Building Mock Galaxy Catalogues to Test the Nature of Gravity

A Thesis presented

by

Mahlet Shiferaw

to

the Department of Astronomy

in partial fulfillment of the requirements

for the degree of

Bachelor of Arts

in the subject of

Astrophysics and Physics

Harvard College

Cambridge, Massachusetts

April 17, 2020

Thesis Advisors: Sownak Bose and Daniel Eisenstein Mahlet Shiferaw

Building Mock Galaxy Catalogues to Test the Nature of Gravity

Abstract

The Theory of General Relativity (GR) is very well-tested on local Solar System scales, but tests on the largest cosmological scales have been limited by the volume and precision of existing galaxy surveys. This situation is expected to change in the coming decade with the advent of several new spectroscopic redshift surveys like DESI and *Euclid*. In this project, we test the nature of gravity on these scales by using cosmological simulations to construct mock galaxy catalogs that mimic surveys as closely as possible. In particular, we focus on ΛCDM and three variants of the $f(R)$ model of modified gravity: F6, F5, F4, each of which enhance the strength of gravity relative to GR with increasing intensity. Because of the inherent nonlinearity of the $f(R)$ model, we use large-scale numerical simulations that self-consistently evolve dark matter particles according to these modified equations of motion. In computing the two-point 3D real-space correlation function of the resulting dark matter halos, we find that due to differences in the intensity of the enhancement of gravity, dark matter halos in F4 are significantly less clustered than GR, F6 is slightly less clustered, and F5 is slightly more clustered. We transform each of these halo catalogues using the Halo Occupation Distribution model, which determines the likelihood of a halo having a certain number of galaxies based upon its mass. Finally, we trim these galaxy catalogues even further by applying survey realism, ensuring that the galaxy distribution in the two cosmologies is identical to the observer.

Contents

Abstract **iii**

Acknowledgments **vii**

Dedication **viii**

1 Introduction **1**

1.1 Tests of GR in the Past, Present, and Future 2

1.1.1 Large-Scale Limitations 3

1.2 The Standard Cosmological Model 4

1.2.1 ΛCDM 5

1.3 Outstanding Questions in ΛCDM 7

1.3.1 Dark Matter 7

1.3.2 Dark Energy 8

1.4 Alternative Theories of Gravity 9

1.4.1 A Short Introduction to $f(R)$ Gravity 10

1.4.2 Determining the "Smoking Gun" for Alternative Theories 11

2 The $f(R)$ Modified Gravity Model **13**

2.1 Modified Gravity Models 13

2.1.1 A Brief Comparison to GR 14

2.1.2 The Hu-Sawicki $f(R)$ Gravity Model 16

2.2 The Simulations 20
2.2.1 The Halo Mass Function 23
2.2.2 The 3D Halo Correlation Function 28

3 The Halo Occupation Distribution Model 34
3.1 Populating Dark Matter Halos with Galaxies 34
3.2 Achieving the Target Number Density 36
3.2.1 Finding the Optimal $M_{\rm min}$ 38
3.3 Assigning the Spatial Distribution of HOD Satellite Galaxies 43
3.3.1 The Navarro–Frenk–White Profile 43
3.3.2 Creating a Mock Galaxy Catalogue 44
3.3.3 Assigning Final Positions and Velocities to Mock Galaxies 47
3.3.4 Intermediary Results 49
3.4 Matching the Projected Clustering 49
3.4.1 Defining the Error 52
3.4.2 Minimizing the Error 53
3.4.3 Final Results 55

4 Creating a Realistic Mock Galaxy Survey 61
4.1 Adding Survey Realism 61
4.1.1 Volume Remapping 62
4.1.2 Determining the Radial Comoving Distance D_C 64
4.1.3 Determining the Remap Vectors 66
4.2 The Final Mock Galaxy Surveys 72

5 Conclusions 77
5.1 Discussion 77
5.2 Future Work 78

5.2.1 Redshift Space Distortions 79
5.2.2 Velocity Power Spectrum 81
5.3 Final Thoughts 82

References **85**

Acknowledgments

First, the author would like to thank co-advisor Sownak Bose for his crucial guidance throughout this thesis, as well as co-advisor Daniel Eisenstein, without whom this thesis would not be possible. The author would also like to thank Ramesh Narayan and John Kovac for acting as professor in overseeing, respectively, the 2019 iteration of Astronomy 99a: Senior Thesis in Astrophysics and the 2020 iteration of Astronomy 99b: Senior Thesis in Astrophysics, during which the author completed this research. Finally, the author would like to thank the external readers Xingang Chen and Douglas Finkbeiner, who each provided valuable feedback on the penultimate draft. The computations in this thesis were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University. The data analysed in this thesis can be made available upon request to the author.

To my mom, who unknowingly sparked my love for space.

To my family, friends, and cats for getting me through quarantine.

And, of course, to Miss Rona. You made this year unforgettable.

Chapter 1

Introduction

The Theory of General Relativity (GR), published by theoretical physicist Albert Einstein in 1915, is the currently accepted theory of gravitation in modern physics: it generalizes special relativity and describes gravity as a unified geometric property of both space and time, or "spacetime" (Einstein 1916). The significance of this theory, and the resulting Einstein field equations (EFEs), cannot be understated. GR shows that black holes are not merely mathematical conjecture, but can be formed by a sufficiently compact mass which deforms spacetime (Montgomery et al. 2009). It famously predicts the existence of gravitational waves, or violent ripples in spacetime produced by massive accelerating objects (Einstein 1918), which were first detected by the Advanced Laser Interferometer Gravitational-Wave Observatory in 2015 (Abbott et al. 2016). Most significantly, GR provides the underlying theory behind current cosmological models of the Universe that solve the EFEs (Einstein 1917; Friedmann 1922, 1924; Bucher & Ni 2015). Clearly, GR has been the source of astonishing scientific advances and has thus transformed the landscape of astronomy and physics over the past century. Further

ramifications of GR are discussed in the following section.

1.1 Tests of GR in the Past, Present, and Future

Over the years, GR has passed a remarkable number of precision tests, starting with basic tests of the fundamental pillars of relativity and gravitation (Will 2014). These include the Michelson-Morley experiment, which famously showed that the speed of light is a universal constant, thus providing the basis of relativistic theory and debunking the popular belief in an "aether" or medium through which light travels (Michelson & Morley 1887). In addition, following Einstein's publication of the Theory of GR, the Eötvös experiment confirmed the Einstein equivalence principle that spacetime is curved (v. Eötvös et al. 1922).

GR has also consistently held up to various tests of post-Newtonian gravity, or solutions of the EFEs for the metric tensor, expressed as corrections to Newtonian gravity in the weak-field. This includes an explanation to the long-standing question of the precessing perihelion in Mercury's orbit, a mystery that had remained unsolved for more than half a century: while various *ad hoc* theories were initially proposed, such as a new planet Vulcan located near the Sun (Le Verrier 1859), Einstein later showed that GR, in place of Newtonian gravity, successfully accounts for the discrepancy (Einstein 1916). A prediction of GR was also famously validated by the classic Eddington experiment: as stated by John Archibald Wheeler, "Spacetime tells matter how to move; matter tells spacetime how to curve". Thus, light passing by the Sun will be deflected, a prediction of GR confirmed by astronomer Arthur Eddington's observation of a total solar eclipse in the days following World War I (Dyson et al. 1920). Tests such as these—of GR as

well as its predictions—have solidified it as the current theory of gravitation.

1.1.1 Large-Scale Limitations

The abundance of these precision tests means that GR has remained extremely well-tested on local Solar System scales. Although these tests have been by definition confined to the weak-field regime, there have also been plentiful tests of GR in the strong-field regime, such as the recent detection of the gravitational redshift in the orbit of a star near the massive black hole candidate Sgr A* (Gravity Collaboration et al. 2018).

However, the same precision measures cannot be said of cosmological tests. These tests do exist: for instance, the clustering of galaxies has successfully survived comparisons between observational data from the Sloan Digital Sky Survey and results from the Millenium Simulation, an N-body dark matter particle simulation which assumes GR (Springel et al. 2005; Nurmi et al. 2013). However, tests of GR on these extreme large scales have historically been limited by the volume and resolution of existing galaxy surveys. Limitations on precision mean that a confident conclusion cannot be drawn in the all-too-common event that the difference between theoretical prediction and observed data is smaller than the experimental spread in error, which can occur due to a combination of statistical and systematic limitations. Fortunately, this is soon expected to change as experimental precision increases and statistical error shrinks with the advent of several new spectroscopic redshifts surveys, such as the ground-based Dark Energy Spectroscopic Instrument (DESI), slated to start taking data in summer 2020 (DESI Collaboration et al. 2016), and the space-based *Euclid*, slated to launch in

June 2022 (Amendola et al. 2018).

