A^i(\chi) W_B^j(\chi)}{\chi^2} P_{AB} \left(k = \frac{\ell + 1/2}{\chi} \right), \quad (6.2)$$

where $W_{A,B}$ are tracer-dependent weighting functions, and P_{AB} is the power spectrum of the two tracers. The weighting function for galaxy shapes is defined as

$$W_\kappa^i(\chi) = \bar{\rho} \int_{z(\chi)}^\infty dz_s \frac{c}{H(z_s)} \frac{n^i(z_s)}{\Sigma_c(z_s, \chi)}, \quad (6.3)$$

where Σ_c is defined as in Eqn. 5.2, without the integral over the source redshifts, $\bar{\rho}$ is the average density of the Universe, and $n^i(z_s)$ is the redshift distribution associated with the i th tomographic source bin. Couple this with the notion that galaxy shapes are an unbiased tracer of the matter distribution, and we can write the formula for $C_{\kappa\kappa}^{ij}$ for two given tomographic source bins i, j in terms of the matter power spectrum as follows

$$C_{\kappa\kappa}^{ij}(\ell) = \int d\chi \frac{W_\kappa^i(\chi) W_\kappa^j(\chi)}{\chi^2} P_{mm} \left(k = \frac{\ell + 1/2}{\chi} \right). \quad (6.4)$$

We calculate the remaining void terms in Appendix A.2, excluding small scales where the void clustering terms depend strongly on the void finder, and find that they are negligible compared to the $C_{\kappa\kappa}$ term in all cases. This simplifies the covariance formula by removing all terms that involve void clustering. Therefore, we are left with

$$\begin{aligned} \text{Cov} [\Delta\Sigma^{ij}(\theta_1), \Delta\Sigma^{mn}(\theta_2)] &= \frac{1}{4\pi f_{\text{sky}} \langle \Sigma_c^{-1} \rangle^j \langle \Sigma_c^{-1} \rangle^n} \\ &\times \int \frac{\ell d\ell}{2\pi} \widehat{J_2(\ell\theta_1)} \widehat{J_2(\ell\theta_2)} \left[C_{\kappa\kappa}^{jn}(\ell) \frac{\delta_{im}^K}{n_v} \right] + \delta_{\theta_1, \theta_2}^K \sigma_{\text{SN}, \Delta\Sigma}^2. \end{aligned} \quad (6.5)$$

This formula only depends on the tomographic void bin indices via the Kronecker delta multiplying the void shot noise term. For cases where $i \neq m$, the covariance is zero. This means that we do not consider covariance between different tomographic void bins, only tomographic source bins. This carries an additional benefit. As we find in Chapter 5, the use of scaled bins (i.e. R/R_v) in the measurement coupled with the wide spread in void sizes results in an ambiguity in terms of what set of radial bins to consider when plugging in values for θ_1 and θ_2 . This ambiguity is not present in more traditional applications of weak lensing, where a single set of radial bins is universally applied to all lenses. Because we find little covariance between tomographic void

bins, and we find negligible clustering signal within a tomographic void bin, we are able to use the methodology proposed and validated in Chapter 5, where we can compute the covariance of each void individually and average them together. To do this, given tomographic source bins i and j , we compute

$$f(\theta_1, \theta_2) = \frac{1}{\langle \Sigma_c^{-1} \rangle^i \langle \Sigma_c^{-1} \rangle^j} \int \frac{\ell d\ell}{2\pi} \widehat{J_2(\ell\theta_1)} \widehat{J_2(\ell\theta_2)} \left[\frac{C_{\kappa\kappa}^{ij}(\ell)}{1} \right] \quad (6.6)$$

for each void individually, where we brought in the $(4\pi f_{\text{sky}})^{-1}$ prefactor into the integral to turn n_v into N_v , or the total number of voids, which is then set to one. We then take the error in the weighed mean, using the same $w_L = \langle \Sigma_c^{-1} \rangle^2$ weights as in the lensing measurement and measured shape noise term in Eqns. 5.3 and 5.6

$$\text{Cov} [\Delta \Sigma^i(\theta_1), \Delta \Sigma^j(\theta_2)] = \frac{\sum_L w_L^2 f(\theta_1, \theta_2)}{(\sum_L w_L)^2} + \sigma_{\text{SN}, \Delta \Sigma}^2. \quad (6.7)$$

The end result of this process, assuming 14 radial bins as is used in this forecast, is a $14N \times 14N$ covariance matrix, where N is the number of tomographic source bins behind a given tomographic void bin, computed from an analytical Gaussian model.

6.2.3 Stacking Source Bins

The methodology outlined thus far will produce a total of 17 measurements from the four tomographic void bins and six tomographic source bins, excluding pairings where a significant portion of the sources exist in front of the voids. However, given that void lensing profiles are difficult to measure with significance, the goal of this forecast is to maximize the significance of the detection. As such, we wish to stack the signal from a given tomographic void bin along all tomographic source bins that are located behind it, producing one lensing measurement for each void bin. We therefore have four different data vectors $\mathbf{x}$, each of dimension $14N$, where N is the number of source bins behind each void bin, and we want to compress them into data vectors $\mathbf{y}$ of dimension 14 through a $14 \times 14N$ weight matrix $\mathbf{B}$

$$\mathbf{y} = \mathbf{B}\mathbf{x} \quad (6.8)$$

The optimal choice of $\mathbf{B}$ for lossless compression is then (Tegmark et al., 1997; Ruggeri et al., 2023)

$$\mathbf{B} = \mathbf{C}^{-1} \mathbf{D} \quad (6.9)$$

where $\mathbf{C}^{-1}$ is the inverse of the covariance matrix of the original data vector, and $\mathbf{D}$ is a matrix of dimension $14N \times 14$. For the purposes of stacking lensing signals, $\mathbf{D}$ is comprised of N vertically stacked 14×14 identity matrices (Blake et al., 2020). This data compression procedure then naturally provides a method to compress the $14N \times 14N$ covariance matrix (Tegmark et al., 1997)

$$\text{Cov}_y = \mathbf{B}\mathbf{C}\mathbf{B}^T. \quad (6.10)$$

An adaptation of this procedure is described in Blake et al. (2020), wherein they take the inverse of only the diagonal component of the covariance to better protect against numerical issues with negligible loss in precision. We find that we can go a single step further in this analysis. If we consider the full $14N \times 14N$ covariance matrix as being built from 14×14 block components, each representing an individual pairing of source bins, we can build a block-diagonal covariance matrix by only taking the diagonal of each block and setting all other entries to zero. We find that inverting this block-diagonal covariance matrix successfully avoids the numerical issues in building the compression weights while including more information than just the variance matrix.

We compare this stacked covariance model to a jackknife estimate in Appendix C. From that analysis, we find that on small scales there exists a discrepancy of up to 20 per cent in the closest redshift void bin (L1, see Table 4.2) between the model and the empirical estimate from jackknife out to $0.5R_v$. This discrepancy becomes negligible by L4 due to the increased shape noise. This likely indicates that the excluded small scale information from $C_{vv}(\ell)$ and $C_{vk}(\ell)$ in Appendix A.2 may not be negligible in all cases. We believe that this effect is likely due to the “one-halo” term of C_{vk} , as voids become anti-correlated with themselves on small scales due to the void exclusion principle, meaning that C_{vv} is unlikely to have an effect at these scales. Properly modelling this effect is outside the scope of this work, and will not have a significant impact in the final results, as the first three radial bins already have the largest error bars and are thus downweighted in the fits.

6.2.4 Void Density Profile Model

In order to assess the significance of the void lensing measurement, we wish to use an analytical void profile model that is fit to the data. The model we choose is the “Universal Void Profile” of Hamaus et al. (2014b, hereafter HSW)

$$\frac{\rho_v(r)}{\bar{\rho}} - 1 = \delta(r) = \delta_c \frac{1 - (r/r_{s1})^\alpha}{1 + (r/R_v)^\beta}. \quad (6.11)$$

which contains four free parameters. δ_c is the value at $r = 0$, or the central underdensity. r_{s1} is the zero-crossing scale, where the void profile crosses from underdense to its overdense wall.

α and β are power-law parameters. This model is the original four-parameter counterpart of the model used in Chapter 5. This model is an empirical model derived from simulated voids using the same VIDE void finder used to build our void catalogue, and has been widely used within the literature for both simulated and data voids (Chan et al., 2014; Barreira et al., 2015; Sánchez et al., 2016; Pollina et al., 2017; Falck et al., 2018; Baker et al., 2018; Fang et al., 2019; Shim et al., 2021; Tavasoli, 2021; Williams et al., 2025; Davies et al., 2025).

In order to fit our data, which is in terms of excess surface mass density ($\Delta\Sigma$), the model needs to be converted from its current form in terms of a unitless 3D overdensity parameter $\delta(r)$. To do so, consider the definition of $\Delta\Sigma$

$$\Delta\Sigma(R/R_v) = \overline{\Sigma(< R/R_v)} - \Sigma(R/R_v), \quad (6.12)$$

where Σ is defined as the projected matter density, and $\overline{\Sigma(< R/R_v)}$ is the average of the projected matter density inside a circle of radius R/R_v . To project the matter density, we can use an Abel transform, as we assume the void density profile averaged across many voids is spherically symmetric

$$\Sigma(R) = 2 \int_R^\infty \frac{\rho(r) r}{\sqrt{R^2 - r^2}} dr. \quad (6.13)$$

Then, noting that $\delta(r) = [\rho(r) - 1]\bar{\rho}^{-1}$, and that constant terms will be subtracted out in Eqn. 6.12, we can write

$$\Sigma(R) = 2\bar{\rho} \int_R^\infty \frac{\delta(r) r}{\sqrt{R^2 - r^2}} dr. \quad (6.14)$$

Therefore, we are able to convert Eqn. 6.11 into an analytical $\Delta\Sigma$ model that we can use to fit our measurement.

6.2.5 Model Fitting Methodology

For the models considered in this analysis, we follow a similar methodology as was done in Chapter 5. We first perform a least-squares fit by finding the set of parameter values that minimize χ^2

$$\chi^2 = \sum_{i,j} (\Delta\Sigma_{i,\text{data}} - \Delta\Sigma_{i,\text{model}}) C_{ij}^{-1} (\Delta\Sigma_{j,\text{data}} - \Delta\Sigma_{j,\text{model}}). \quad (6.15)$$

where C_{ij}^{-1} is the inverse of the covariance matrix for the corresponding measurement. We additionally perform a least-squares fit for a null model, wherein all bins are set to zero, resulting in a measurement of χ_0^2 . Then, since the null model is a nested model of the HSW model,

according to Wilks' Theorem (Wilks, 1938), the distribution of $\Delta\chi^2 = \chi_0^2 - \chi^2$ is itself a χ -squared distribution, with degrees of freedom equal to the difference of the nested models, which in this case is four. The area under the tail of the distribution given the value of $\Delta\chi^2$ from the model fit then reports the significance of the detection.

To explore the posterior around the best fit, we additionally run an MCMC analysis using the *emcee* (Foreman-Mackey et al., 2013) package. We use a set of flat uniform priors for the values of δ_c and r_{s1} . We also enforce a prior of $0 < \alpha < \beta$ to select for proper behaviour of the void profile in limiting cases. On small scales, the slope of the profile is determined by the value of α . From prior works we typically expect VIDE voids to exhibit a flat central underdensity that then curves upwards into its overdense wall. Cases where $\alpha \leq 1$ produces voids without these features. However, we find for a number of the catalogues considered here that the best fitting profiles prefer $\alpha < 1$, and as such do not enforce a prior of $\alpha > 1$ for the best fits in order to observe the fuller extent of the posterior. On the other hand, the large scale behaviour of the void profile is determined by $\alpha - \beta$, and in cases where $\alpha \geq \beta$, the profile does not properly return to zero on large scales. This results in unphysical cases where voids exhibit constant to infinite densities on very large scales, and thus need to be excluded in the prior. The MCMC analysis was run on each measurement individually with 20 walkers each taking 30 000 steps.

Lastly, because the data points represent a measurement without shape noise, we can take them to be the mean of a Gaussian distribution represented by the realistic covariance matrix that includes a shape noise component. This allows us to draw different realizations of the data vector from the covariance. Each of these data vectors is then an approximation of a realistic measurement expected from the *Euclid* survey. We can then perform a least-squares fit to each of these realizations to estimate their χ^2 distributions and detection significances. We perform this step 500 times for each catalogue.

6.3 Results

We present the void lensing measurement for the four tomographic void bins in Fig. 6.1, and the three void size bins in Fig. 6.2. The left column shows the $\Delta\Sigma$ measurements without shape noise in black points, with the least-squares HSW fit in the black line. The blue lines represent 200 individual samples from the MCMC posterior. The error bars are the square root of the diagonal of the covariance matrix with shape noise for each measurement, while the data points themselves are not scattered by shape noise. The plots on the right represent the corresponding 3D void overdensity profiles, $\delta_m(r)$, plotted by taking the corresponding parameter values from the $\Delta\Sigma$ fits and plugging them into the HSW model in Eqn. 6.11. We note that the overdensity profile from two to three r/R_v is not plotted on the right hand side to highlight the non-zero portion of the

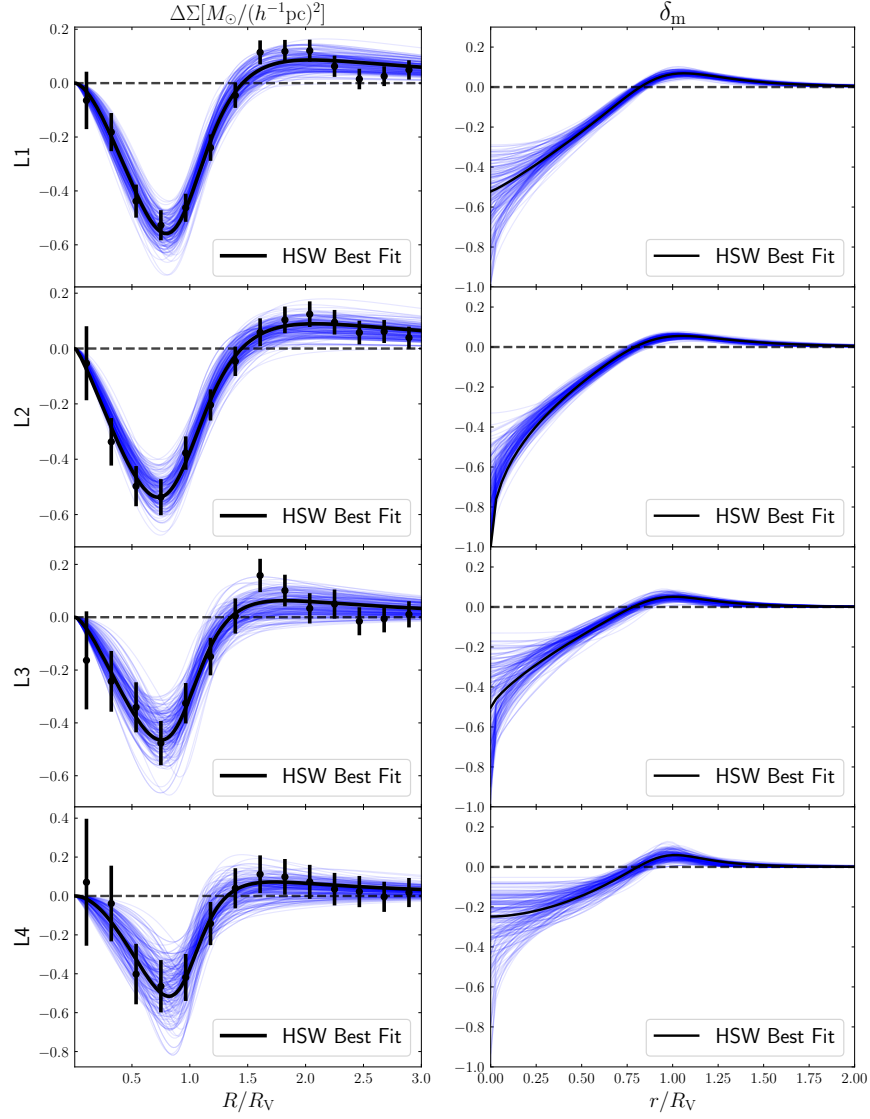


Figure 6.1: Void lensing profiles from the tomographic void samples, stacked along all corresponding source bins. The black points represent the lensing measurement from Flagship without additional scatter from shape noise, while the error bars are the square root of a covariance matrix with a realistic shape noise component added. The black line represents the least-squares model fit to these data and error bars, and the blue lines are samples from the MCMC posterior. The left column displays the measurements and model fits in $\Delta\Sigma$ units, while the right column displays the equivalent model fits in 3D underdensity δ_m . The region from two to three r/R_V is excluded in the right-hand plots to focus on the non-zero portions of the plot, unlike in the left-hand column.