Until that happens, however, GR remains essentially untested on these largest of cosmological scales. The interest in doing so is driven by the nature of the standard cosmological model ΛCDM, which is further discussed in detail in Section 1.2.1. Briefly, this model requires the existence of a dark matter particle and some cosmological constant Λ, or dark energy. The nature of both dark matter and dark energy remain a mystery, thus driving cosmologists to ask whether gravity could, in fact, be different from GR. After all, the structure that is observed in our Universe cannot automatically inform us of its underlying cause without further investigation. All of this motivates the crucial research questions behind this thesis: how can GR be tested in this regime? Until observational data progresses, how can we take advantage of theoretical research methods to answer this question? Observationally, what impact does this have on cosmology? Are alternate theories of gravity viable? What is the true nature of dark energy?

1.2 The Standard Cosmological Model

Before investigating GR, we must understand the standard cosmological model it encompasses. Shortly after the genesis of GR, the standard model of modern cosmology was birthed by obtaining the exact solutions to the EFEs (Ishak 2019). These equations have thousands of exact solutions, but only a small number of them have any real physical meaning (Stephani et al. 2003). For instance, Delgaty & Lake (1998) found that only 16 out of 127 candidate solutions passed basic tests of physical acceptability, such as isotropy of pressure. Furthermore, of that small population with physical relevance, only a select few have applications to cosmology, including the current standard model.

This model combines the idea of expanding universes (Friedman 1922; Lemaître & Eddington 1931) with the observational expectation that spacetime is globally isotropic and homogenous (Robertson 1935; Walker 1937), known as the cosmological principle (Barrow 1989). Doing so results in the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, which describes the geometry of space.

Using the EFEs in combination with this metric, cosmologists are able to obtain an expression for the Hubble parameter of the Universe at any given redshift in terms of the Hubble constant and the density parameters for matter, radiation, curvature, and dark energy. Although there are slight discrepancies in the measurement of the Hubble constant (Jackson 2007; Freedman & Madore 2010), astronomers have been able to precisely measure the curvature of the Universe as well as the values of each of the cosmological density parameters, primarily through data obtained from the Planck observations (Planck Collaboration et al. 2016). With these measurements, it becomes possible to describe the cosmic mass-energy budget and thus map the entire evolution history and large-scale structure of the Universe exactly.

1.2.1 ΛCDM

This evolution history is described by the ΛCDM model, where Λ is the cosmological constant, or dark energy, responsible for the accelerated expansion of the Universe and CDM stands for Cold Dark Matter, a particle responsible for large-scale gravitational effects which are unaccounted for by ordinary matter. As a result of previous precision observations, as mentioned in Section 1.2, astronomers know the Universe to be nearly spatially flat. Additionally, due to observations of galactic redshifts, there is strong

empirical support for describing the evolution of the Universe using the "Hot" Big Bang Model, which generally states that the Universe expanded from an initially highly dense state to its current, very low-density state (Ryden 2003).

ΛCDM is thus a variant of the "Hot" Big Bang Model with a unique history of cosmological parameter domination. This history begins with an early Universe which was radiation-dominated and full of extremely hot, ionized baryonic gaseous matter. Eventually, around redshift $z \sim 3400$, a transition occurred as the Universe became matter-dominated rather than radiation-dominated. Additionally, as the Universe expanded, this hot, ionized baryonic gaseous matter eventually cooled so that electrons and ions were able to combine and create neutral hydrogen.

In this period, known as the epoch of recombination at redshift $z \sim 1100$, photons no longer scattered off of free electrons, as they had previously (Peebles 1968; Zeldovich et al. 1968; Seager et al. 1999). Because light was able to freely stream throughout the Universe for the first time, a "dense fog" was finally lifted, effectively making the Universe transparent in a process known as photon decoupling. The relic radiation from these first photons is known as the Cosmic Microwave Background. Today, due to the previous opacity of the universe, astronomers are unable to observe anything past the surface of last scattering, which marks the point from which a typical CMB photon last scattered off of an electron.

After photon decoupling and prior to the formation of the first stars and galaxies, the Universe was plunged into an extended epoch appropriately named the "Dark Ages" until redshift $z \sim 30$ (Miralda-Escudé 2003). Eventually, as structure began to form via gravitational collapse, the first stars caused the once-neutral cosmic gas to return

to its previously ionized state. This epoch of reionization, which occurred within the redshift range of $6 < z < 15$, also marked a shift in which baryons became more visible for the next millions of years (Zaroubi 2012). Once these early galaxies had formed, the Universe began its modern dark-energy-dominated era at redshift $z \sim 1$, a shift that is responsible for the current accelerated expansion of the Universe (Frieman et al. 2008). This most recent stage in the formation history of the Universe is where the major differences between GR and alternate theories of gravity are most apparent.

1.3 Outstanding Questions in ΛCDM

Due to overwhelming empirical evidence in their favor, dark matter and dark energy have been shown to be necessary components of any viable cosmological model (Trimble 1987). However, both still pose problems for cosmology in the frameworks of GR and ΛCDM, as it remains unknown what exactly these components are in detail (Debono & Smoot 2016).

1.3.1 Dark Matter

Dark matter is a mysterious, invisible particle which makes up the overwhelming majority of matter in the Universe. Although it was theorized years beforehand (Zwicky 1937), the existence of dark matter was not confirmed until the 1970s, after flat rotation curves of galaxies were observed (Ostriker et al. 1974; Rubin et al. 1980). These flat rotation curves did not match theoretical expectations of decreasing velocity, and implied that galactic mass increases linearly with radius (de Swart et al. 2017). However, this

resulted in a mismatch between the observed and theoretical mass of these galaxies, a discrepancy which ultimately suggested the existence of some sort "missing mass" around the galaxies (Bertone & Hooper 2018). Following Zwicky (1933)'s first usage of the term "dunkle Materie", this missing mass later became known as dark matter, a type of particle which cannot be observed directly, but that does interact gravitationally with photons and baryons.

In the ΛCDM model, this particle is theorized to be a CDM particle: a cold, collisionless dark matter particle with essentially no velocity dispersion. This particle creates a bottom-up model of hierarchical structure formation, wherein low-mass galaxies form before high-mass galaxies, thus growing in size through accretion and merging. This bottom-up model very well matches key measurements of the evolution of large-scale structure in our Universe, as described in Section 1.2.1. However, there is one major tension concerning ΛCDM: despite these aforementioned successes, as well as extensive experimental efforts on the parts of numerous particle physicists, a CDM particle has yet to be detected (Bertone et al. 2005).

1.3.2 Dark Energy

Dark energy, or the cosmological constant Λ, poses yet another conundrum in ΛCDM. From observations of supernova measurements, it was famously found that the Universe is actually undergoing accelerated expansion (Riess et al. 1998; Perlmutter et al. 1999). This discovery suggested the additional existence of a dark energy component that dominates the mass-energy budget of the Universe. Additionally, a multitude of cosmological observations beyond supernovae also support that the FLRW expansion

rate has a late-time acceleration (Blake et al. 2011a,b, 2012).

Despite these agreements, there have been huge discrepancies in attempting to measure the value of the cosmological constant. Namely, astronomical observations suggest a cosmological constant which is many orders of magnitude smaller than the value obtained from quantum field calculation (Weinberg 1989). Additionally, it is not yet understood why the energy density of the cosmological constant is the same exact order of magnitude as the density of matter: had the energy density been any larger, structure in the Universe would not have been able to form (Ishak 2007). Indeed, it is not even understood why dark energy dominates during the present-day rather than earlier in the past, as that also would have inhibited structure formation.

1.4 Alternative Theories of Gravity

These lingering questions regarding ΛCDM, as well as the observational limits on testing GR cosmologically, as overviewed in Section 1.1.1, have led to various proposals of alternative theories of gravity, including scalar-tensor gravity, quadratic gravity, and more (Clifton et al. 2012; Berti et al. 2015). In this thesis, we consider an alternate theory of modified gravity, the $f(R)$ gravity model, where R is the Ricci scalar, an invariant quantity which characterizes the curvature of a spacetime manifold. This is not to be confused with Modified Newtonian Dynamics (MOND), which does not include any dark matter, and instead modifies Newtonian gravitation to account for the missing mass in galactic rotation curves (Milgrom 1983). We provide a short description of $f(R)$ gravity below, but a full overview is located in Chapter 2.

1.4.1 A Short Introduction to $f(R)$ Gravity

Modified theories of gravity were originally motivated by the desire to do away with the cosmological constant Λ. Thus, in modified gravity, dark matter still exists and is often assumed to be a CDM particle. However, most models that are still currently viable still require an explicit cosmological constant to drive accelerated expansion. Despite this, $f(R)$ gravity does aim to largely replace the cosmological constant Λ: rather than relying on a dark energy-like component to the Universe, $f(R)$ gravity proposes that the accelerated expansion of the Universe is primarily driven by deviations from GR. As a result, the investigation of modified gravity is motivated by the desire to test deviations from GR due to differences in the underlying strength of gravity.

However, because GR is so well-tested on local, Solar System scales, viable models of modified gravity must incorporate a "screening" mechanism. In the case of the $f(R)$ gravity model, this is achieved through the "chameleon" mechanism, which suppresses this enhancement of gravity in high-density regions like the Solar System (Khoury & Weltman 2004; Mota & Shaw 2007; Bose et al. 2017). The law of gravitation therefore only deviates from GR in low-density regions, which are typically only found on large cosmological scales. Thus, voids are made to be even more low-density than they already are, leading to overall accelerated expansion of the Universe. At the same time, due to the high density and large radius of curvature R on local scales, the chameleon mechanism suppresses this enhancement of gravity in such regions. Consequently, $f(R)$ gravity effectively reduces to the GR limit in places like the Solar System, thereby matching current and past observations of the Solar System.

In this thesis, we focus on the most-widely researched form of modified gravity,

known as the Hu-Sawicki $f(R)$ gravity model (Hu & Sawicki 2007). In this model of $f(R)$ gravity, a CDM particle is assumed, as well as a chameleon mechanism that suppresses the enhancement of gravity in highly-dense regions. This mechanism makes $f(R)$ gravity an interesting and viable alternative to GR, motivating our core question in this thesis: how can we distinguish $f(R)$ gravity from GR in our Universe? What is the "smoking gun" that we can look for in observations?

1.4.2 Determining the "Smoking Gun" for Alternative Theories

Because of the inherent nonlinearity in the $f(R)$ model, we rely on large-scale numerical simulations that self-consistently evolve dark matter particles according to the $f(R)$ modified equations of motion. We focus on four N-body cosmological simulations in particular: one of GR, as well as three variants of $f(R)$ gravity: F4, F5, and F6, each of which enhance the strength of gravity relative to GR with increasing intensity. However, in reality, astronomers do not observe the dark matter halos that form in N-body simulations; they observe galaxies. We are concerned with using proxies that an observational cosmologist might use, and therefore aim to address the following question: if we make the same selections that are made in a redshift survey, will there be any differences that still persist at the end of the process that might be representative of differences in gravity?

To accomplish this, we take halo catalogues produced from each simulation set and transform them into galaxy catalogues that mimic true galaxy surveys. We do this by populating the halo catalogues with galaxies using the Halo Occupation Distribution (HOD) model (Kravtsov et al. 2004). After creating a pipeline to automate this process,

we apply realism to the galaxy catalogues to create mock surveys. For this step, we take the 3D coordinates of the galaxy catalogue and project them onto the sky in RA and DEC coordinates. We then apply the same redshift selection, sky mask, and incompleteness that real galaxy surveys contain; in particular, we mimic the SDSS-III LOWZ survey, which focuses on nearby galaxies only, at redshifts $z \lessapprox 0.4$ (Manera et al. 2014). With these mock surveys in hand, we are finally equipped to identify any differences between the two models, as these residual differences are representative of the underlying differences in gravity.

The remainder of this thesis is organized as follows: in Chapter 2, we review the theory behind GR and the Hu-Sawicki $f(R)$ gravity model in further detail. We also discuss the N-body simulations used for each of these halo catalogues, such as their size, volume, and resolution limit. Additionally, we present and discuss the Halo Mass Function and Correlation Function for GR and $f(R)$. In Chapter 3, we review the HOD model used to assign positions and velocities to galaxies in each of the halo catalogues. We present and discuss the Galaxy Correlation Function and Matter Power Spectrum obtained from the HOD galaxies. In Chapter 4, we outline the procedure for adding realism to the galaxy catalogues. We compare the mock galaxy surveys between $f(R)$ and GR, examine redshift space distortions, and present final results. Finally, in Chapter 5, we present our conclusions and discuss future steps for further researching this topic.

Chapter 2

The $f(R)$ Modified Gravity Model

2.1 Modified Gravity Models

Modified gravity models are popular alternatives to Einstein's Theory of General Relativity. They include models ranging from Scalar-Tensor, Einstein-Aether, and Bimetric theories, as well as $f(R)$ and other general higher-order theories (Clifton et al. 2012). A commonality of these models is their presumption that the accelerated expansion of the Universe is caused by a large-scale deviation from GR, rather than a cosmological constant Λ or some unknown form of matter, such as dark energy (Joyce et al. 2016). In most modified gravity theories, this modification is due to what is commonly called a "fifth force", which is mediated by a new scalar field. Because of the existence of this fifth force, these models describe a Universe which is only made up of ordinary and dark matter, often presumed to be a cold dark matter particle.

Although the fifth force is a large-scale deviation from GR, its effect is still very

much felt on smaller scales comparable to the typical separation between galaxies. As mentioned in Section 1.4.1, in order to reconcile these changes with observations of the Solar System, modified gravity theories that wish to remain viable often contain an inherent "chameleon" or "screening" mechanism, which acts to suppress the modification of gravity, or in this case, the fifth force, on these scales. In the $f(R)$ gravity model we consider, this suppression of the fifth force is achieved by hiding scalar degrees of freedom in high density regions (Navarro & Acoleyen 2007), the details of which are discussed in Section 2.1.2. In the following subsection, however, we first begin with an overview of GR, and how it compares to the $f(R)$ gravity model.

2.1.1 A Brief Comparison to GR

The Einstein Field Equations (EFEs) of GR are described by Equation 2.1 below, where $G_{\mu\nu}$ is the Einstein tensor representing the curvature of spacetime, Λ is the cosmological constant, $g_{\mu\nu}$ is the metric tensor, G is the gravitational constant, and $T_{\mu\nu}$ is the energy momentum tensor (Janssen et al. 2007).

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{2.1}$$

The FLRW metric is given by Equation 2.2 below, where s is the line element, $a(t)$ is the scale factor, k is the curvature of spacetime, and r is the radial coordinate.

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right) \tag{2.2}$$

Solving the EFEs of Equation 2.1 for this metric, Equation 2.2, as well as the energy

momentum tensor $T_{\mu\nu}$ results in the dynamical Friedmann equations, the first of which can be written in terms of the cosmological parameters below:

$$H^2(a) = H_0^2 \left[\Omega_m^0 a^{-3} + \Omega_r^0 a^{-4} + \Omega_k^0 a^{-2} + \Omega_{de}^0 a^{-3(1+w)}\right] \tag{2.3}$$

Here, w is the dark energy equation of state variable, and the cosmological parameters are as follows: $H(a)$ is the Hubble parameter, H_0 is the Hubble constant, and Ω_m^0, Ω_r^0, Ω_k^0, and Ω_{de}^0 are the density parameters for matter, radiation, curvature, and dark energy, respectively. In contrast, for an FLRW curved background metric in a $f(R)$ gravity model, the dynamical Friedmann equations can be written as Equation 2.4 below (Ishak 2019):

$$H^2 = \frac{1}{3F}\left[8\pi G\bar{\rho} - \frac{1}{2}(f - RF) - 3H\dot{F}\right] - \frac{\kappa}{a^2} \tag{2.4a}$$

$$\dot{H} = -\frac{1}{2F}(8\pi G(\bar{\rho} + \bar{P}) + \ddot{F} - H\dot{F}) + \frac{\kappa}{a^2} \tag{2.4b}$$

Here, the $f(R)$ gravity model generalizes GR by using the general function $f(R)$ in place of the Ricci scalar R used in GR. As such, $F \equiv f_R$, where f_R refers to the scalar field, and the Ricci scalar R can be written as $R = 6(2H2 + \dot{H} + \kappa/a^2)$. Additionally, κ represents the curvature of spatial sections, and overdots such as $\dot{H}$ symbolize a derivative with respect to cosmic time t. In addition to the dynamical Friedmann equations, this modification to gravity ultimately generalizes the EFEs of Equation 2.1, as well as the dynamical Friedmann equations of Equation 2.3. In the following subsection, we discuss these alterations in further detail.