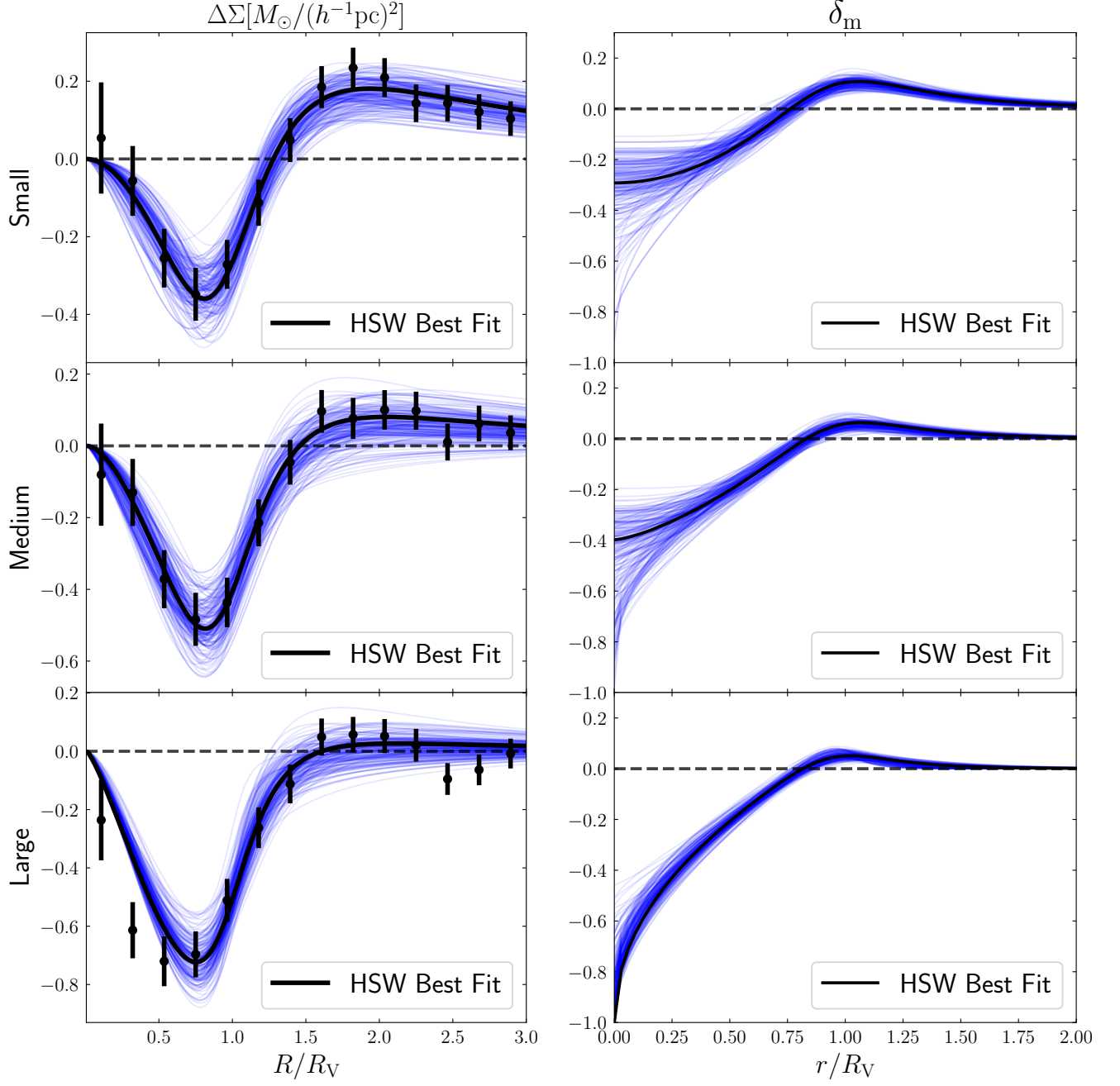


Figure 6.2: Same as Fig. 6.1, but featuring the void size catalogues.

Table 6.1: Fit statistics and best fit model parameters for all void catalogues considered here. We report the average and 1σ scatter of χ^2 and the average significance (labelled Sig.) across 500 samples from a “true” mean from the raw output of Flagship and a covariance that includes an additional realistic shape noise component not present in the data. We additionally report the χ^2_{true} for the fit of the HSW model to the “true” mean with realistic covariance for comparison. It is expected that the model overfits the data in these cases. The spread in the significance across the 500 samples is $\pm 1\sigma$ by construction. We additionally report the posterior means and 1σ errors from an MCMC posterior for the four HSW parameters found by fitting to the “true” lensing signal shown in Figs. 6.1 and 6.2. For L2 and Large, which feature skewed distributions in δ_c , we quote the 95 per cent upper bound, with the best-fit value in parentheses.

Catalogue	χ^2/dof	$\chi^2_{\text{true}}/\text{dof}$	Sig.	δ_c	r_{s1}/R_v	α	β
L1	$(20 \pm 7)/10$	9.8/10	12σ	$-0.52^{+0.15}_{-0.30}$	$0.82^{+0.04}_{-0.03}$	$1.13^{+0.62}_{-0.51}$	$7.4^{+1.5}_{-1.0}$
L2	$(15 \pm 6)/10$	3.2/10	11σ	$< -0.38 (-1.0)$	$0.79^{+0.05}_{-0.03}$	$0.43^{+0.51}_{-0.01}$	$6.5^{+2.1}_{-0.7}$
L3	$(17 \pm 7)/10$	7.2/10	7.1σ	$-0.50^{+0.27}_{-0.36}$	$0.78^{+0.06}_{-0.04}$	$0.75^{+1.2}_{-0.38}$	$8.4^{+5.9}_{-1.4}$
L4	$(12 \pm 5)/10$	1.5/10	4.7σ	$-0.25^{+0.12}_{-0.45}$	$0.81^{+0.08}_{-0.06}$	$1.8^{+2.5}_{-1.3}$	10^{+8}_{-2}
Small	$(15 \pm 7)/10$	4.7/10	8.3σ	$-0.29^{+0.11}_{-0.23}$	$0.76^{+0.05}_{-0.05}$	$2.0^{+1.1}_{-1.0}$	$6.9^{+1.4}_{-1.3}$
Medium	$(16 \pm 6)/10$	5.5/10	8.2σ	$-0.40^{+0.13}_{-0.32}$	$0.83^{+0.05}_{-0.04}$	$1.4^{+0.84}_{-0.77}$	$7.6^{+2.4}_{-1.4}$
Large	$(31 \pm 11)/10$	20/10	11σ	$< -0.56 (-1.0)$	$0.82^{+0.05}_{-0.03}$	$0.48^{+0.25}_{-0.03}$	$8.6^{+7.1}_{-1.2}$

profile, but is plotted in the left column. The best fitting parameter values are given in Table 6.1. We additionally report the mean and standard deviation for the χ^2 of the 500 data vector samples as described in the previous section. For the detection significance, which is the area under the tail of a χ^2 distribution, we report as the equivalent area under both tails of a Gaussian distribution for convenience. We note that by construction, the spread in the significance detection across the 500 data vector samples will always be $\pm 1\sigma$ since the significance is defined relative to the covariance used to draw the samples. The quality of the fits observed in the figures are given by χ^2_{true} in Table 6.1. Since the covariance includes the shape noise scatter that is not present in the data, the expectation is for the model to overfit the data, and we report this statistic to compare to the average χ^2 values across the 500 realistic samples.

For all plots, we observe underdensities, with minimum $\Delta\Sigma$ values occurring just short of the void radius, as expected. We forecast the most significant detection of the void lensing profile will occur within the closest tomographic void bin at the 12σ level. Subsequent bins report less significant detections, as expected from fewer background tomographic source bins, with levels at 11σ , 7.1σ , and 4.7σ . For the void size catalogues, we detect at a 8.3σ , 8.2σ , and 11σ from Small to Large, respectively. We do however find that, on average, the HSW fits to L1 and Large are rejected. For L1, this is primarily driven by the fits to the outermost bins, while it is driven by the innermost bins for Large. This behaviour follows the same trend as the χ^2_{true} values, where L1 and especially Large exhibit relatively worse fits than other catalogues. In the case of Large, the fit to the unscattered shapes is rejected ($p < 0.05$). This indicates that the poor fits are a product of the fit to the data without shape noise, rather than the realistically-scattered samples.

From the corresponding 3D overdensity plots, we do find that for L2, L3 and Large, the overdensity profiles exhibit unexpected small-scale behaviour, wherein the profiles steeply rise from their central values at $r = 0$ instead of starting out relatively flat as is expected from most void profile studies in simulations and observations. This behaviour is observed in the HSW model fits when $\alpha < 1$, however we do not restrict it for illustrative and comparative purposes. We find that this behaviour in the inner region is driven by the first four bins in $\Delta\Sigma$. This means that the fit to the α parameter is sensitive to the noisiest bins. We note in particular for Large that the model does not seem to capture the depth of the $\Delta\Sigma$ plot in the second and third bins, producing very poor values for both χ^2 across the 500 realistic sample measurements and χ^2_{true} . We additionally find a degeneracy between α and δ_c when $\alpha < 1$ that causes δ_c to drop to the lowest value at the prior bound when α is sufficiently less than one. This effect is most clearly seen in the right-hand plots for L2 and Large. Extending the δ_c prior bound to unphysical values results in a convergence for L2 at around $\delta_c \approx -2$, while Large does not show convergence, even at extreme values. This seems to indicate an inflexibility of the HSW model when fit to specifically $\Delta\Sigma$ profiles whose innermost slope is too steep, whether by physical effects or by statistical fluctuations.

Comparing the tomographic measurements by themselves, we find little evolution of the void profile with redshift. All best fit values for the zero-crossing scale, r_{s_1} as well as β are consistent with each other for all four tomographic bins. Visually, aside from the small scale behaviour, we do not observe significant differences in the best fitting profiles. On the other hand, we do observe some visual differences between the smallest and largest voids in Fig. 6.2. Visually, we find that the overdensity profiles for the Small catalogue exhibit stronger compensation walls at $r = R_v$ than the Large void catalogue. This is in agreement with expectations that smaller voids are more likely to be “void-in-clouds,” where a small-scale underdensity exists within a larger-scale overdense region. We additionally find that the Large catalogue strongly prefers fits with $\alpha < 1$, while the same is not true for Small and Medium. As mentioned previously, this is driven by the behaviour of the first three to four bins of the measurement, where in particular the minimum values are significantly more negative in Large (≈ -0.7) compared to the other two catalogues at around values of -0.3 to -0.5 . We find that the difference between the Small and Large catalogues is significant at 4.4σ .

We additionally provide the corner plots for both sets of catalogues shown in Fig. 6.3 and Fig. 6.4. For the tomographic catalogues, we do not find significant change in the parameter space with redshift. We do find that L3 and L4 have extended tails into high values of β and low values of δ_c . We additionally find that L2 prefers low values of δ_c due to the low- α degeneracy, remarkably more so than L3. On the other hand, Fig. 6.4 does show a more significant change in parameters with void size. This is demonstrated in comparing Small to Large in the δ_c and α parameters. As mentioned prior, the larger void catalogues prefer α fits to be smaller than 1, while smaller void catalogues prefer $\alpha > 1$. Given the banana-shaped contour in the δ_c - α plane, this then directly implies that larger void catalogues would prefer much lower values of δ_c than smaller void catalogues. We do additionally find that void size has little effect on r_{s_1} and β , beyond a slight preference for smaller voids to have a lower r_{s_1} , and larger voids having an extended tail to high β .