2.1.2 The Hu-Sawicki $f(R)$ Gravity Model

In this thesis, we consider the Hu-Sawicki gravity model, which, like other $f(R)$ gravity models, is based on deviations from GR. In GR, the EFEs contain derivatives that are, at maximum, of the second order (Lovelock 1971). This means that GR can easily be modified by simply allowing for higher order derivatives in the field equations. The $f(R)$ gravity models do just that, changing the EFEs from second-order to fourth-order derivatives. This is accomplished by generalizing the standard Einstein-Hilbert action of GR, which is described by Equation 2.5 below:

$$S = \frac{1}{16\pi G}\int \sqrt{-g}(R - 2\Lambda)d^4x + \int \mathcal{L}_m\left(g_{\mu\nu}, \psi\right) d^4x \tag{2.5}$$

Here, G is Newton's gravitational constant, g the determinant of the metric tensor $g_{\mu\nu}$, and $\mathcal{L}_m$ the Lagrangian density of the matter fields, ψ. This equation follows from the action integral of classical mechanics. In the Hu-Sawicki $f(R)$ Gravity Model however, Equation 2.5 is generalized by replacing the Ricci scalar R with an algebraic function $f(R)$. This modification is described by Equation 2.6 below:

$$S = \frac{1}{16\pi G}\int \mathrm{d}^4x\sqrt{|g|}[R + f(R)] + \int \mathrm{d}^4x\mathcal{L}_m \tag{2.6}$$

In taking the derivative of the Lagrangian $\mathcal{L}_m$ with respect to the metric, one can obtain the modified Poisson equations, which, in effect, describe the dynamical changes made to $f(R)$ gravity. These modified Poisson equations are described by Equation 2.7 below:

$$\nabla^2\Phi = \frac{16\pi G}{3}\delta\rho_m - \frac{1}{6}\delta R\left(f_R\right) \tag{2.7a}$$

$$\nabla^2 f_R = \frac{1}{3c^2}\left[\delta R\left(f_R\right) - 8\pi G\delta\rho_m\right] \tag{2.7b}$$

Here, the scalar field $f_R \equiv \mathrm{d}f(R)/\mathrm{d}R$ has been introduced in order to relate the gravitational potential Φ to the density and curvature at a certain position. These quantities are defined in Equation 2.8 below, where the bar refers to the cosmic mean of a variable:

$$\delta\rho_m \equiv \rho_m - \bar{\rho}_m \tag{2.8a}$$

$$\delta R \equiv R - \bar{R} \tag{2.8b}$$

Finally, we can define $f(R)$ itself, using the Hu-Sawicki gravity model (Hu & Sawicki 2007). Although this model is by no means the only choice, it has desirable properties, and is described by Equation 2.9 below:

$$f(R) = -m^2\frac{c_1\left(-R/m^2\right)^n}{c_2\left(-R/m^2\right)^n + 1} \tag{2.9}$$

The mass scale m can be defined in the following equation, where $\Omega_m h^2$ is the physical matter density, inferred from the Cosmic Microwave Background using GR, and $\bar{\rho}_0 = \bar{\rho}(\ln a = 0)$ is the current average density:

$$m^2 \equiv \frac{\kappa^2\bar{\rho}_0}{3} = (8315\ \mathrm{Mpc})^{-2}\left(\frac{\Omega_m h^2}{0.13}\right) \tag{2.10}$$

The remaining variables, n, c_1, and c_2 are dimensionless model parameters. The specific form of $f(R)$ expressed in Equation 2.9 has been chosen so that it abides by the following limits:

$$\lim_{R\to\infty} f(R) = \text{ const.} \tag{2.11a}$$

$$\lim_{R\to 0} f(R) = 0 \tag{2.11b}$$

It is absolutely necessary that $f(R)$ satisfy these limits, as it ensures that a number of essential observational properties are met. First, $f(R)$ must achieve accelerated expansion without the use of an explicit cosmological constant. Furthermore, it must do so with an expansion history that is close to that of ΛCDM. Lastly, it must preserve the ΛCDM as a limiting case, in order to pass the numerous Solar-System tests of GR. This is achieved through the aforementioned screening mechanism, which can be seen in action in Equation 2.12:

$$\lim_{m^2/R\to 0} f(R) \approx -\frac{c_1}{c_2} m^2 + \frac{c_1}{c_2^2} m^2 \left(\frac{m^2}{R}\right)^n \tag{2.12}$$

In this equation, $f(R)$ has been expanded for the limiting case of high curvature R, equivalently small m^2/R. For cases of high density or spatial curvature, like the Solar System, it becomes clear that the second term will vanish and $f(R)$ will become a constant, thus satisfying the first requirement of Equation 2.11. Therefore, although there is no explicit cosmological constant chosen in the $f(R)$ gravity model, in the limiting case of high curvature, the model will still reduce to GR, thus satisfying the many high-precision Solar System tests of GR.

Finally, we are able to relate the scalar field f_R to the function $f(R)$ using Equation 2.13, as outlined below:

$$f_R = -n\frac{c_1}{c_2^2}\frac{(-R/m^2)^{(n+1)}}{\left[1+(-R/m^2)^n\right]^2} \approx -n\frac{c_1}{c_2^2}\left(\frac{m^2}{-R}\right)^{(n+1)} \tag{2.13}$$

In the above equation, R has been approximated in the following manner:

$$-\bar{R} \approx 8\pi G\bar{\rho}_m - 2\bar{f}(R) = 3m^2\left[a^{-3} + \frac{2}{3}\frac{c_1}{c_2}\right] \tag{2.14}$$

In order for $f(R)$ to preserve the exact cosmological background expansion history of ΛCDM, this models sets $c_1/c_2 = 6\Omega_\Lambda/\Omega_m$. Here, Ω_m is the current fractional density of non-relativistic matter and Ω_Λ is the current fractional density of the cosmological constant for ΛCDM. With this, it is ensured that the Hu-Sawicki model of $f(R)$ gravity reproduces the ΛCDM background expansion history. Additionally, in this thesis, we assume that $n = 1$, which is the choice adopted most commonly. Finally, in the present epoch, the Hu-Sawicki $f(R)$ parameter becomes f_{R0}, which is defined in Equation 2.15 below:

$$f_{R0} \equiv f_R(\ln a = 0) \ll 1 \tag{2.15a}$$

$$f_{R0} \approx -n\frac{c_1}{c_2^2}\left(\frac{12}{\Omega_m} - 9\right)^{-n-1} \tag{2.15b}$$

The absolute value of this parameter $|f_{R0}|$ defines the efficiency of the chameleon mechanism in the $f(R)$ gravity model.

2.2 The Simulations

The chameleon mechanism is critical for maintaining the viability of this theory of modified gravity. In fact, linear theory would overpredict the effect of modified gravity on small scales, precisely because it would not capture the chameleon screening. As a result of this, the $f(R)$ model is inherently nonlinear, and depending on the density of the region of the Universe in question, gravity will behave very differently. This variation in the strength of gravity causes strong nonlinearities in the modified field equations, making analytical studies of the $f(R)$ modified gravity model extremely challenging. As a result, for this research we rely on large N-body cosmological simulations which self-consistently evolve dark matter particles according to the modified equations of motion.

In this thesis, we use the Extended LEnsing PHysics using ANalaytic ray Tracing (ELEPHANT) simulations (Cautun et al. 2018). These simulations were originally performed using the ECOSMOG code (Li et al. 2012), which itself is based on the simulation code RAMSES (Teyssier 2002). Within the simulations, the dark matter halos are identified using the publicly-available ROCKSTAR halo finder code (Behroozi et al. 2012). The exact parameters and technical specifications of these simulations are included in Table 2.1. We use F4, F5, and F6 to refer to a Hu-Sawicki $f(R)$ parameter $|f_{R0}|$ of 10^{-4}, 10^{-5}, and 10^{-6} respectively. Additionally, every dark matter halo in each simulation contains the parameters listed in Table 2.2, such as virial mass, velocity, root mean square velocity, virial radius, and position.

In creating the mock galaxy survey, we apply realism cuts that mimic parameters from the LOWZ sample of the 2015 Baryon Oscillation Spectroscopic Survey (BOSS)

Parameter	Physical Meaning	Value
Ω_m	present fractional matter density	0.281
Ω_Λ	$1 - \Omega_m$	0.719
h	$H_0/\left(100\text{kms}^{-1}\text{Mpc}^{-1}\right)$	0.697
n_s	primordial power spectral index	0.971
σ_8	r.m.s. linear density fluctuation	0.820
n	HS $f(R)$ parameter	1.0
$\lvert f_{R0}\rvert$	HS $f(R)$ parameter	10^{-6}, 10^{-5}, 10^{-4}
L_{box}	simulation box size	1024 Mpc/h
N_{p}	simulation particle number	1024^3
m_{p}	simulation particle mass	7.78×10^{10} $\text{M}_\odot/h$
N_{dc}	domain grid cell number	1024^3
N_{ref}	refinement criterion	$8, 8, 8, 8, 8, 8, 8, 8 \ldots$

Table 2.1:: The various parameters and technical specifications of the $f(R)$ and GR large cosmological N-body dark matter simulations used in this thesis.