6.3.1 Integrated Density Contrast

Due to the uncertainty of the measurement close to the centre of the void and its effects on the model fits, it is difficult to directly assess the depth of the central underdensity and gauge how “empty” these voids are. Instead, we follow the methodology presented in Chapter 5 and examine the integrated density contrast $\Delta(r)$. This quantity is defined as the normalized integral over the 3D overdensity profile $\delta_m(r)$

$$\Delta(r) = \frac{3}{r^3} \int_0^r \delta(y) y^2 dy. \quad (6.16)$$

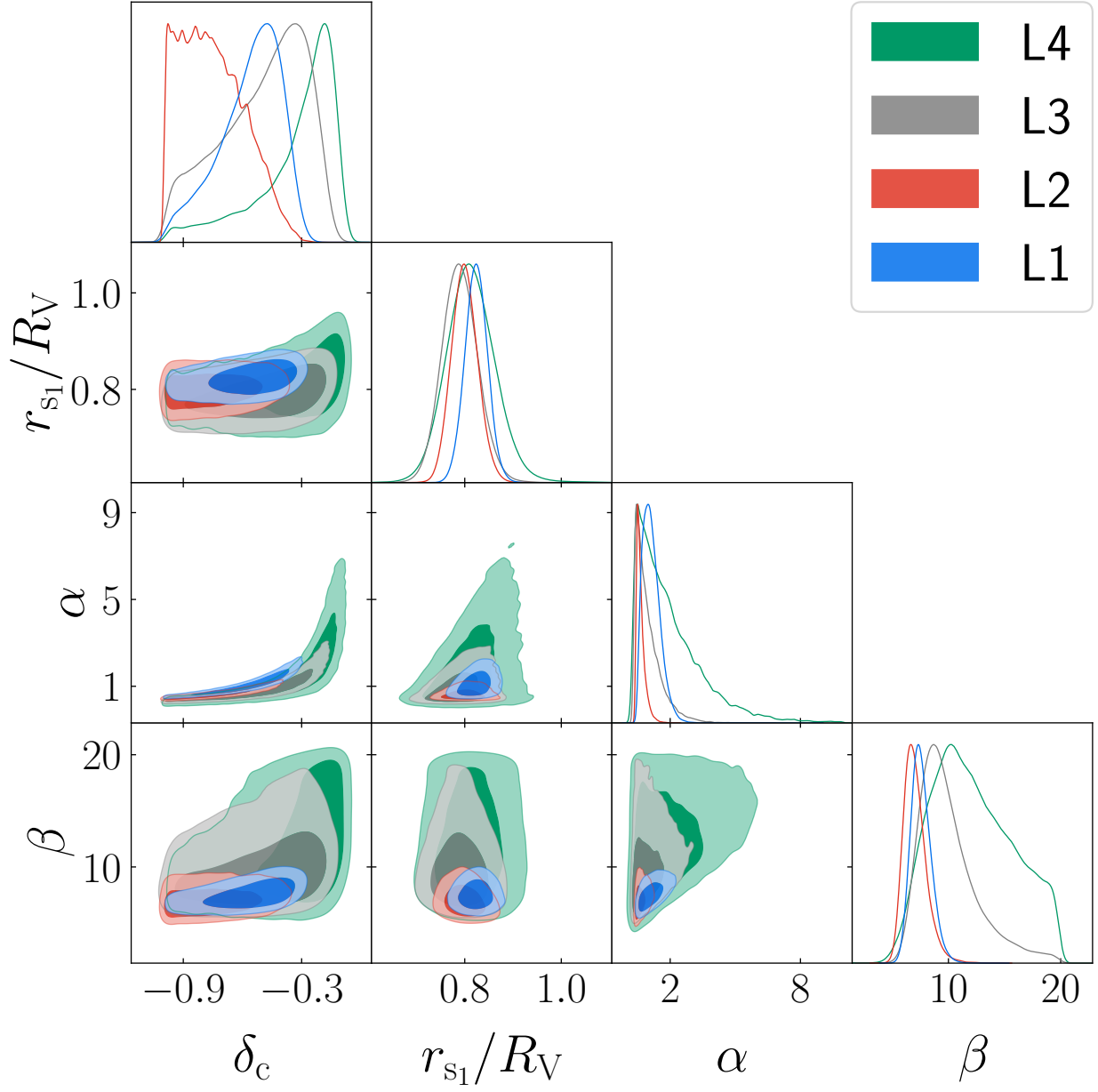


Figure 6.3: Corner plot of the MCMC posterior of the HSW model fits for the tomographic catalogues.

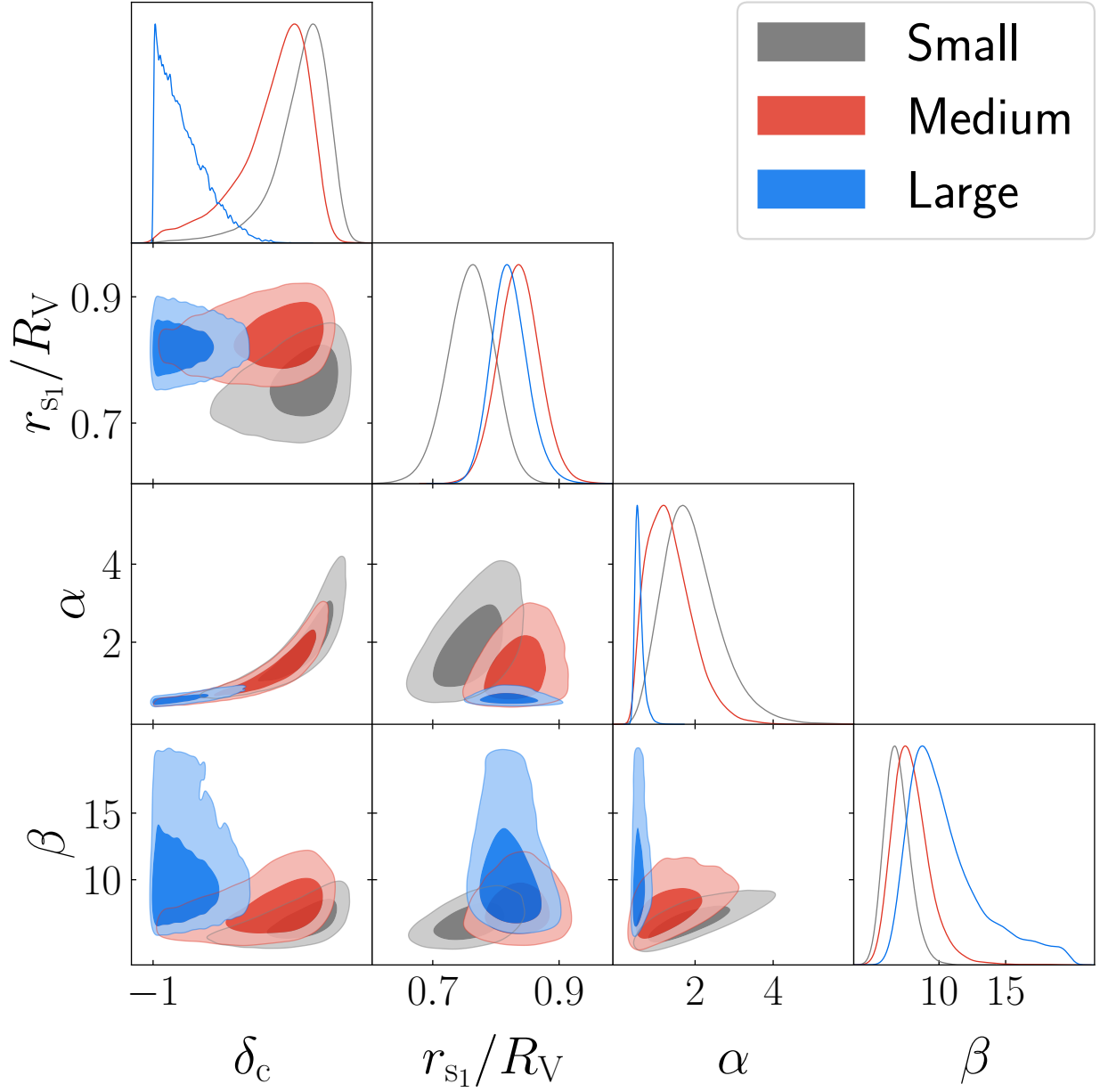


Figure 6.4: Same as Fig. 6.3, but featuring the void size catalogues.

Table 6.2: Integrated density contrasts at three different locations along the void profiles. The error bars are the 1σ scatter taken from individual MCMC samples

Cat.	$\Delta(r = 0.5R_v)$	$\Delta(r = 1.0R_v)$	$\Delta(r = 3.0R_v)$
L1	-0.30 ± 0.03	-0.059 ± 0.012	0.0050 ± 0.0017
L2	-0.29 ± 0.03	-0.044 ± 0.011	0.0050 ± 0.0019
L3	-0.21 ± 0.03	-0.029 ± 0.013	0.0022 ± 0.0014
L4	-0.18 ± 0.05	-0.031 ± 0.016	0.0020 ± 0.0015
Small	-0.22 ± 0.04	-0.013 ± 0.018	0.0143 ± 0.0037
Medium	-0.27 ± 0.04	-0.055 ± 0.015	0.0047 ± 0.0025
Large	-0.32 ± 0.03	-0.058 ± 0.012	0.0011 ± 0.0014

We examine Δ computed at three different locations: half the void radius, at the void radius, and at three times the void radius. The half-radius calculation will provide a more precise assessment of the depth of the central underdensity, while the full-radius and 3 times the radius calculations will provide information about the compensation of the void, or how well the overdense wall could “fill in” the void. An integrated density contrast consistent with zero indicates that a void is compensated at that scale, while a negative or positive value indicates under- or over-compensated, respectively. To perform this calculation, we plug in the overdensity profile corresponding to the best fit HSW model fit to the $\Delta\Sigma$ measurements into Eqn. 6.16. We then compute the error as the standard deviation of the scatter when computing the same quantity for MCMC samples. The results are shown in Table 6.2.

Considering the half-radius measurements of the tomographic catalogues first, we find consistency between L1 and L2, and consistency between L3 and L4, with a small discrepancy between the two groupings. It is expected that higher redshift voids are shallower than lower redshift voids, as gravity has not had as much time to empty higher redshift voids out, so this behaviour is expected. For the void size catalogues, we find a similarly weakly significant relationship between void size and void underdensity. Large is deeper than Small at the 1σ level, with Medium bridging the gap. This general trend is in agreement with trends observed in central underdensities from non-merged voids with void size, from both simulations (Nadathur & Hotchkiss, 2015b; Schuster et al., 2023) and data (Fang et al., 2019). Considering the measurements at the void radius, we find a similar trend as the half-radius measurement for the tomographic catalogues. For the void size catalogues, we observe that Small is significantly different from both Medium and Large now, and is consistent with being compensated at this scale, while Medium and Large are still consistent with each other. At $3R_v$ for the tomographic catalogues, we find that all catalogues are significant to marginally significantly larger than zero, indicating that these voids are marginally overcompensated and on average exist within slightly overdense regions on larger scales. Look-

ing to the void size catalogues, we can observe that this overdensity primarily originates from voids in the Small void catalogues, as their integrated density contrast is significantly larger than zero at these scales, while Medium is marginally larger than zero, and Large is consistent with compensated. This relationship shows a clear trend of void compensation with void size that is detectable from a *Euclid* DR2-like survey.

6.4 Discussion

From analyzing the weak lensing signal around voids in the full octant of the Flagship simulation, we determine the forecasted significance of detecting the void lensing profile in the closest redshift bin to be on average around 12σ . Given that the depth of the *Euclid* survey is expected to remain relatively unchanged between data releases, to a first approximation we can scale the significance of these results by the change in the area. Since the area of Flagship is roughly the expected area of the second data release, which is around one-third the size of the expected full *Euclid* survey, we expect the detection significance to scale by $\sqrt{3}$, resulting in a significance of 21σ , with a secondary detection in the second redshift bin of 19σ . At the time of writing, the expected area of the first data release is around 1900 deg^2 , which scales the significance of L1 to 6.92σ . The most significant detection of void lensing of 3D voids identified in spectroscopic tracers is reported in Chapter 5, where we find a detection of 6.2σ from voids in the BOSS survey. Therefore, we forecast that both DR2 and DR3 of the *Euclid* survey may represent significant improvements in measuring the void lensing profile. We additionally find a SNR of 11.8 for the L1 catalogue to compare to the 14.0 SNR measured from the ZOBOV voids in Fang et al. (2019). This indicates that the line-of-sight smearing caused by the use of photometric redshifts in void finding produces a very pronounced projected underdensity. For cases where isotropy is not needed, such as modified gravity or neutrino studies, projected underdensities might produce stronger signals.

However, a number of systematics were not modelled in this analysis that may have a significant impact on the measurements. Primarily, this analysis was done on a simple, contiguous octant of the Flagship simulation. Given that ZOBOV algorithms operate by joining bordering underdense cells together with a watershed algorithm, the presence of holes in the survey mask can split a single void into multiple smaller voids. This creates a selection bias towards smaller voids whose boundaries artificially exhibit properties of larger voids. Alternatively, voids sufficiently close to survey boundaries can be removed, but doing so would decrease the amount and quality of the void data, thereby the significance of the detection would degrade with increasing complexity of the survey mask. Therefore, these forecasts should be taken as optimistic.

In Chapter 5, we used a void catalogue found using *Revolver* to measure the void lensing

profiles of BOSS with source data from the UNIONS survey. Because both *Revolver* and *VIDE* are based on the ZOBOV algorithm, the main relevant difference between the two catalogues is their definition of the void centre. *Revolver* uses a circumcentre definition, wherein the center of a void is defined as the center of the largest empty circumscribed sphere, as opposed to the barycentre definition used in *VIDE*. Using the circumcentre definition, the authors found that fitting the HSW model to their void lensing measurements resulted in $\alpha < 1$ fits for each of their catalogues. In response, we attempted a variation of the model proposed by [Barreira et al. \(2015\)](#), wherein the void radius in the denominator of Eqn. 6.11 becomes a free parameter. This was done with the justification that the HSW model was empirically found using barycentre voids. This modification to the model only marginally improved the fits of that chapter, but the resulting models all fit $\alpha \gtrsim 1$. Given that a few of the fits in this chapter also prefer $\alpha < 1$, but with barycentre voids, it is then interesting to perform the same fits, but while freeing the void radius parameter, in line with the modification proposed by [Barreira et al. \(2015\)](#). Since these two models are nested, we can then perform the same Wilks' Theorem analysis to determine if the difference between the fits is significant. Performing this analysis, we do not find a sufficiently significant change in the goodness-of-fit across all seven catalogues, nor do we find in cases where $\alpha < 1$ that the modification produces fits with $\alpha \gtrsim 1$, unlike in the case of Chapter 5. As expected, the HSW model performs well in modelling the void profiles of *VIDE* voids.

In regards to future work, the detection significance of void lensing profiles in this forecast is remarkable given that the high redshift of the NISP catalogue results in a methodology where the expected peak of the VIS source catalogue is at a lower redshift than the spectroscopic void lenses. This unavoidably reduces the expected source density quite significantly from 28.8 arcmin^{-2} to 4.2 arcmin^{-2} . However, in the future, spectroscopic surveys like the Dark Energy Spectroscopic Instrument (DESI) survey ([DESI Collaboration et al., 2016](#)) will offer a low redshift galaxy sample for 3D ZOBOV voids that overlaps with the full DR3 *Euclid* VIS imaging survey. These lower redshift void catalogues can then take advantage of the large galaxy density to obtain tighter constraints. As such, it becomes important to develop strategies to extract useful information from void lensing profiles, which trace the underlying matter distribution in underdense environments. One such strategy was explored in [Fang et al. \(2019\)](#) and Chapter 5, wherein they used void-galaxy profile measurements with a free bias parameter as a model to fit to the void lensing profile. Doing so results in an estimate of the clustering bias of galaxies with respect to matter relative to underdense environments. As explored in simulations by [Pollina et al. \(2017\)](#), the value of this bias is stronger than the large-scale galaxy bias within small voids, but equal in mid-to-large voids.