Parameter	Physical Meaning	Units
M_{200}	the virial mass	$M_{\odot}$
$V_{\rm max}$	the maximum velocity	Mpc/h/s
$V_{\rm RMS}$	the root mean square velocity	Mpc/h/s
R_{200}	the virial radius	Mpc/h
R_s	the scale radius	Mpc/h
$c_{\rm NFW}$	the NFW profile concentration parameter	–
X	x-coordinate	Mpc/h
Y	y-coordinate	Mpc/h
Z	z-coordinate	Mpc/h
V_X	x-component of the velocity	Mpc/h/s
V_Y	y-component of the velocity	Mpc/h/s
V_Z	z-component of the velocity	Mpc/h/s

Table 2.2:: The various parameters defined for each halo in the $f(R)$ and GR large cosmological N-body dark matter simulations used in this thesis.

Data Release, which is part of the Sloan Digital Sky Survey III (SDSS-III) (Reid et al. 2016). The LOWZ sample has a redshift range of $0.15 < z < 0.43$, the median of which is $\bar{z} = 0.34$. As a result, in this thesis we focus on two snapshots from the ELEPHANT simulations, one at $z = 0$ for F4, F5, and GR, and one at $z = 0.34$ for F5, F6, and GR. When computing the halo mass function (HMF) and two-point 3D real-space halo correlation function, we include an analysis of all datasets at both redshifts. This is mostly to demonstrate how strong the maximum difference between the $f(R)$ gravity model and GR can be. However, because F4 deviates so strongly from GR and is in contention with a number of observational constraints (Lombriser 2014), we focus solely on F5, F6, and GR at $z = 0.34$ when using the Halo Occupation Distribution (HOD) model to populate dark matter halos with galaxies in Chapter 3.

2.2.1 The Halo Mass Function

This section discusses the HMF for F4, F5, and GR at $z = 0$, depicted in Figure 2.1, as well as the HMF for F5, F6, and GR at $z = 0.34$, depicted in Figure 2.2. The HMF plots the number density of halos at a certain mass, and in both of these figures the grey, shaded region denotes a "resolution limit", wherein the number of halos at low masses begin to significantly drop due to finite mass resolution in the simulation. In other words, if the N-body simulations included enough low-mass halos, the HMF would monotonically increase for the entirety of the plot, rather than displaying irregular behavior in the beginning. Additionally, although there are no error bars on this figure, the errors are so small that they can be ignored.

In Figures 2.1 and 2.2, we calculate the HMF for 408,022 of the most massive halos

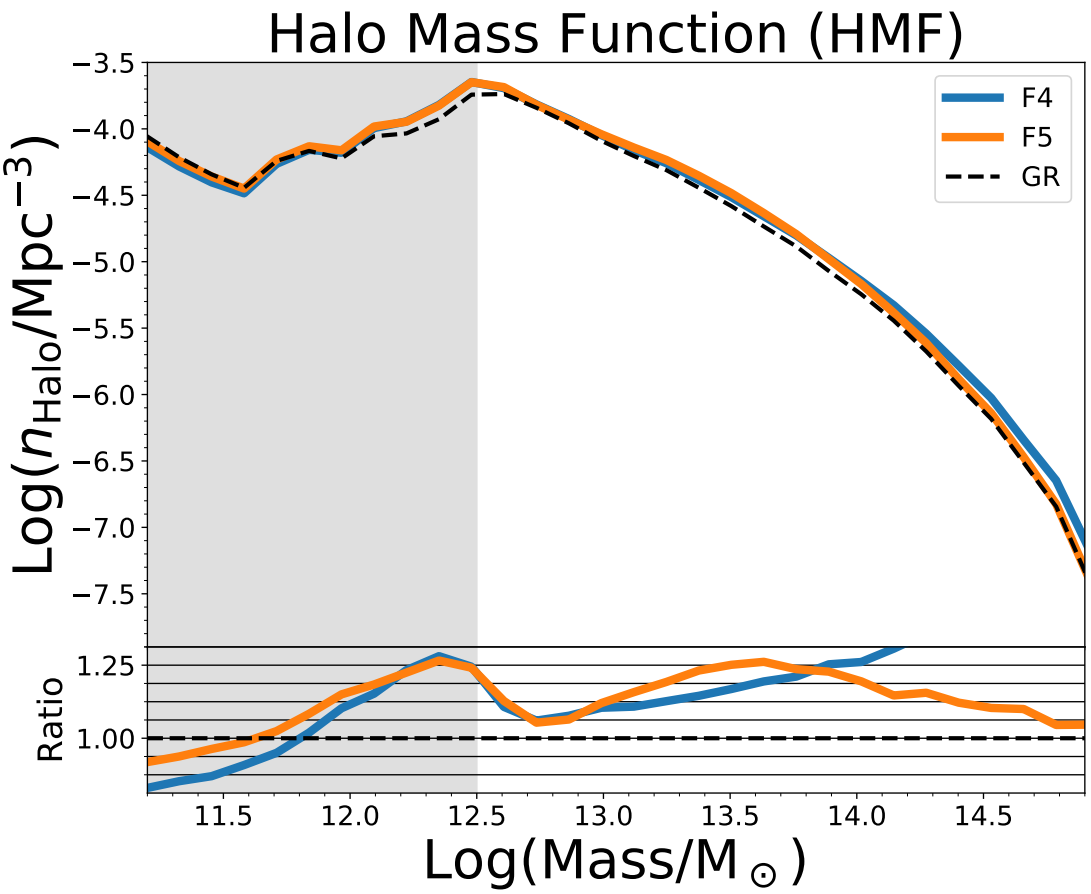

Figure 2.1: *Top Panel*: The halo mass function (HMF) at $z = 0$ for dark matter halos in F4 (blue), F5 (orange), and GR (black). The grey shaded region denotes a "resolution limit", where the number of low-mass halos begins to significantly drop due to finite mass resolution in the simulation. *Bottom panel*: The ratios of the HMF of F4 and F5 to GR.

in each simulation, ensuring a target number density of 3.8×10^{-4} halos/$(\mathrm{Mpc}/h)^3$. This target was chosen in order to closely mimic the number density of galaxies in the LOWZ catalogue of BOSS, which hovered around 10^{-4} galaxies/$(\mathrm{Mpc}/h)^3$ (Reid et al. 2016). We purposefully do not randomly select halos: while that is a theoretically valid comparison, it would not be the most meaningful one to make. This is due to the fact that since the halo population is dominated by low-mass halos, a random selection of 408,022 objects preferentially selects more low mass halos. In a true galaxy survey, on the other hand, which compiles a galaxy sample of fixed number density, we are instead more likely to observe the brightest galaxies, and correspondingly most massive dark matter halos.

Thus, we choose to bring our comparison more in line with what might be done in a galaxy survey by selecting a population at fixed abundance. In other words, we calculate the HMF for a population of halos which has first been sorted in descending order of halo mass. Doing so leads to Figure 2.1, where we can see a clear distinction between F4, F5, and GR. Due to the nature of their screening mechanisms, the HMF of F4 is greater than that of F5, which is in turn greater than that of GR. In examining F4, it is important to keep in mind that the screening mechanism is inefficient on all scales. This means that the enhancement of gravity is not suppressed, so that dark matter halos will merge and grow in size. As a result, dark matter halos in both over-dense (surrounded by many other halos) and under-dense (located in a void-like area) regions of the Universe are more massive than their counterparts in GR. Because this, there are less low-mass halos than there would be in GR, and the HMF for F4 becomes shifted to the right in comparison.

On the other hand, examining F5 leads to a slightly different interpretation, as its screening mechanism is more efficient than that of F4. This means that it is

actually inefficient only in under-dense regions of the Universe, and is roughly equal to GR in over-dense regions of the Universe. However, it is not equivalent to GR in these over-densities, meaning that the enhancement of gravity is still not completely suppressed. Thus, dark matter halos in over-dense regions are more massive in F5 than their counterparts in GR, but are still of a lower mass than those in F4, because the screening mechanism is not as weak. This relationship is visible in Figure 2.1, as the HMF for F5 is shifted to the right of that for GR, but to the left of the HMF for F4. In Figure 2.2, we are able to examine the HMF for F5, F6, and GR at $z = 0.34$. Exactly as in Figure 2.1, which depicts a snapshot at $z = 0$, the HMF for F5 is shifted to the right of the HMF for GR. It is encouraging that these comparisons remain consistent between the two snapshots, as our consequent analysis should apply to present-day as well.