Additionally, void matter profiles are particularly sensitive to modified gravity, which acts to create deeper voids, and neutrino mass and warm dark matter, which act to free stream matter back into the central underdensity. However, extracting this information from the void lensing

profile is not simple, as the profile strongly depends on the specific void finder and tracer sample chosen. This complicates analytic modelling of the profile from first principles. An alternative would be to use an emulator trained on a suite of cosmological simulations. This is a technique that is seeing preliminary success in void-galaxy profile and void size function contexts (Fraser et al., 2025; Lehman et al., 2025). The essential idea is that by training a machine learning model on simulated voids identified in a consistent manner across different cosmologies, the emulator will train on changes in the void profile with cosmology, providing us with a cosmological model that can be fit to a given void lensing profile from data. A particularly strong motivation to perform this analysis is the work of (Massara et al., 2015), who demonstrates that the largest voids are expected to have inner regions that deviate on the order of 5 to 20 per cent in 0.6 eV cosmologies. From Table 6.2, we find integrated density values out to half the void radius that are constrained to 10 per cent errors in Large, and 13 per cent errors in Medium, suggesting that it may be possible to constrain the upper bound on neutrino mass with *Euclid* void lensing. Future work will be needed to quantify this more accurately, however.

6.5 Conclusions

Euclid will be a promising survey for the detection of the void lensing profile and its potential future as a cosmological probe. We optimistically forecast here that the combination of NISP-like and VIS-like surveys of v1.8.4 of the *Euclid* Flagship simulation are expected to detect the void lensing profile at 6.9σ using the lowest-redshift quarter of the void catalogue at the first of three data releases, corresponding to 1900 deg^2 . This detection signal increases to around 12σ then 21σ at the second and third data releases of *Euclid* survey, which can be directly compared to the 6.2σ detection of Chapter 5 with voids in the BOSS survey. In performing this forecast measurement, we extend the analytical covariance model in Chapter 5 to a lensing tomography setup, and stack the information along different source bins to increase the constraining power of the measurement. We find little consistent change in the void lensing profiles with redshift, but do find significant change in the void lensing profiles with void size. Small voids are shallower and more overcompensated than large voids, consistent with expectations and forecasted to be detectable at a DR2-sized survey at 4.4σ .

Chapter 7

Conclusions and Future Work

The future of cosmology depends on the construction of deeper and wider surveys to gather more and higher quality data to measure the cosmological parameters with greater and greater precision. The future of cosmology also depends on the development of new tools and probes to extract more information from pre-existing data sets. Cosmic voids, regions of the night sky that are underdense in luminous tracers, are an ideal candidate to complement the cosmological information that we have obtained from those luminous tracers. The goal of this thesis is to measure a new cosmic void probe: void matter profiles as measured by weak gravitational lensing. By demonstrating that this profile can be measured with sufficient significance in both current and future data sets, our goal is to justify future efforts to develop the tools needed to extract the cosmological information from this new signal.

In Chapter 5, we present the methodology and results of measuring the void lensing signal from BOSS voids and UNIONS sources. This measurement is the most significant detection of void lensing from spectroscopic ZOBOV voids to date at 6.2σ significance from the Full catalogue. The change in void profiles as a function of void radius was also detectable at 2.3σ . We additionally present and validate an analytical Gaussian covariance model for the first time in void lensing analyses. We also demonstrate the need for mock simulated voids over a uniform random sample to measure lingering additive systematics that need to be subtracted from the lensing signal. We find that the empirical HSW model for the void profiles produces better small-scale interpretations of circumcentre voids when modified with the B+15 modification. This does come at a cost of a complex posterior in the model fits. We lastly measure the galaxy bias relative to void environments, finding $b_g = 2.45 \pm 0.36$, consistent with the large-scale galaxy bias measurement of Sugiyama et al. (2022), but with large enough scatter to be unable to detect any trend with void size as expected in (Pollina et al., 2017). We do find that the linear bias model is not rejected when comparing void lensing measurements with void-galaxy cross-correlations.

In Chapter 6, we extend the methodology of the previous chapter to a tomographic lensing analysis, wherein photometric redshifts allows for the construction of line of sight bins. These bins then allow for a better weighting scheme and the exclusion of foreground galaxy shapes that would weaken the signal. In this chapter, we forecast a *Euclid* detection of the void lensing profile of 6.9σ at the time of the first data release in the closest void redshift bin, then 12σ and 21σ for the second and final releases. We find that unlike in Chapter 5, the unmodified HSW model performs well in fitting to barycentre voids, though the model seems unable to fit some behaviours of $\Delta\Sigma$ profiles, as can be seen in the Large catalogue fits. We additionally validate the tomographic version of the covariance model with a jackknife estimate. We do however find a small-scale discrepancy likely due to the assumed-negligible $C_{V\kappa}$ term, which would require additional simulated modelling of the void profiles at the particle level to account for. We do find that the integrated density measurements for *Euclid* DR2 may constrain the upper bound of neutrino mass, but future work needs to be performed to properly quantify this.

7.1 Future Work

As mentioned in Chapters 5 and 6, additional future galaxy surveys such as DESI, LSST, and Roman will offer more coverage of the night sky in both galaxy imaging for weak lensing and spectroscopic galaxy tracers for void finding. There are typically a few different ways weak lensing constraints can be improved at the survey level, namely by increasing the survey area, the number density of source galaxies, the number density of lens objects, and the depth of the source galaxies relative to the lenses. The increased relative depth of sources is important as it increases the strength of the lensing effect per source galaxy by improving $\langle\Sigma_c^{-1}\rangle$. In particular, DESI will offer significant overlap with UNIONS, as well as later releases of the *Euclid* survey. The lower redshift range of DESI when compared to the *Euclid* NISP survey will also allow for better void lensing constraints using VIS sources. The high redshift range of NISP leaves the peak of the *Euclid* source distribution in the foreground, significantly reducing the expected galaxy shape number density for NISP void lensing. Further, the eventual overlap between *Euclid* VIS and BOSS will also tighten constraints compared to the UNIONS survey due to the higher redshift peak of the *Euclid* survey compared to UNIONS, as well as the greatly increased number density of sources, from 7.3 arcmin^{-2} to an expected 28.8 arcmin^{-2} . The release of future UNIONS data sets, with more area coverage and eventual photometric redshifts, will also significantly improve constraints on the matter profiles.

One of the principal reasons for measuring and studying the void lensing profile is to exploit the signal for its cosmological information. Other void probes have been found to offer complementary information to other cosmological probes, both underdense and overdense, in the form of

breaking parameter degeneracies (see, for example, Fig 7 in [Pelliciari et al., 2023](#), for simulated void size function and halo mass function). For void lensing in particular, [Salcedo et al. \(2025\)](#) finds from a Fisher forecast that the combination of void lensing from ZOBOV voids with void size function measurements produce moderately tighter constraints on Ω_m when compared to void size function alone and 4.5 times improvement on the constraints of σ_8 , using forecasting numbers from Stage-III surveys like UNIONS or DESI. Stage-IV surveys like *Euclid* will provide even tighter constraints.

As stated previously, void statistics are also sensitive to modified gravity models due to their insensitivity to screening mechanisms that are needed to reproduce GR gravity on solar system scales. In particular for void lensing, certain modified gravity models, such as Galileon modified gravity ([Nicolis et al., 2009](#)) have an additional contribution to the gravitational lensing potential that is unique from the contribution to the matter distribution. Therefore, void lensing will be able to provide unique constraints on different variants of these models ([Barreira et al., 2015, 2017](#); [Baker et al., 2018](#); [Su et al., 2023](#)). Further, [Su et al. \(2023\)](#) optimistically forecasts the expected SNR of the overlapping region of *Euclid* VIS (optimistically assuming a 30 arcsec^2 galaxy shape density which was not possible with NISP voids) and DESI, as well as the overlap of LSST and DESI to suggest that the difference between Cubic Galileon gravity and standard GR gravity will be detectable with high SNR (~ 40) for three-dimensional voids.

However, we note that neither the results in Chapter 5 nor the results in Chapter 6 for *Euclid* Data Release 2 in measuring the void profile were able to exceed the SNR of the photometric voids in [Fang et al. \(2019\)](#). Therefore, it is likely that two-dimensional projected voids or ZOBOV voids identified in photometric galaxy surveys like UNIONS or *Euclid* VIS alone may produce stronger lensing signals than what is presented in this thesis, at the cost of isotropic assumptions. We further find from the results of [Cautun et al. \(2018\)](#) and [Paillas et al. \(2019\)](#) that three-dimensional voids in some modified gravity models show very weak constraints of the deviation from GR gravity. Projected voids, such as tunnels or troughs, on the other hand, exhibit typical $\Delta\Sigma$ profiles with deeper minima, and this deeper minima results in stronger detections of the deviations of modified gravity from GR. Examples of $\Delta\Sigma$ profiles for a specific model of modified gravity, normal-branch Dvali–Gabadadze–Porrati braneworld ([Dvali et al., 2000](#)), compared to GR are shown in Figs. 7.1 for the 3D void finders and 7.2 for projected void finders.

We therefore find two different void classes from which we can extract useful cosmological information from their lensing signal: three-dimensional voids and projected voids. The first class corresponds to the spectroscopic ZOBOV voids presented in this thesis, and maybe useful when used in combination with other void probes where the spherical shape assumption is necessary. As an example, some cosmological models for the void-galaxy cross-correlation require an input of the integrated matter profile, such as the template-fitting models of [Nadathur & Percival \(2018\)](#), which were used in [Nadathur et al. \(2020\)](#) and [Woodfinden et al. \(2022\)](#). For those analyses,

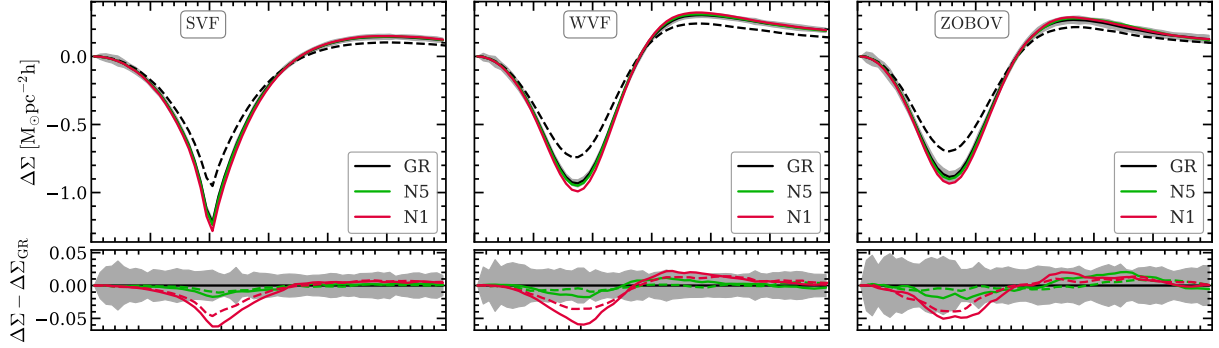


Figure 7.1: $\Delta\Sigma$ plots from [Paillas et al. \(2019\)](#) demonstrating deviations of normal-branch Dvali–Gabadadze–Porrati braneworld gravity models from GR gravity in the three void finders pictured in Fig. 3.2. The black solid line indicates the profile at $z = 0$, and the dashed line indicates the profile at $z = 1$. The green and red lines indicate different strengths of the same nDGP gravity model. Image Credit: [Paillas et al. \(2019\)](#)

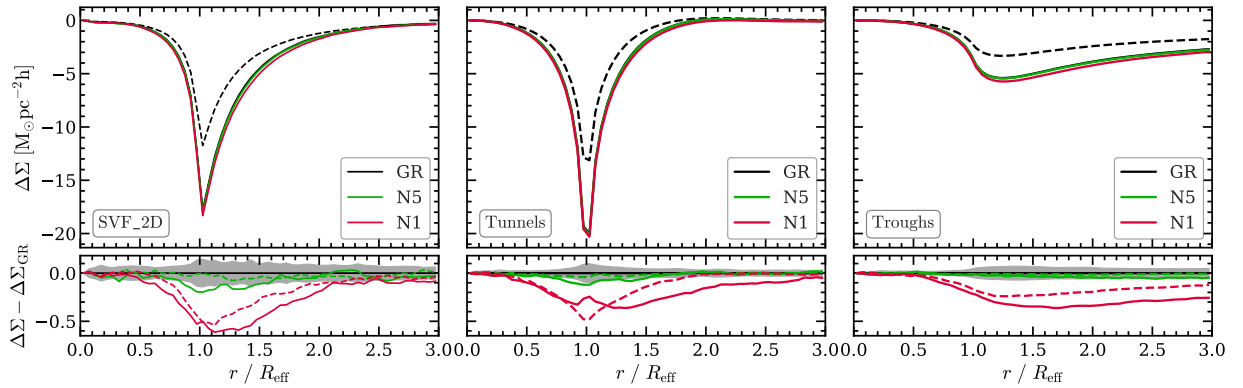


Figure 7.2: Same as Fig. 7.1, but for the two-dimensional void finders considered in [Paillas et al. \(2019\)](#). Image Credit: [Paillas et al. \(2019\)](#)

templates of the matter profile were generated from simulations. Void lensing results will probe the matter profile directly, and, while they may be too noisy to use directly, will provide useful priors to base the templates on. Further, if the galaxy bias is found to be environment dependent as predicted from [Pollina et al. \(2017\)](#), the combination of void matter profiles from three-dimensional void lensing and void-galaxy cross-correlations would be useful to constrain galaxy bias as a nuisance parameter. Furthermore, [Bromley & Geller \(2025\)](#) demonstrate a relationship between the matter outflow rate and the three-dimensional void depth, which is sensitive to the values of H_0 , Ω_m and Ω_Λ . From the forecasts of [Salcedo et al. \(2025\)](#), it is also expected that void lensing from these profiles will be complementary when used in combination with other void probes to produce tighter constraints, while the optimistic forecasts of [Su et al. \(2023\)](#) suggest that these voids will still be able to potentially constrain modified gravity models. Additionally, Stage-III survey simulation results presented in [Xu et al. \(2026\)](#) three-dimensional voids suggests that neutrino mass constraints might be possible with three-dimensional voids using Stage-IV imaging surveys. These voids will also be more closely related to the idealized spherical voids typically considered in void models. The other class will be projected voids, like the 2D voids in [Fang et al. \(2019\)](#) or the troughs from [Gruen et al. \(2015\)](#). These voids are expected to carry stronger lensing signals and modified gravity constraints, and can be applied to photometric imaging surveys like UNIONS or *Euclid* VIS directly.