Now, in examining the HMF for F6, it is important to note that the screening mechanism is at its most efficient of all three considered $f(R)$ models. As a result, halos in over-dense regions of the Universe in F6 are nearly exactly equal to their counterparts in GR. However, this also means that these halos are less massive than their counterparts in F5, which still undergo some enhancement of gravity in over-dense regions. On the other hand, halos in under-dense regions of the Universe in F6 continue to experience a stronger, more enhanced degree of Newtonian Gravity. Because of this, these halos become more massive than their counterparts in GR. This relationship is visible in the bottom panel of Figure 2.2, where a slight increase in the HMF of F6 compared to GR can be seen. At about 10^13 $M_\odot$, however, the HMF for both GR and F6 appear to become roughly identical, signifying that this cutoff is likely the point at which high-mass halos in F6 become dominated by halos in over-dense regions, rather than under-dense regions. Furthermore, it is clear that the HMF for F6 consistently remains below that of

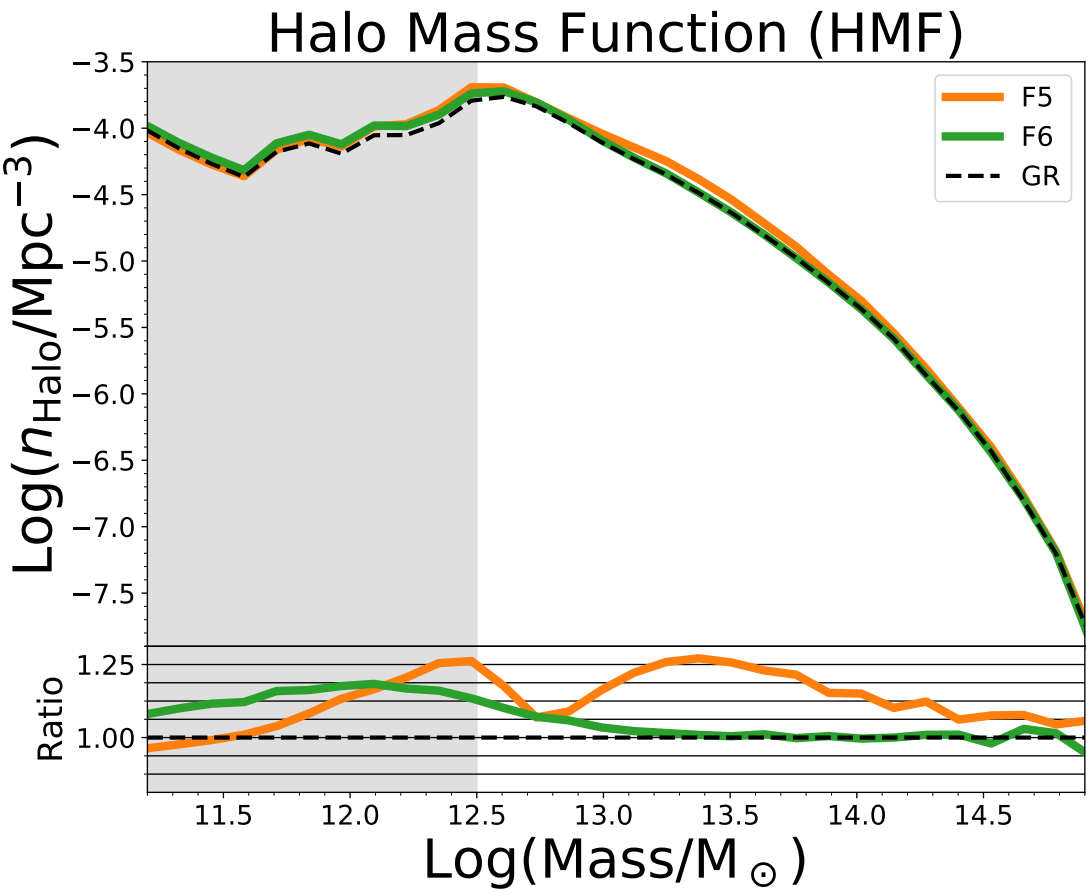

Figure 2.2: *Top Panel*: The halo mass function (HMF) at $z = 0.34$ for dark matter halos in F5 (orange), F6 (green), and GR (black). The grey shaded region denotes a "resolution limit", where the number of low-mass halos begins to significantly drop due to finite mass resolution in the simulation. *Bottom panel*: The ratios of the HMF of F5 and F6 to GR.

F5, as expected, due to the comparative efficiency of the screening mechanism.

2.2.2 The 3D Halo Correlation Function

To measure the large-scale clustering of halos in these simulations, we use the spatial two-point correlation function $\xi(r)$. This function describes the clustering properties of galaxies in three dimensions as a function of pair separation and is defined in Equation 2.16 below:

$$N(r)\mathrm{d}V = N_0[1 + \xi(r)]\mathrm{d}V \tag{2.16}$$

Here, $N(r)\mathrm{d}V$ represents the number of galaxies in the volume element $\mathrm{d}V$ at distance r from any galaxy, N_0 represents the average number density of galaxies in space, and $\xi(r)$ represents the excess number of galaxies at a distance r away from each galaxy. In this thesis, we use the publicly available Python code CORRFUNC (Sinha & Garrison 2020) to carry out both the two-point 2D projected correlation function $w_p(p)$ and the two-point 3D real-space correlation function $\xi(r)$. In CORRFUNC, $\xi(r)$ is defined using the "natural estimator", or Landy-Szalay formula, as described by Equation 2.17 below (Landy & Szalay 1993):

$$\xi(r) = \frac{\mathrm{DD}}{\mathrm{RR}} - 1 \tag{2.17}$$

Here, DD refers to the total number of "galaxy-galaxy" pairs. However, in order to compare to the galaxy data, the correlation function naturally requires a term which concerns random data. Thus, RR refers to the total number of pairs given a completely

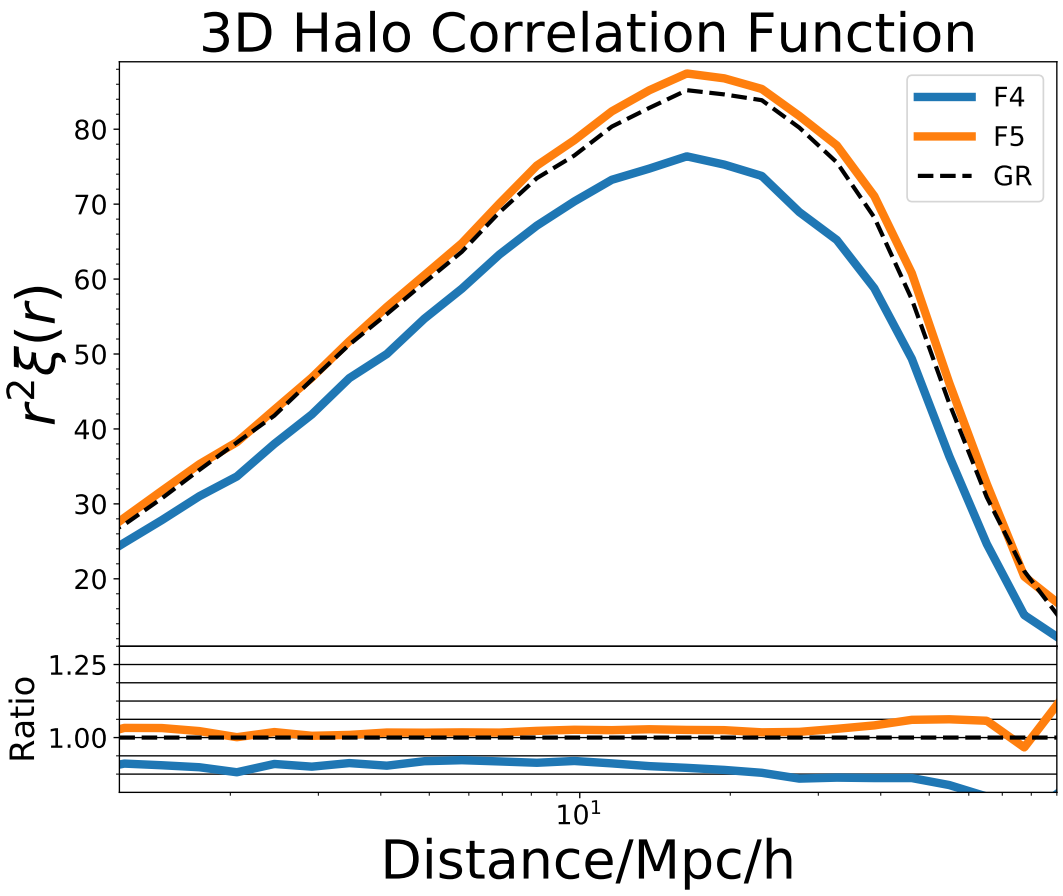

Figure 2.3: *Top Panel*: The two-point 3D real-space correlation function $r^2\xi(r)$ at $z = 0$, calculated at fixed abundance for dark matter halos in F4 (blue), F5 (orange), and GR (black). 408,022 halos selected in order of descending mass in a (1024 Mpc/h)^3 volume box make a number density of 3.8×10^{-4} halos/(Mpc/h)^3. *Bottom panel*: The ratios of $r^2\xi(r)$ of F4 and F5 to GR.