Previous attempts to extract cosmological information out of void probes has relied on models built from idealized voids. In the case of the void size function, connecting observed voids to these models requires detailed cleaning of the data set ([Ronconi et al., 2019](#)) to bring observed voids into closer agreement with assumptions made in the idealized case. Since the overall shape of the void matter distribution is heavily dependent on the void finder, which can be seen in part in the discussions of circumcentre and barycentre voids throughout this thesis, and given the void profile's dependence on properties of the tracer distribution such as the galaxy bias measured in Chapter 5, it is therefore unlikely that an analytical model of the void profile derived from first principles will adequately fit observed voids. However, emulators and other machine learning methods, such as implicit likelihood inference used in [Thiele et al. \(2024\)](#), can be used on simulation suites like *AbacusSummit* and *Quijote*. These simulation suites are constructed of a multitude of different individual simulation boxes, each varying either initial conditions or cosmological parameters. This methodology would in principle allow for a machine learning model to train on void lensing profiles found using the same void finder and similar tracer properties, such as completeness, survey footprint, bias, and sparsity, with only the cosmological parameters of interest changing. Both *AbacusSummit* and *Quijote* have simulation boxes with different neutrino masses, allowing for neutrino mass constraints as well. This is particularly notable given that the constraints found for integrated density in Chapter 6 are somewhat smaller than the simulated deviations for the heaviest neutrino cosmologies (0.6 eV) tested in ([Massara](#)

et al., 2015), suggesting potential constraints on the upper bound of neutrino masses. This emulator methodology has seen some initial success for void-galaxy cross-correlation models in Fraser et al. (2025), and a void emulator for void density profiles of VIDE voids has been built in Lehman et al. (2025), which was able to reproduce the cosmology of a separate simulation. Future examination and study will likely be required to apply this methodology to data, as well as to apply this methodology to projected voids, as they may present challenges unique from their ZOBOV counterparts.

In conclusion, we find that the lensing signal around cosmic voids identified with spectroscopic tracers currently is, and will be in the future, significantly detectable from galaxy surveys. Through the use of machine learning models and simulation suites, this novel probe now has the capability to open new avenues into investigating the distribution of matter within the Universe, as well as helping to answer some of the important lingering questions with our standard cosmological model, such as the nature of the neutrino mass hierarchy or deviations from General Relativity. With the advent of next generation surveys such as *Euclid* promising wider and deeper catalogues usable for both void finding and weak lensing, the future of exploring the darkest parts of the Universe is very bright.

References

- Abazajian K. N., et al., 2011, [Astroparticle Physics](#), **35**, 177
- Adame A. G., et al., 2025, [JCAP](#), **2025**, 021
- Adey D., et al., 2018, [PRL](#), **121**, 241805
- Agarwal S., Feldman H. A., 2011, [MNRAS](#), **410**, 1647
- Aharmim B., et al., 2013, [Phys. Rev. C](#), **88**, 025501
- Alam S., et al., 2017, [MNRAS](#), **470**, 2617
- Alam S., et al., 2021, [JCAP](#), **2021**, 050
- Albrecht A., et al., 2006, [arXiv e-prints](#), pp astro-ph/0609591
- Amendola L., Frieman J. A., Waga I., 1999, [MNRAS](#), **309**, 465
- Arsenov N., Kovács A., Pérez Sar M., Bogdán Á., Sinigaglia F., Kitaura F.-S., Favole G., Slavcheva-Mihova L., 2025, [arXiv e-prints](#), p. [arXiv:2509.17696](#)
- Aubert M., et al., 2022, [MNRAS](#), **513**, 186
- Bahcall N. A., Lubin L. M., Dorman V., 1995, [ApJL](#), **447**, L81
- Baker T., Clampitt J., Jain B., Trodden M., 2018, [Phys. Rev. D](#), **98**, 023511
- Barreira A., Cautun M., Li B., Baugh C. M., Pascoli S., 2015, [JCAP](#), **08**, 028
- Barreira A., Bose S., Li B., Llinares C., 2017, [JCAP](#), **02**, 031
- Bartelmann M., Schneider P., 2001, [Phys. Rep.](#), **340**, 291

Bayer A. E., et al., 2021, [ApJ](#), 919, 24

Becker M. R., et al., 2016, [Phys. Rev. D](#), 94, 022002

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, [ApJ](#), 762, 109

Benítez N., 2000, [ApJ](#), 536, 571

de Bernardis P., et al., 2000, [Nature](#), 404, 955

Blake C., et al., 2020, [AAP](#), 642, A158

Blanton M. R., et al., 2003, [ApJ](#), 592, 819

Blanton M. R., Lupton R. H., Schlegel D. J., Strauss M. A., Brinkmann J., Fukugita M., Loveday J., 2005, [ApJ](#), 631, 208

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, [ApJ](#), 379, 440

Bonici M., Carbone C., Davini S., et al., 2023, [AAP](#), 670, A47

Brandenberger R., Mukhanov V., Sornborger A., 1993, [PRD](#), 48, 1629

Bromley B. C., Geller M. J., 2025, [JCAP](#), 2025, 011

Brouwer M. M., et al., 2018, [MNRAS](#), 481, 5189

Burger P. A., et al., 2023, [A&A](#), 669, A69

Cai Y.-C., Padilla N., Li B., 2015, [MNRAS](#), 451, 1036

Cai Y.-C., Neyrinck M., Mao Q., Peacock J. A., Szapudi I., Berlind A. A., 2016, [MNRAS](#), 466, 3364

Camacho-Ciurana G., Lee, P. Arsenov, N. Kovács, A. Szapudi, I. Csabai, I. 2024, [A&A](#), 689, A171

Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, [MNRAS](#), 447, 646

Cautun M., Paillas E., Cai Y.-C., Bose S., Armijo J., Li B., Padilla N., 2018, [MNRAS](#), 476, 3195

Chan K. C., Hamaus N., Desjacques V., 2014, [PRD](#), 90, 103521

Chantavat T., Sawangwit U., Wandelt B. D., 2017, [ApJ](#), 836, 156

Clampitt J., Jain B., 2015, [MNRAS](#), 454, 3357

Clampitt J., Jain B., Sánchez C., 2016, [MNRAS](#), 456, 4425

Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, [ApJL](#), 648, L109

Colberg J. M., et al., 2008, [MNRAS](#), 387, 933

Contarini S., et al., 2022, [A&A](#), 667, A162

Contarini S., Pisani A., Hamaus N., Marulli F., Moscardini L., Baldi M., 2023, [ApJ](#), 953, 46

Contarini S., Pisani, A. Hamaus, N. Marulli, F. Moscardini, L. Baldi, M. 2024, [A&A](#), 682, A20

Crocce M., Castander F. J., Gaztañaga E., Fosalba P., Carretero J., 2015, [MNRAS](#), 453, 1513

Curtis O., McDonough B., Brainerd T. G., 2025, [ApJ](#), 985, 244

DESI Collaboration et al., 2016, [arXiv:1611.00036](#),

Davies C. T., Klein M., Fumagalli A., Mohr J. J., 2025, [arXiv:2510.16799](#),

Dawson K. S., et al., 2012, [AJ](#), 145, 10

De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, [MNRAS](#), 459, 3078

Diemer B., 2018, [ApJS](#), 239, 35

Doré O., et al., 2014, [arXiv e-prints](#), p. [arXiv:1412.4872](#)

Dvali G., Gabadadze G., Porrati M., 2000, [Physics Letters B](#), 484, 112

Eifler T., et al., 2021, [MNRAS](#), 507, 1746

Einstein A., 1916, [Annalen der Physik](#), 354, 769

Eisenstein D. J., et al., 2011, [AJ](#), 142, 72

Euclid Collaboration: Castander F., Fosalba P., Stadel J., et al., 2025, [A&A](#), 697, A5

Euclid Collaboration: Cropper M., Al-Bahlawan A., Amiaux J., et al., 2025, [A&A](#), 697, A2

Euclid Collaboration: Jahnke K., Gillard W., Schirmer M., et al., 2025, [A&A](#), 697, A3

Euclid Collaboration: Mellier Y., Abdurro'uf Acevedo Barroso J., et al., 2025, [A&A](#), 697, A1

Euclid Collaboration: Pocino A., Tutusaus I., Castander F. J., et al., 2021, [AAP](#), 655, A44

Euclid Collaboration et al., 2025, [AAP](#), 697, A1

Faber S. M., Gallagher J. S., 1979, [ARA&A](#), 17, 135

Falck B., Koyama K., Zhao G.-B., Cautun M., 2018, [MNRAS](#), 475, 3262

Fang Y., et al., 2019, [MNRAS](#), 490, 3573

Farrens S., et al., 2022, [A&A](#), 664, A141

Fixsen D. J., 2009, [ApJ](#), 707, 916

Flaugher B., et al., 2015, [AJ](#), 150, 150

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, [PASP](#), 125, 306

Fosalba P., Gaztañaga E., Castander F. J., Manera M., 2008, [MNRAS](#), 391, 435

Fraser T. S., Paillas E., Percival W. J., Nadathur S., Radinović S., Winther H. A., 2025, [JCAP](#), 06, 001

Friedmann A., 1922, [Zeitschrift fur Physik](#), 10, 377

Gillis B. R., et al., 2013, [MNRAS](#), 431, 1439

Gregory S. A., Thompson L. A., 1978, [ApJ](#), 222, 784

Gruen D., et al., 2015, [MNRAS](#), 455, 3367

Gruen D., et al., 2018, [Phys. Rev. D](#), 98, 023507

Guerrini S., et al., 2025, [AAP](#), 700, A215

Guinot A., et al., 2022, [A&A](#), 666, A162

Guth A. H., 1981, [PRD](#), 23, 347

Gwyn S., et al., 2025, [AJ](#), 170, 324

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, [ApJ](#), 622, 759

Hamaus N., Wandelt B. D., Sutter P. M., Lavaux G., Warren M. S., 2014a, [Phys. Rev. Lett.](#), 112, 041304

Hamaus N., Sutter P. M., Wandelt B. D., 2014b, [Phys. Rev. Lett.](#), 112, 251302

Hamaus N., Pisani A., Choi J.-A., Lavaux G., Wandelt B. D., Weller J., 2020, [JCAP](#), 12, 023

Hamaus N., Aubert M., Pisani A., et al., 2022, [AAP](#), 658, A20

Hang Q., Alam S., Cai Y.-C., Peacock J. A., 2021, [MNRAS](#), 507, 510

Hawken A. J., Aubert M., Pisani A., Cousinou M.-C., Escoffier S., Nadathur S., Rossi G., Schneider D. P., 2020, [JCAP](#), 06, 012

Heymans C., et al., 2012, [MNRAS](#), 427, 146

Hildebrandt H., et al., 2012, [MNRAS](#), 421, 2355

Hoyle B., et al., 2018, [MNRAS](#), 478, 592

Hu W., Sawicki I., 2007, [PRD](#), 76, 064004

Hubble E., 1929, [Proceedings of the National Academy of Science](#), 15, 168

Huchra J., Davis M., Latham D., Tonry J., 1983, [ApJS](#), 52, 89

Huff E., Mandelbaum R., 2017, [arXiv:1702.02600](#)

Icke V., 1984, [MNRAS](#), 206, 1P

Ivezić Ž., et al., 2019, [ApJ](#), 873, 111

Jõeveer M., Einasto J., Tago E., 1978, [MNRAS](#), 185, 357

Jamieson D., Loverde M., 2019, [PRD](#), 100, 123528

Jeffrey N., et al., 2021, [MNRAS](#), 505, 4626

Jeong D., Komatsu E., Jain B., 2009, [Phys. Rev. D](#), 80, 123527

de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A., 2013, [Experimental Astronomy](#), 35, 25

Kaiser N., 1984, [ApJL](#), 284, L9

Khoury J., Weltman A., 2004, [PRL](#), **93**, 171104
 Kitaura F.-S., et al., 2016a, [PRL](#), **116**, 171301
 Kitaura F.-S., et al., 2016b, [MNRAS](#), **456**, 4156
 Kovács A., et al., 2022, [MNRAS](#), **515**, 4417
 Kovács A., Beck R., Smith A., Rácz G., Csabai I., Szapudi I., 2022, [MNRAS](#), **513**, 15
 Krause E., Eifler T., 2017, [MNRAS](#), **470**, 2100
 Krause E., Chang T.-C., Doré O., Umetsu K., 2013, [ApJL](#), **762**, L20
 Kreisch C. D., Pisani A., Carbone C., Liu J., Hawken A. J., Massara E., Spergel D. N., Wandelt B. D., 2019, [MNRAS](#), **488**, 4413
 Kreisch C. D., Pisani A., Villaescusa-Navarro F., Spergel D. N., Wandelt B. D., Hamaus N., Bayer A. E., 2022, [ApJ](#), **935**, 100
 Kuijken K., et al., 2015, [MNRAS](#), **454**, 3500
 Kuijken K., et al., 2019, [AAP](#), **625**, A2
 Landy S. D., Szalay A. S., 1993, [ApJ](#), **412**, 64
 Laureijs R., et al., 2011, [ESA/SRE\(2011\)12](#), p. [arXiv:1110.3193](#)
 Lavaux G., Wandelt B. D., 2012, [ApJ](#), **754**, 109
 Lehman K., Schuster N., Lucie-Smith L., Hamaus N., Davies C. T., Dolag K., 2025, [arXiv e-prints](#), p. [arXiv:2502.05262](#)
 Lemaître G., 1931, [MNRAS](#), **91**, 490
 Lesgourgues J., Pastor S., 2006, [Phys. Rep.](#), **429**, 307
 Li Q., et al., 2024, [ApJL](#), **969**, L25
 Liang Y., Zhao C., Chuang C.-H., Kitaura F.-S., Tao C., 2016, [MNRAS](#), **459**, 4020
 Liaudat, T. Bonnin, J. Starck, J.-L. Schmitz, M. A. Guinot, A. Kilbinger, M. Gwyn, S. D. J. 2021, [A&A](#), **646**, A27

Limber D. N., 1953, [ApJ](#), 117, 134
 LoVerde M., Afshordi N., 2008, [Phys. Rev. D](#), 78, 123506
 Mandelbaum R., et al., 2005, [MNRAS](#), 361, 1287
 Massara E., Villaescusa-Navarro F., Viel M., Sutter P., 2015, [JCAP](#), 11, 018
 Massara E., Percival W. J., Dalal N., Nadathur S., Radinović S., Winther H. A., Woodfinden A., 2022, [MNRAS](#), 517, 4458
 Melchior P., Sutter P. M., Sheldon E. S., Krause E., Wandelt B. D., 2014, [MNRAS](#), 440, 2922
 Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution, [doi:10.1017/CBO9780511807244](https://doi.org/10.1017/CBO9780511807244).
 Nadathur S., Hotchkiss S., 2015a, [MNRAS](#), 454, 889
 Nadathur S., Hotchkiss S., 2015b, [MNRAS](#), 454, 2228
 Nadathur S., Percival W. J., 2018, [MNRAS](#), 483, 3472
 Nadathur S., Hotchkiss S., Diego J. M., Iliev I. T., Gottlöber S., Watson W. A., Yepes G., 2015, [MNRAS](#), 449, 3997
 Nadathur S., Carter P. M., Percival W. J., Winther H. A., Bautista J. E., 2019a, [Phys. Rev. D](#), 100, 023504
 Nadathur S., Carter P., Percival W. J., 2019b, [MNRAS](#), 482, 2459
 Nadathur S., et al., 2020, [MNRAS](#), 499, 4140
 Neyrinck M. C., 2008, [MNRAS](#), 386, 2101
 Nicolis A., Rattazzi R., Trincherini E., 2009, [PRD](#), 79, 064036
 Nusser A., Davis M., 1994, [ApJL](#), 421, L1
 Ostriker J. P., Peebles P. J. E., 1973, [ApJ](#), 186, 467
 Padilla N. D., Ceccarelli L., Lambas D. G., 2005, [MNRAS](#), 363, 977
 Paillas E., Cautun M., Li B., Cai Y.-C., Padilla N., Armijo J., Bose S., 2019, [MNRAS](#), 484, 1149

Pellicciari D., Contarini S., Marulli F., Moscardini L., Giocoli C., Lesci G. F., Dolag K., 2023, [MNRAS](#), 522, 152

Penzias A. A., Wilson R. W., 1965, [ApJ](#), 142, 419

Perico E. L. D., Voivodic R., Lima M., Mota D. F., 2019, [AAP](#), 632, A52

Perlmutter S., et al., 1999, [ApJ](#), 517, 565

Planck Collaboration 2020, [AAP](#), 641, A6

Planck Collaboration et al., 2016, [AAP](#), 594, A1

Platen E., van de Weygaert R., Jones B. J. T., 2007, [MNRAS](#), 380, 551

Pollina G., Hamaus N., Dolag K., Weller J., Baldi M., Moscardini L., 2017, [MNRAS](#), 469, 787

Pollina G., et al., 2019, [MNRAS](#), 487, 2836

Potter D., Stadel J., Teyssier R., 2017, [Computational Astrophysics and Cosmology](#), 4, 2

Radinović S., Nadathur S., Winther H. A., et al., 2023, [AAP](#), 677, A78

Raghunathan S., Nadathur S., Sherwin B. D., Whitehorn N., 2020, [ApJ](#), 890, 168

Reid B., et al., 2015, [MNRAS](#), 455, 1553

Riess A. G., et al., 1998, [AJ](#), 116, 1009

Ronconi T., Contarini S., Marulli F., Baldi M., Moscardini L., 2019, [MNRAS](#), 488, 5075

Ross A. J., et al., 2016, [MNRAS](#), 464, 1168

Rubin V. C., Ford Jr. W. K., 1970, [ApJ](#), 159, 379

Ruggeri R., et al., 2023, [MNRAS](#), 525, 3865

Salcedo A. N., Pisani A., Hamaus N., 2025, [arXiv:2504.08221](#)

Sartori S., et al., 2025, [AAP](#), 700, A17

Schuster N., Hamaus N., Pisani A., Carbone C., Kreisch C. D., Pollina G., Weller J., 2019, [JCAP](#), 12, 055

Schuster N., Hamaus N., Dolag K., Weller J., 2023, [JCAP](#), 05, 031

Shankaranarayanan S., Johnson J. P., 2022, [General Relativity and Gravitation](#), **54**, 44

Sheldon E. S., et al., 2004, [AJ](#), **127**, 2544

Sheth R. K., van de Weygaert R., 2004, [MNRAS](#), **350**, 517

Shim J., Park C., Kim J., Hwang H. S., 2021, [ApJ](#), **908**, 211

Shirasaki M., Takada M., 2018, [MNRAS](#), **478**, 4277

Singh S., Mandelbaum R., Seljak U., Slosar A., Vazquez Gonzalez J., 2017, [MNRAS](#), **471**, 3827

Sinha M., Garrison L. H., 2020, [MNRAS](#), **491**, 3022

Smoot G. F., et al., 1992, [ApJL](#), **396**, L1

Speagle J. S., 2020, [MNRAS](#), **493**, 3132

Su C., et al., 2023, [ApJ](#), **951**, 64

Sugiyama S., et al., 2022, [Phys. Rev. D](#), **105**, 123537

Sutter P. M., Lavaux G., Hamaus N., Wandelt B. D., Weinberg D. H., Warren M. S., 2014, [MNRAS](#), **442**, 462

Sutter P. M., et al., 2015, [Astronomy and Computing](#), **9**, 1

Sánchez C., et al., 2016, [MNRAS](#), **465**, 746

Tavasoli S., 2021, [ApJL](#), **916**, L24

Tegmark M., Taylor A. N., Heavens A. F., 1997, [ApJ](#), **480**, 22

Thiele L., Massara E., Pisani A., Hahn C., Spergel D. N., Ho S., Wandelt B., 2024, [ApJ](#), **969**, 89

Vainshtein A. I., 1972, [Physics Letters B](#), **39**, 393

Vielzeuf P., et al., 2021, [MNRAS](#), **500**, 464

Vielzeuf P., Calabrese M., Carbone C., Fabbian G., Baccigalupi C., 2023, [JCAP](#), **08**, 010

White M., 2014, [MNRAS](#), **439**, 3630

Wilks S. S., 1938, [The Annals of Mathematical Statistics](#), **9**, 60

Williams M. J., Macpherson H. J., Wiltshire D. L., Stevens C., 2025, [MNRAS](#), **536**, 2645

Woodfinden A., Nadathur S., Percival W. J., Radinovic S., Massara E., Winther H. A., 2022, [MNRAS](#), **516**, 4307

Wright A. H., Hildebrandt H., van den Busch J. L., Heymans C., 2020, [AAP](#), **637**, A100

Wu H.-Y., Weinberg D. H., Salcedo A. N., Wibking B. D., Zu Y., 2019, [MNRAS](#), **490**, 2606

Xu W., Zhao C., Su C., Shan H., Liu Y., 2026, [arXiv e-prints](#), p. [arXiv:2603.09278](#)

Zehavi I., et al., 2011, [ApJ](#), **736**, 59

Zel'dovich Y. B., 1970, [AAP](#), **5**, 84

Zhang Z., et al., 2024, [A&A](#), **691**, A75

APPENDICES

Appendix A

Comparing Covariance Terms

A.1 BOSS and UNIONS Covariance

As discussed in Sect. 5.2.3, the full analytic covariance matrix without simplifications can be written as

$$\begin{aligned} \text{Cov} [\Delta\Sigma (\theta_1), \Delta\Sigma (\theta_2)] &= \frac{1}{4\pi f_{\text{sky}} \langle \Sigma_c^{-1} \rangle^2} \int \frac{d\ell}{2\pi} \overline{J_2(\ell\theta_1)} \overline{J_2(\ell\theta_2)} \\ &\times \left[\frac{C_{\kappa\kappa}(\ell)}{n_V} + C_{VV}(\ell) C_{\kappa\kappa}(\ell) + C_{VV}(\ell) \frac{\sigma_\epsilon^2}{n_g} + C_{V\kappa}^2(\ell) \right] \\ &+ \delta_{\theta_1, \theta_2}^K \sigma_{\text{SN}, \Delta\Sigma}^2, \quad (\text{A.1}) \end{aligned}$$

where we have expanded the binomial multiplication seen in Eq. (5.5) and converted from the units of shear to the units of $\Delta\Sigma$. We now want to compare the contributions from each of the terms in square brackets to the total covariance. In order to obtain analytic formulas for each of the angular power spectra, $C_{AB}(\ell)$, here, we use Limber's approximation (Limber, 1953; LoVerde & Afshordi, 2008) to write the general formula for any tracers A,B:

$$C_{AB}(\ell) = \int d\chi \frac{W_A(\chi) W_B(\chi)}{\chi^2} P_{AB} \left(k = \frac{\ell + 1/2}{\chi} \right), \quad (\text{A.2})$$

where $W_{A,B}$ are tracer-dependent weighting functions. For example, if we write

$$W_\kappa(\chi) = c\bar{\rho} \int_{z(\chi)}^\infty dz_s \frac{n(z_s)}{H(z_s) \Sigma_c(z_s, \chi)}, \quad (\text{A.3})$$

and we note that $P_{\kappa\kappa} = P_{\text{mm}}$ as lensing is sensitive to the underlying matter distribution, we end up with the definition of $C_{\kappa\kappa}$ that we used in Eq. (5.8).

In order to define W_V , we follow the definition used in [Bonici et al. \(2023\)](#):

$$W_V(z) = \frac{H(z)}{c} n_V(z) b_V, \quad (\text{A.4})$$

where b_V is the void bias. Incorporating the void bias into the weighting function here allows us then to write the void-void power spectrum, P_{VV} like so:

$$P_{VV}(k, z) = [1 - S_N(k)] P_{\text{mm}}(k, z), \quad (\text{A.5})$$

where $S_N(k)$ is a low- k pass filter to cut out small scale information. This is done to avoid the need to model the effects of the void exclusion principle, wherein two voids cannot overlap too strongly, else they are treated as a single void by a void finder. Since we are only interested in an order of magnitude estimate here to gauge the level of contribution these terms have to our covariance, modelling the small-scale information including the void exclusion principle here is outside of the scope of this work. The low- k pass filter is the smoothstep function:

$$S_N(k) = \begin{cases} 0 & \text{if } \frac{k}{k_{\text{ref}}} \leq a \\ \left(\frac{k}{k_{\text{ref}}}\right)^{N+1} \sum_{n=0}^N \binom{2N+1}{N-n} \binom{N+n}{n} \left(-\frac{k}{k_{\text{ref}}}\right)^n & \text{if } a \leq \frac{k}{k_{\text{ref}}} \leq b, \\ 1 & \text{if } \frac{k}{k_{\text{ref}}} \geq b \end{cases}, \quad (\text{A.6})$$

with parameters set similar to those in [Bonici et al. \(2023\)](#): $a = 1$, $b = 1.8$, $N = 3$, and the reference scale k_{ref} is determined to be the scale of the average void radius: $k_{\text{ref}} = 2\pi/R_V$.

Therefore, in order to compute terms involving voids as tracers, we need an estimate of the void bias. We initially attempt to measure it from our data. If we measure the void-void clustering signal ξ_{VV} , we can compute the matter-matter correlation function from the fiducial cosmology. Since we are cutting out the small scale information, these two quantities are related by the void bias factor:

$$\xi_{VV}(r) = b_V^2 \xi_{\text{mm}}(r) \quad (\text{A.7})$$

We measure the void-void clustering signal using `pycorr` ([Sinha & Garrison, 2020](#)), which uses jackknife to determine the error bars, then compute the matter-matter correlation function using `Colossus` ([Diemer, 2018](#)) and the fiducial cosmology. However, as seen in [Clampitt et al. \(2016\)](#), the void bias is only measurable in configuration space between scales corresponding to twice the average void radius (due to the void exclusion) and the baryon acoustic oscillation (BAO) peak, which affects the matter correlation function. Because of these effects, [Clampitt et al. \(2016\)](#)

measured their void bias from $2R_V$ to $80h^{-1}\text{Mpc}$. Unfortunately, because our voids are traced by LRGs, which are a sparse and highly biased tracer sample, our average (unweighted) void radius is around $40 h^{-1}\text{Mpc}$. This means that we would have no region within which to safely determine the void bias. However, again, as we are only interested in an order of magnitude estimate here, we instead attempt to measure the void bias in the region 2 to $5R_V$. We then take the computed matter-matter correlation function, multiply it by a free parameter of b_V^2 , and then perform a least-squares fit to the void-void correlation measurement from `pycorr`. The results of this fit are shown in Fig. A.1. We can observe that the large size of an average void in our catalogues means that the primary positive matter-matter correlation signal that we fit is the BAO peak. From studies that have attempted to measure the BAO signal from void clustering (Liang et al., 2016; Kitaura et al., 2016a), it was demonstrated that non-overlapping voids do not carry the BAO signal. Since ZOBOV voids are by construction non-overlapping, we do not expect the BAO peak to be included in our clustering profile. However, excluding the BAO peak results means that the primary matter-matter signal on these large scales is negative, resulting in unphysical negative fits to the square of the void bias. The same holds true even when fitting to smaller scales like $1.5R_V$. We therefore come to the conclusion that we are unable to constrain a reasonable estimate of the void bias from void-void clustering.

Instead, since we are interested in an order of magnitude estimate, we can use the results of Jamieson & Loverde (2019), who measured the void bias in Separate Universe Simulations from VIDE voids. They measure a void bias at a redshift of 0.5 that ranges from roughly 1.5 at the smallest voids, to -1 at the largest, crossing 0 at roughly $60 h^{-1}\text{Mpc}$. These large-scale void bias measurements are of similar order as those of Hamaus et al. (2014a), who used simulated voids with tracer densities matching SDSS DR7. As such, we assume a void bias value of 1.5 for Small, 0.5 for Full, LOWZ, and CMASS, and a value of -1 for Large. We assume a negative value for Large despite its average void radius coinciding with the $b_V = 0$ size as we find from the integrated density contrast that this catalogue is expected to be undercompensated. A bias of zero would indicate compensation. In contrast, we find that Full, LOWZ, and CMASS are compensated, indicating very small values of the void bias. We choose to set the void bias to 0.5 as a conservative estimate for this order of magnitude approximation for $C_{V\kappa}$ and C_{VV} .

Using these values of the void bias, we can then compute C_{VV} and $C_{V\kappa}$ using Eqs. (A.2), (A.3), and (A.4), using Colossus (Diemer, 2018) to compute the linear matter power spectrum. Then, we can compute the four terms that exist inside of the square brackets of Eq. (A.1). We note that it is not clear what a reasonable estimate of the sky fraction would be, as we include voids whose centre exists outside of the overlap of the BOSS and UNIONS footprints. Given that the void binning is so large in this case, these voids will still have a significant number of lens-source pairs, but create a “fuzzy” edge around the intersection where void centres exist while sources do not. In initial testing, we found that the comparison to random covariance, which we used for

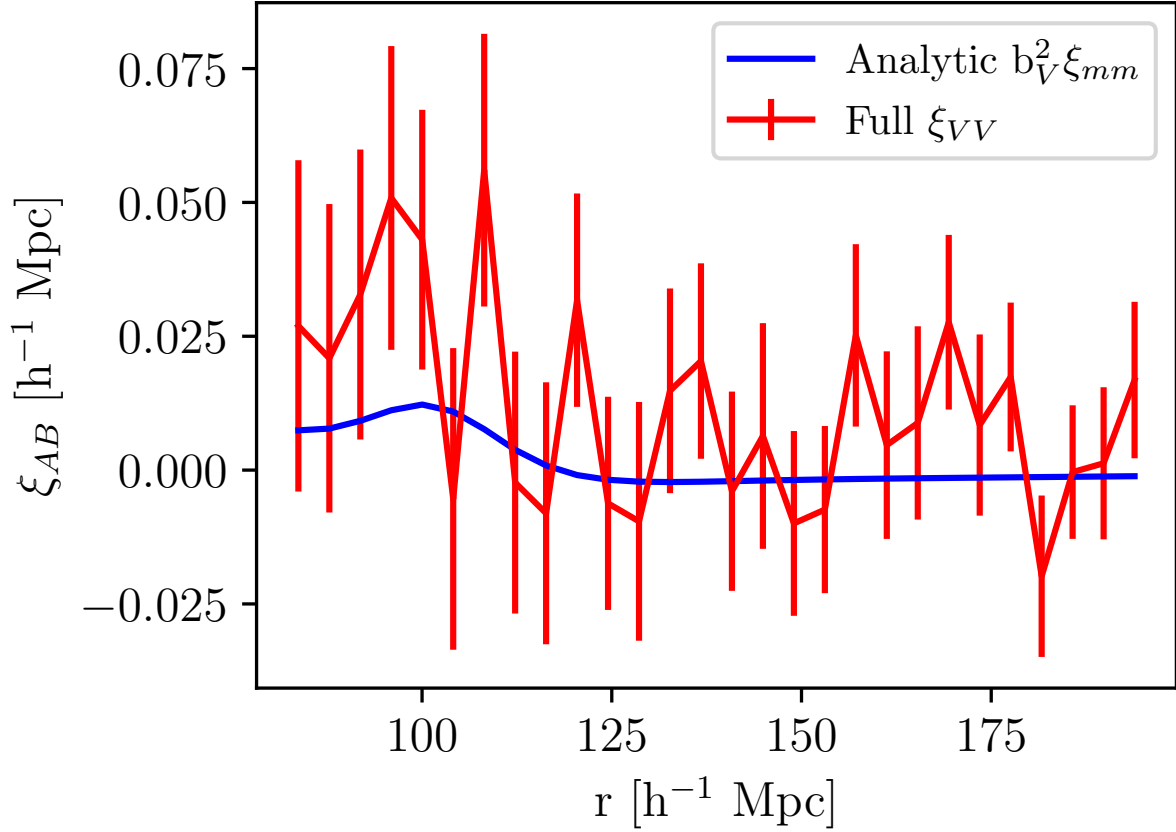


Figure A.1: Measurement of the void bias from the BOSS void sample using the void-void correlation measurement from the Full catalogue and an analytical calculation of the matter-matter correlation function. The fit is performed on scales larger than two times the average void radius to remove the void exclusion effect from the analysis.

validation in Fig 5.2 differed by about 10 per cent in all bins when using the strict intersection of the two surveys. For the purposes of these plots, we choose a sky fraction value that corresponds to the strict intersection, with the understanding that this will be a slight overestimation in the final covariance terms. This area problem is avoided in the final covariance model in Eqn. 5.10 under the assumption that $C_{V\kappa}$ and C_{VV} are negligible. We plot all four terms in Fig. A.2 over a range of scales, where we observe that the $C_{\kappa\kappa}/N_V$ term is larger than all other terms on the plotted scales by more than an order of magnitude. This then provides reasonable justification to exclude the other three terms from the covariance formula.