random dataset with the same number density as that of the galaxies. This term can be computed analytically using Equation 2.18 below, where RR_i denotes the expected number of random pairs in bin i, N the total number of points, V_i the volume of bin i, $\bar{\rho}$ the average density of the entire simulation box, and L the box length:

$$\mathrm{RR}_i = NV_i\bar{\rho} = NV_i\frac{N-1}{L^3} \tag{2.18}$$

The two-point 3D real-space correlation function is depicted in Figure 2.3 for F4, F5, and GR at $z = 0$. In order to emphasize the differences in this clustering statistic on all scales, $r^2\xi(r)$ is plotted, rather than $\xi(r)$. Like the HMF, the correlation function is plotted for 408,022 halos at fixed abundance, achieving a target number density of 3.8×10^{-4} halos/$(\mathrm{Mpc}/h)^3$. As discussed in Section 2.2.1, because the screening mechanism for F4 is inefficient on all scales, the HMF is shifted to the right of that for GR. This inefficient screening mechanism also means that gravity is enhanced in over-dense regions of the Universe, leading to an increased number of dark matter halo mergers.

The direct consequence of such mergers is depicted in Figure 2.4: where there are a cluster of low-mass dark matter halos in GR, their counterparts in F4 will merge together into a few isolated high-mass halos. As a result, an over-dense region of the Universe in GR can become an under-dense region in F4, causing an overall shift towards less clustering of dark matter halos. This effect is exacerbated by the enhancement of gravity in under-dense regions of the Universe, where halos in F4 become larger than their counterparts in GR. This shift in mass is significant: because the correlation function is calculated using fixed abundance, more under-dense halos are sampled in F4, therefore

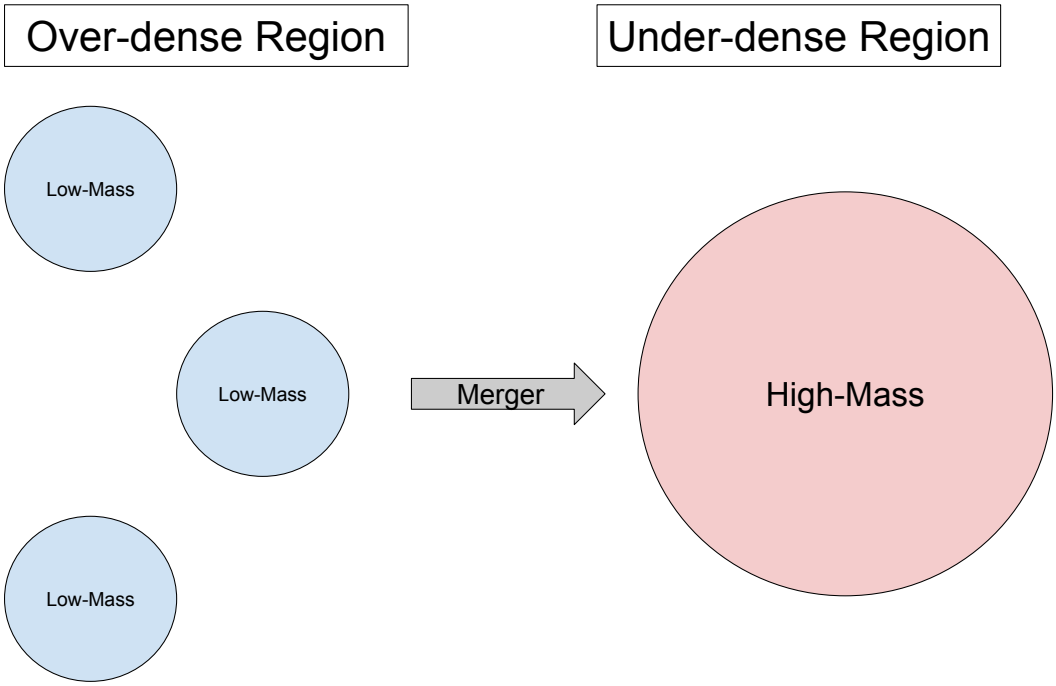

Figure 2.4: *Left*: A cluster of low-mass dark matter halos in GR, located in an over-dense region of the Universe. *Right*: Due to the inefficient screening mechanism in F4, over-dense regions will undergo an enhancement of gravity. Thus, these halos will merge together into just one high-mass halo, located in what is now an under-dense region of the Universe.

causing the dramatic drop in clustering displayed by F4 in Figure 2.3.

On the other hand, F5 appears to have slightly stronger clustering than GR. This is because, as discussed in Section 2.2.1, the screening mechanism for F5 is such that gravity is roughly equal to GR in over-dense regions, but still not exact. This slight enhancement of Newtonian Gravity means that for F5, there are a higher number of high-mass dark matter halos located in over-dense regions of the Universe. Thus, just as the HMF is shifted to the right of that for GR, fixed abundance means that the correlation function for F5 will be shifted slightly upwards of that for GR.

In Figure 2.5, the two-point 3D real-space correlation function $r^2\xi(r)$ is depicted for F5, F6, and GR at $z = 0.34$. 408,022 halos are plotted at fixed abundance for a target number density of 3.8×10^{-4} halos/(Mpc/h)^3. As in Figure 2.5, we see that F5 displays comparatively more clustering than GR. However, F6 counter-intuitively appears less clustered than GR, albeit less so than F4, even though it is the closest to GR out of all three $f(R)$ models. In fact, this subversion is a direct result of the efficiency of the screening mechanism found in F6. As discussed in Section 2.2.1, F6 has the most efficient screening mechanism out of all the considered $f(R)$ models. As a result, its gravity is exactly equal to that of GR in over-dense regions of the Universe. However, because there is still an enhancement of gravity in under-dense regions, these halos become larger than their counterparts in GR. Thus, with fixed abundance we find that F6 samples more halos in under-dense regions of the Universe than GR. As such, F6 desplays slightly less clustering than GR, but still more so than F4, as displayed in Figure 2.5. It is encouraging that these results are all consistent with each other, and that their physical interpretation is consistent with our expectation of the relative behavior of F4, F5, F6, and GR.