A.2 Euclid Flagship Covariance

We want to also compare the covariance terms for the *Euclid* Flagship data as well, given the increased quality of data and the higher redshift of these simulated void catalogues. Using the full tomographic covariance formula presented in Eqn. 6.1, with general tomographic C_{AB} terms given in Eqn. 6.2, we define the weighting function for void tracers in the same manner as Eqn. A.4, only accounting for the different $n_V(z)$ and b_V terms that correspond to the different tomographic void bins. We apply the same low- k pass filter in Eqn. A.6 to the power spectrum for all terms that involve voids. This means that small-scale information, which is sensitive to the void finder through void exclusion (Hamaus et al., 2014a), is removed.

We additionally attempt to measure the void bias in the Flagship void catalogues by comparing the void-void clustering measurement using `pycorr` to the matter-matter correlation function computed by `Colossus` on scales larger than $2R_V$. We find similar results as in Appendix A.1, in that the void exclusion region is too large to fit any data reliably between it and the BAO scale. Attempts to fit scales larger than this result in unphysical negative void bias squared values. We find that it is therefore difficult to measure void bias from void-void clustering of ZOBOV voids in configuration space from any survey due to their large size resulting in large void exclusion regions. This is in contrast to the results of Clampitt et al. (2016) and Bonici et al. (2023), who both used projected void finders, which on average have smaller radius distributions, pushing the void exclusion region to smaller scales.

We again use the results of Jamieson & Loverde (2019) to provide estimates of the void bias, and we match their void biases to our average void sizes to determine the value of the void bias used in this covariance model. Once we have these values, we can now compute and compare the contributions to the covariance from void tracers. We plot only the case of the first void and source tomographic bins in Fig. A.3, but we observe similar order of magnitude differences across different void and source tomographic bins. In this plot, we additionally bring in the prefactor of $(4\pi f_{\text{sky}})^{-1}$, which is the survey area of the full octant into each covariance term in the expansion

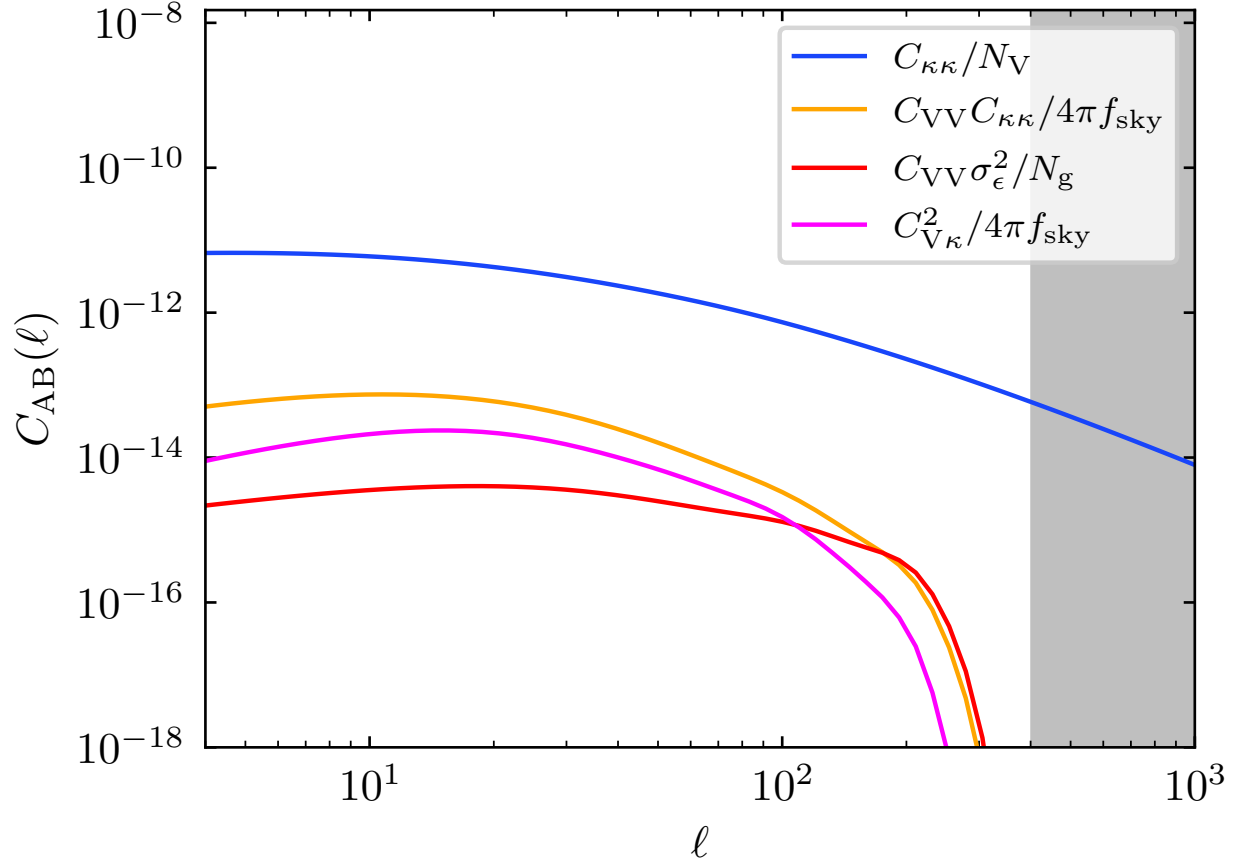


Figure A.2: Comparison plot of the different expanded terms of the analytic covariance model for the Full void catalogue from the BOSS data. The non-grey region includes all scales between the maxima of the Bessel window functions for the largest ($3R_V$) and smallest bins used in the Full catalogue.

of Eqn. 6.1. This has the effect of turning the area number density of voids and galaxies, n_g , and n_V , into the total number count across the simulation. From this plot, we observe that the C_{KK}/N_V term dominates across all scales by at least an order of magnitude, and is therefore the only term of relevance of the four. However, we still note that this analysis excludes small-scale void correlations, which are expected to be small for void clustering given the void exclusion principle, but may not be the case for $C_{V\kappa}(\ell)$

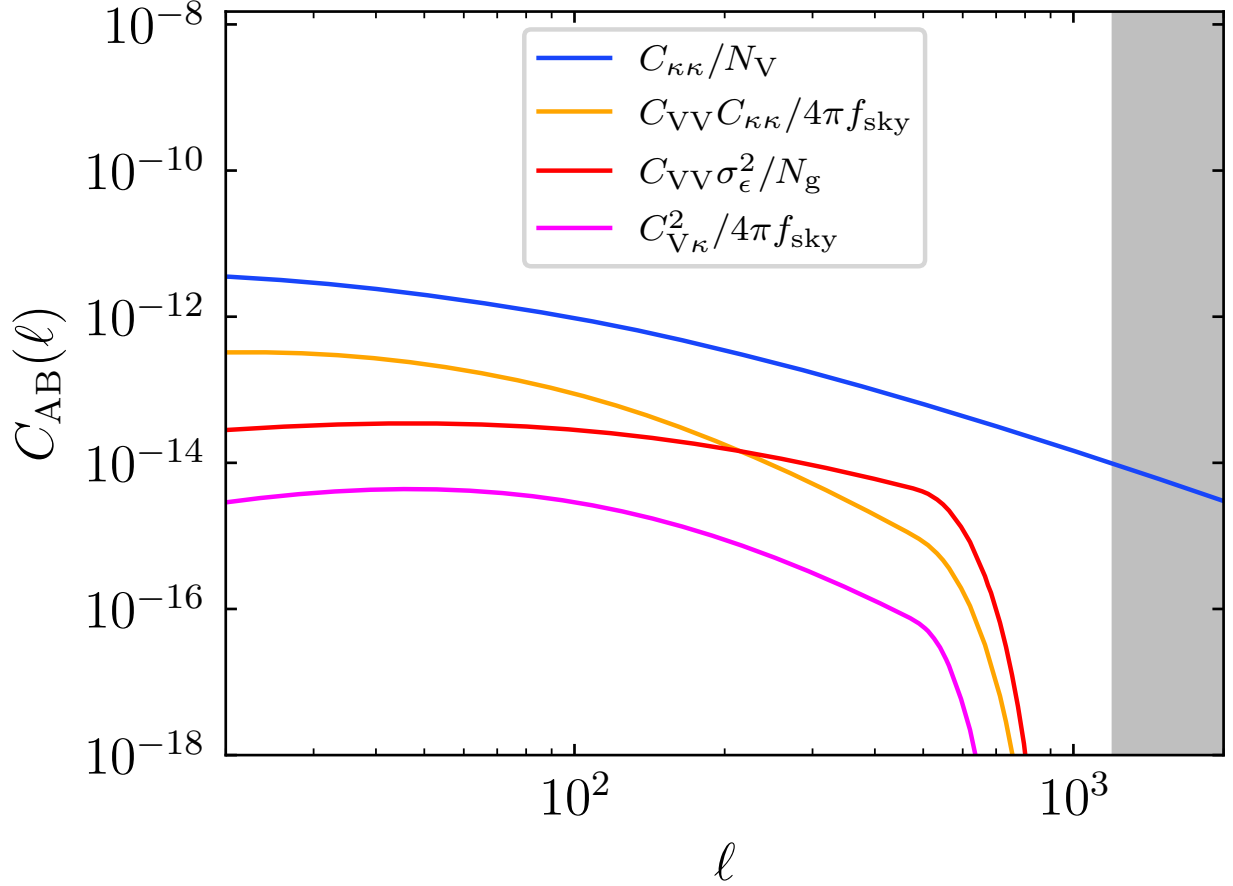


Figure A.3: Comparison of the different angular power spectra terms entering into the tomographic covariance model in Eqn. 6.1 across a range of scales for the lowest redshift void and source bin in the *Euclid* Flagship catalogues. We note that angular power spectra terms that include a void tracer are cut off at small scales to avoid modelling regions that are dependent on the void finder. The greyed out regions indicate scales beyond the projected binning for an average-sized void.

Appendix B

Investigating the Five-Parameter HSW Model Posterior

In this section we attempt to explore properties of the posterior of the HSW model with the B+15 modification (see Eq. 5.14) to void matter profiles of circumcentre voids, as was done in Chapter 5. Generally, the four-parameter model of [Hamaus et al. \(2014b\)](#) has been used in the literature (with the exception of modified gravity studies), which originally fixed $r_{s_2} = R_V$. The decision to use this modified model for the analysis in Chapter 5 was motivated by the findings of [Nadathur & Hotchkiss \(2015a\)](#), who found qualitatively that the five-parameter model produces a better fit to circumcentre voids, while the HSW model was initially empirically determined using barycentre voids. This was further motivated by the particular results after fitting the four-parameter HSW model to the measurements in Chapter 5 that will be detailed later in this section.

As encountered in Fig. 5.6, the resulting posterior of the five-parameter fit to the Full catalogue (with similar complexities present in all catalogues but Small) is fairly complex and non-Gaussian. It further appears that some parameters, such as r_{s_1}/R_V and β contain regions where these values are unconstrained, while others, such as δ_c and r_{s_2}/R_V are much more constrained and Gaussian-like. An additional peculiarity present in this corner plot is the location of the least-squares best fits is often two or more σ away from the peak of the marginalized histograms. Since these posteriors were built from the steps of an ensemble of MCMC walkers, a natural question to ask is whether the chains have sufficiently converged. To validate this, we used the *dynesty* package ([Speagle, 2020](#)), which runs a nested sampler algorithm to estimate the posterior as opposed to a random-walk approach. This represents an independent constraint on the posterior which additionally has the advantage of a natural definition of a stopping criterion. Comparing the resulting contours, we found little change in the contour shape, indicating that our MCMC analysis was converged.

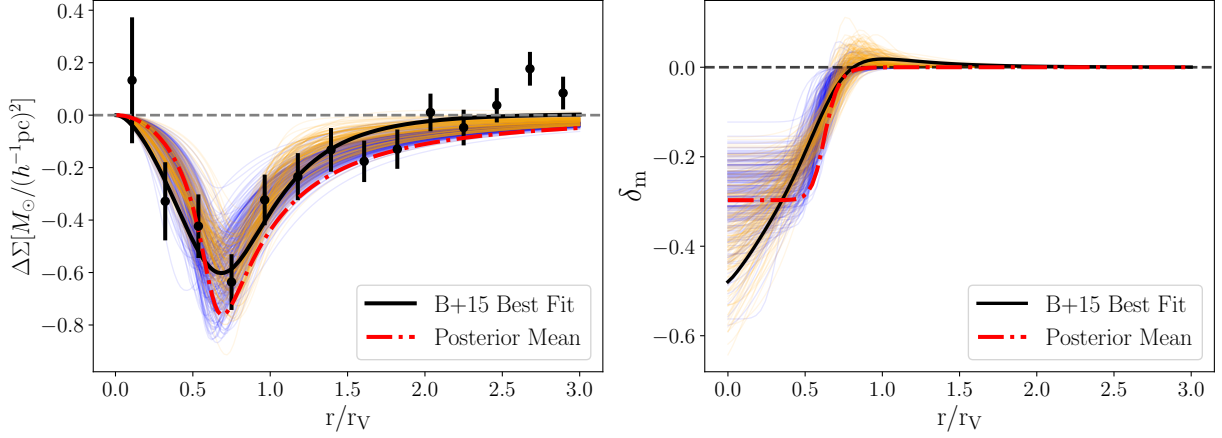


Figure B.1: The $\Delta\Sigma$ and δ_m plots of the Full sample as shown in Fig. 5.4. The orange lines represent 200 MCMC samples from the posterior region that includes the best-fit model, represented by the black line. The blue lines represent 200 MCMC samples from the posterior region where the $\delta_m(r/r_v)$ profile is step-function-like, which occurs when r_{s_2} is sufficiently smaller than r_{s_1} . Additionally, we plot the model corresponding to the posterior mean in the red dot-dashed line, measured by taking the average of MCMC samples in each model parameter.

A potential explanation for the behaviour of this posterior comes from studying the shape of the contour in the $r_{s_1}/R_V - r_{s_2}/R_V$ plane. This L-shaped contour seems to indicate that there are two distinct regions of parameter space that are fit here – one with large values of r_{s_2}/R_V , and one with small values. The large-valued region, which includes the best fit, seems to provide a narrow, constrained region of the r_{s_1} parameter, while the small-valued region seems only able to place a lower bound, with a mostly uniform sampling density that appears to extend past the upper prior bound, if allowed. Looking at the $\delta_c - r_{s_1}/R_V$ and the $\delta_c - \alpha$ contours, we see similar contour shapes, denoting a distinction between high values of δ_c and low values, where the high value regions indicate unconstrained regions of the other parameter, while low value regions contain the best fits and are more constrained.

Examining the formula for the five-parameter model itself, we note an interesting relationship between the r_{s_1} and r_{s_2} parameters. If r_{s_2} is sufficiently smaller than r_{s_1} , then the profile sharply transitions to 0 from the effect of the denominator at a scale before the profile crosses 0 from the effect of the numerator. The resulting 3D density profile appears more like a step-function, featuring a flat underdense region, a sharp near-vertical transition to 0, then a flat, 0 region on large scales. This step-function-like profile is relatively insensitive to changes in α and β , who normally are largely responsible for the slope from underdense to 0, and the slope from overdense

to 0 in a typical void profile. Further, it is insensitive to changes in r_{s_1} above some sufficient lower bound, as changing where the profile crosses 0 after the denominator forces it to asymptotically approach 0 is immaterial to the overall profile's shape. Therefore, from this limiting behaviour of the formula itself, if there is a combination of values of δ_c and r_{s_2}/R_V that fit the data with r_{s_2} sufficiently smaller than r_{s_1} , then the other three parameters are unconstrained.

However, it can be seen from the 3D density plots of the least-squares best fit in Figs. 5.4 and 5.5 that the best-fit profiles do not exhibit these step-function-like behaviours, and according to the corner plot for the Full fit, this best-fit lies within the smaller, more constrained region associated with larger values of r_{s_2} . This indicates that there seem to be two distinct regions of parameter space fit by this posterior. One region where the fit is primarily determined by δ_c and r_{s_2} and is unconstrained by the rest of the parameters, and one where the fit is determined by most if not all of the parameters. The typical difference in χ^2 between these two regions is roughly 3, so both regions represent acceptably likely fits to the data. However, due to the fact that these corner plots are constructed from steps taken from the ensemble of walkers of the MCMC analysis, the former region is over-represented given the unconstrained nature of some of the parameters, and the fact that this unconstrained region is flat in likelihood space. This would explain why the best fitting model is located away from the mean of the marginalized histograms, particularly in δ_c and r_{s_2} .

We additionally noted in Sect. 5.3 that visually, in the 3D density profiles, the best fits did not seem fully represented by the scatter presented by the MCMC samples. Noting the two-region posterior, we expect the scatter to be more representative if we isolate the region that contains the best fit. In order to demonstrate this effect, we make a cut in r_{s_2} at a value of 0.73 to separate these two regions for the Full catalogue. The value of 0.73 was chosen to primarily exclude the horizontal region of the r_{s_1} - r_{s_2} contour. We show the results of this cut in Fig. B.1. The MCMC samples with $r_{s_2} > 0.73$ are coloured orange to represent the region of the posterior that includes the best fits. The blue samples represent the region whose fits are determined by only δ_c and r_{s_2} . For this plot, we used only the Full measurement, and plotted 200 samples from each posterior region. While difficult to tell from the $\Delta\Sigma$ plot, which indicates that both posterior regions fit the data well, we can clearly see in the 3D density plot that the scatter from the orange samples better represents the best fit line, while the scatter in the samples shown in blue features a steeper transition from underdense to overdense, more representative of the step-function-like posterior region.

We additionally plot the five-parameter HSW model corresponding to the mean of the projected 1D posteriors, taken as the average of the MCMC samples in each of the five parameters individually, in the red dot-dashed lines in Fig. B.1. We can observe in the plot on the right that the mean of the posterior is representative of the step-function-like profiles shown by the subset of the MCMC samples in blue, indicating that the larger volume of the step-function-like

posterior region is more prominently featured within the average. We do however observe that the posterior mean represents an extreme value of the blue scatter on mid-to-large scales in the left plot. This is likely due to the skewed nature of some of the marginalized histograms in the contour plot (notably α). It is for these reasons we opted to plot the best fitting model as opposed to the posterior mean in Figs. 5.4 and 5.5.

The existence of these two posterior regions then raises the question of whether this is an effect of the use of *Revolver* voids or whether it is due to the increased flexibility offered by the five-parameter model from the B+15 modification over the four-parameter HSW model. It can be seen in the corner plot of Fig. 9 of Fang et al. (2019) that the four-parameter model used on VIDE void lensing signals does not produce the step-function-like parameter region, as $r_{s_1}/R_V < 1$, which for the four-parameter model fits (where $r_{s_2}/R_V = 1$) implies $r_{s_1} < r_{s_2}$. We see similar results in corner plots produced by fitting our measurements to the four-parameter model, indicating that this secondary parameter region is a result of the flexibility of the five-parameter model. Further, we find that the weak compensating walls that are a consequence of the circumcentre voids do contribute to how acceptable these types of step-function-like fits are. The Small catalogue, which has a significant overdense wall, does not have these step-function-like MCMC samples in the right-hand column of Fig. 5.5, and we find no similar complex contours in Fig. 5.7. We note that these types of profiles with almost no walls have not been directly observed in the literature for ZOBOV voids, save for the largest voids of a catalogue. In principle, this region could be excluded by the priors, potentially by a prior on the relationship between r_{s_1} and r_{s_2} . However, we additionally observe from the differences in best-fit r_{s_1} and r_{s_2} values in Table 5.1, that the value at which r_{s_2} is sufficiently smaller than r_{s_1} is dependent on other properties, particularly in the case of Large, which exhibits the largest difference between the two. In order to construct a robust set of priors that can exclude this step-function-like region from the five-parameter fit, more investigation is needed.

With this additional region of parameter space in mind, we also wanted to quantitatively evaluate the difference between our measurement as fit by the original four-parameter HSW model versus the fit provided by the B+15 modification. We can see from the best fit results shown in Table 5.1 that profiles prefer values of r_{s_2}/R_V that are slightly less than 1, with the exception of LOWZ and Small voids. Intuitively, this suggests that these measurements prefer the increased flexibility of the B+15 modification. Since these are nested models, we can test this by applying Wilk's Theorem to compare the significance in the best-fitting χ^2 between the four- and five-parameter HSW models. From this, the only significantly improved fit comes from the Full measurement, at a marginal significance of 1.9σ . For the other samples, the improvement is not significant, indicating that the additional degree of freedom does not provide a strong change in the quality of the fits. However, there is a caveat with the four-parameter model fits. With every catalogue, all four-parameter models have α at the prior boundary ($\alpha = 1$), and upon relaxing this

boundary, all fit to values $\alpha < 1$, which produces voids with concave behaviour in the innermost region. Voids with these types of profiles have never been observed in simulations or in void measurements in the literature. It is therefore likely that the underlying cause is the inability of the data to constrain the innermost regions of the voids particularly well. This is likely due to the smallest bins in $\Delta\Sigma$ having the largest error bars, and therefore accepting a wide range of possible inner regions. This effect couples with the weak walls of circumcentre void stacks to produce weak constraints on α . We observe this effect represented by extreme values of both α and δ_c as seen in the four-parameter model and seemingly to a lesser extent (via the linear slopes) in the five-parameter modification. Because of the slightly better interpretations in the innermost regions, we opt to present the results of five-parameter HSW model over the four-parameter, but we note that both four- and five-parameter models can be used interchangeably for our results.

Appendix C

Jackknife Verification

In order to validate the analytical covariance model used in Chapter 6, we measure the jackknife covariance as an empirical estimate for comparison. We subdivide the projected void plane into 100 equal-area patches and measure the lensing signal from each patch. We then combine the lensing signal from 99 patches 100 times, excluding a different patch each time. Then, the jackknife covariance is calculated as

$$C_{jk} = \frac{N_p - 1}{N_p} \sum_k (x_{ki} - \bar{x})(x_{kj} - \bar{x}). \quad (\text{C.1})$$

We plot the resulting diagonals of the jackknife covariance (as the off-diagonal elements of the jackknife are too noisy for easy comparison) in Fig. C.1. We only show the case for the L1 catalogue, as it has the smallest covariance and is thus the most sensitive to discrepancies. We do note that we find consistent behaviour across all catalogues. We additionally plot the diagonal of the corresponding analytical covariance matrix without shape noise. From these two plots, we notice two discrepancies. On large scales, we find that the covariances disagree. This is an expected phenomenon from measuring the jackknife covariance without subtracting the signal around random points in the lens plane (Singh et al., 2017; Wu et al., 2019). Given that we observe in Fig. A.3 that the $C_{vv}(\ell)$ and $C_{vk}(\ell)$ terms are subdominant on large scales, we believe the excess variance is due to this jackknife effect. However, we additionally note a disagreement on small scales, with agreement in the intermediate bins. Since we apply a cutoff to the power spectra for void terms in the covariance, small-scale contributions are not included in the model. Given the void exclusion principle, we do not expect void clustering to be the source of this additional variance. Instead, it is likely due to a small contribution from a “one-halo” term of $C_{vk}(\ell)$. We observe both small-scale and large-scale discrepancies across all seven sub-catalogues considered in this work.

To assess the order of magnitude of this discrepancy in the full error budget, we perform the jackknife-analytical covariance comparison including shape noise, which is the dominant source of noise in this analysis. For the jackknife measurement, this involves adding an additional Gaussian scatter to both components of the galaxy shape output from Flagship before performing the measurement. The result is additionally plotted in Fig. C.1. Compared to the around two to four times larger error bars in the without shape noise case, the small scale discrepancy contributes roughly 20 per cent larger error bars in the first three of fourteen total radial bins. Given that these bins already have the largest error bars due to containing the least number of void-source pairs, it is unlikely that these bins contribute heavily to the model fits, and therefore to the overall significance. We further find that this discrepancy decreases to negligible levels overall at the L4 catalogue, due to being dwarfed by the shape noise component. We additionally note that this discrepancy would result in a roughly 20 per cent increase in the error bars for the Large measurement, but in order to produce an acceptable model fit to the data shown in the $\Delta\Sigma$ plot in Fig. 5.5, we would require error bars more than five times the size. Properly modelling this small-scale effect without an analytical prescription would require modelling the covariance in void profiles at the matter particle level in the simulation, which is beyond the scope of this work.

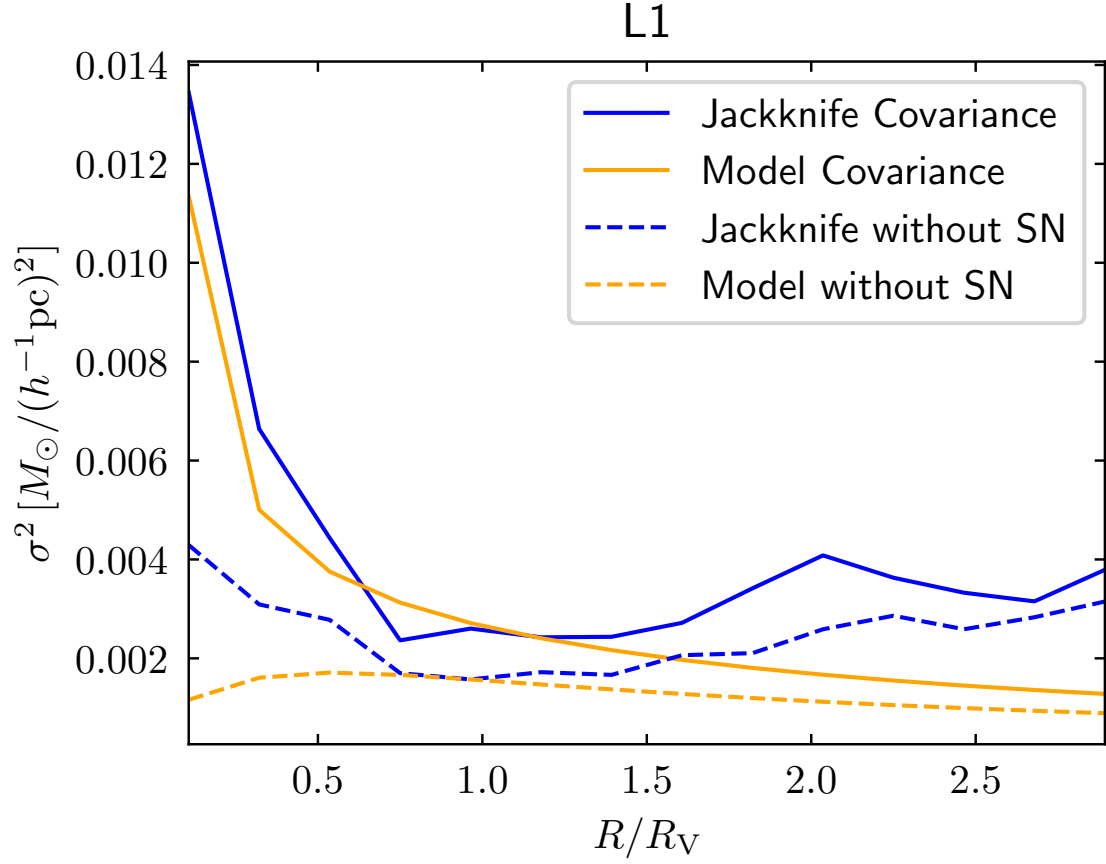


Figure C.1: Diagonal of the jackknife (blue) and analytical (orange) covariances. The dashed lines show the case where no diagonal shape noise component is added to the analytical model, and the Flagship galaxy shapes are not scattered by a random realization of shape noise. The solid lines show the case where both covariances are affected by shape noise.