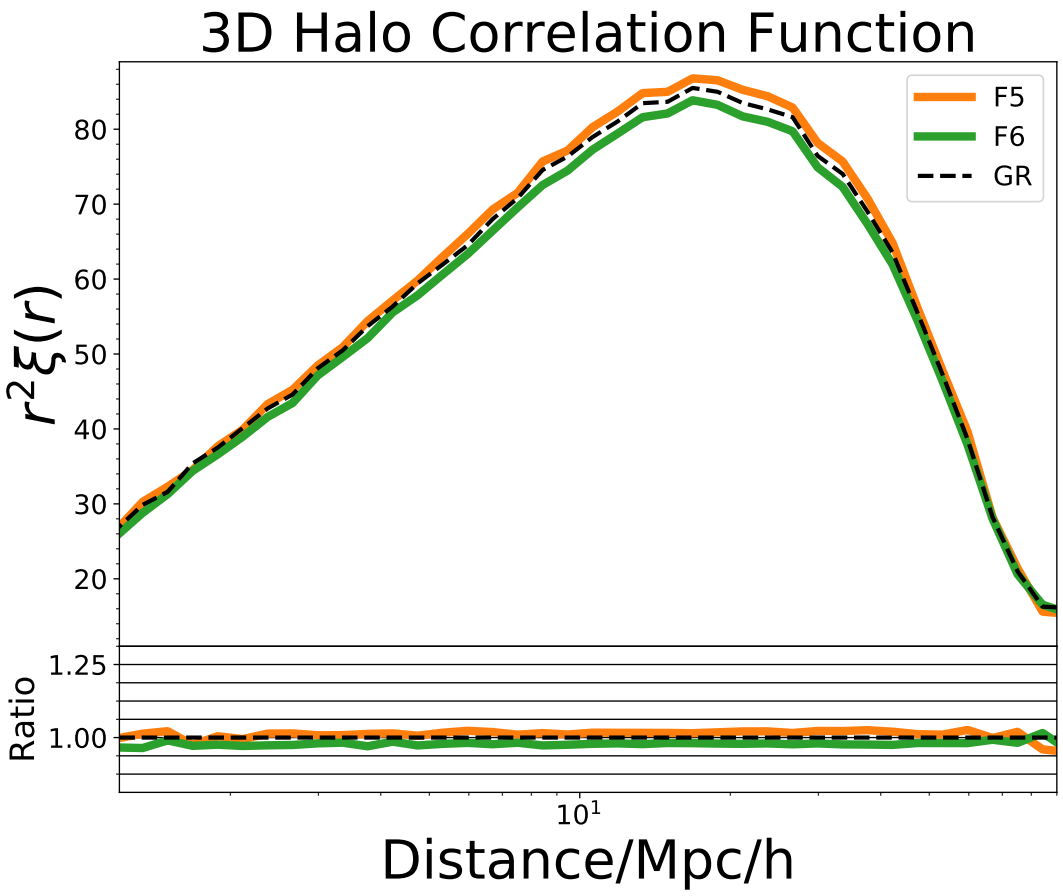

Figure 2.5: *Top Panel*: The two-point 3D real-space correlation function $r^2\xi(r)$ at $z = 0.34$, calculated at fixed abundance for dark matter halos in F5 (orange), F6 (green), and GR (black). 408,022 halos selected in order of descending mass in a $(1024\ \mathrm{Mpc}/h)^3$ volume box make a number density of 3.8×10^{-4} halos/$(\mathrm{Mpc}/h)^3$. *Bottom panel*: The ratios of $r^2\xi(r)$ of F5 and F6 to GR.

Chapter 3

The Halo Occupation Distribution Model

3.1 Populating Dark Matter Halos with Galaxies

In this chapter, we move away from simply examining the $f(R)$ gravity model at the halo level, and finally begin populating the dark matter halos with galaxies. From this point on, we no longer consider F4, and instead focus only on F5, F6, and GR at $z = 0.34$. To construct mock galaxy catalogues from these dark matter only simulations, we implement the HOD model. In order to discuss the HOD model, it is first important to become familiar with the structure of a dark matter halo, depicted in Figure 3.1, as well as the corresponding terminology. We assume that every dark matter potentially halo encompasses a central galaxy which is located at its center. In addition to this central galaxy, there may be smaller gravitationally-bound halos located within the radius of this larger dark matter "host" halo. These virialized clumps of dark matter, or subhalos,

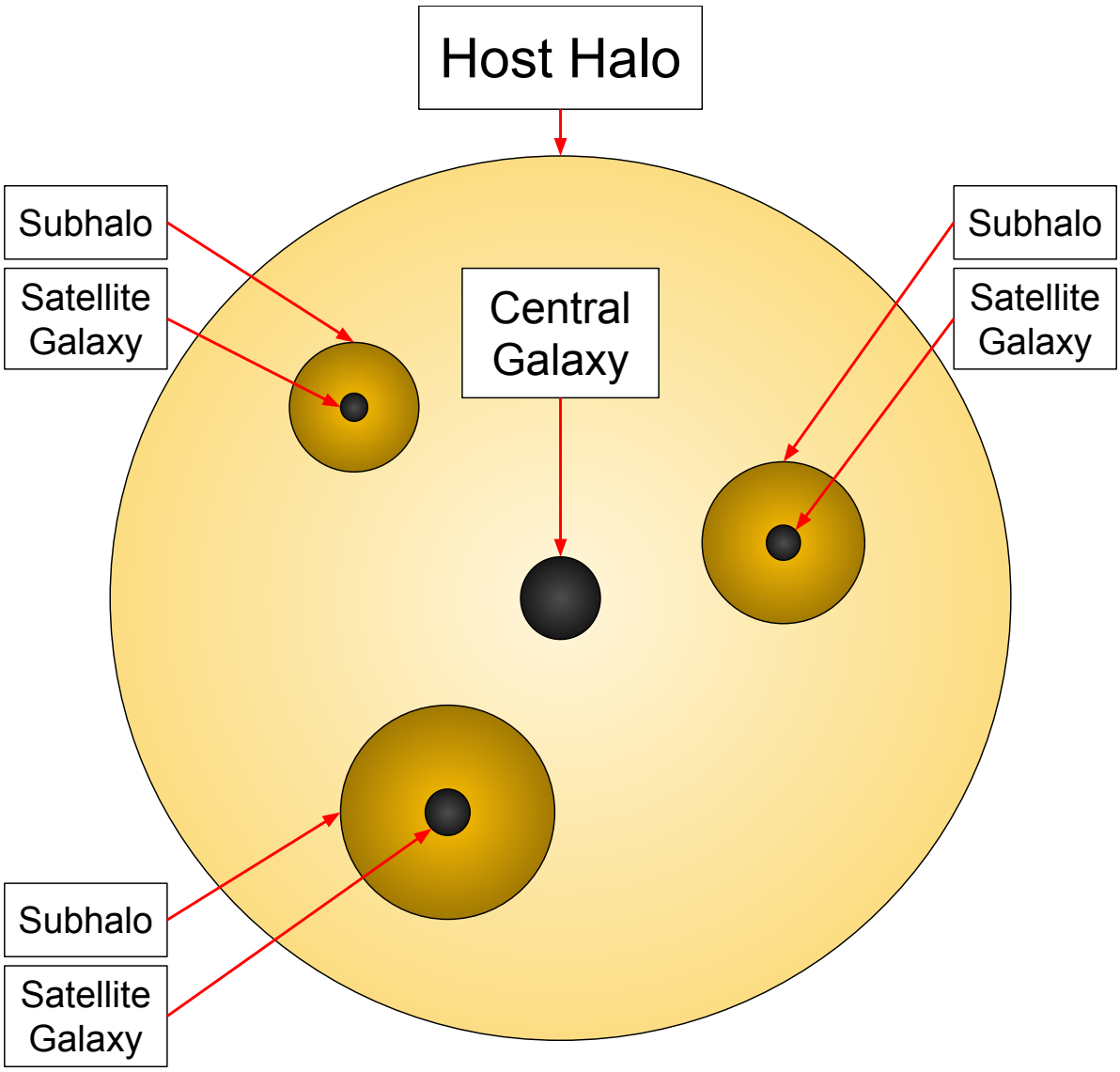

Figure 3.1: A dark matter "host" halo, which encompasses a central galaxy at its center, as well as many smaller subhalos. Each gravitationally-bound clump of dark matter, or subhalo, surrounds a satellite galaxy which is situated at its center.

thus orbit within the gravitational potential of the "host" halo. If a subhalo exists, it then encompasses its own satellite galaxy, situated at its center.

Thus, the HOD model is a probability distribution which describes the likelihood for a halo of mass M to host a number $N_{\rm cen}$ and $N_{\rm sat}$ of central and satellite galaxies. The result is an HOD model which can be described by just five basic parameters (Zheng et al. 2005), as shown in Equation 3.1 below:

$$\langle N_{\rm cen}\,(M)\rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm min}}{\sigma_{\log M}}\right)\right] \tag{3.1a}$$

$$\langle N_{\rm sat}(M)\rangle = \langle N_{\rm cen}\rangle \left(\frac{M - M_0}{M_1}\right)^{\alpha} \tag{3.1b}$$

Here, $\langle N_{\rm cen}\,(M)\rangle$ and $\langle N_{\rm sat}\,(M)\rangle$ refer to the mean occupation distributions for central and satellite galaxies, respectively, for a dark matter halo of mass M. The five basic HOD parameters are thus $M_{\rm min}$, $\sigma_{\log M}$, M_0, M_1, and α, each of which operate according to the functionality described in Table 3.1.

3.2 Achieving the Target Number Density

In order to create a suitable mock galaxy survey, we must ensure that the galaxy population achieves the same number density profile as the LOWZ sample from BOSS, part of SDSS-III, as depicted in Figure 3.2. To do this, we must first upscale the number of mock galaxies by setting a target number density of $\sim 8.8 \times 10^{-4}$ galaxies/$({\rm Mpc}/h)^3$. Doing so will assure that the final desired number density is still achieved, even after the appropriate survey realism cuts, such as sky completeness and redshift selection, are

HOD Parameter	Physical Meaning
$M_{\rm min}$	the minimum mass of halos that can host a central galaxy
$\sigma_{\log M}$	the characteristic width of this transition
M_0	the truncation mass of halos that can host a satellite galaxy
M_1	a normalization factor
α	the power-law slope

Table 3.1:: The five basic HOD parameters which describe the mean occupation function for central and satellite galaxies, respectively $\langle N_{\rm cen}\,(M)\rangle$ and $\langle N_{\rm sat}\,(M)\rangle$, as well as their physical meanings.

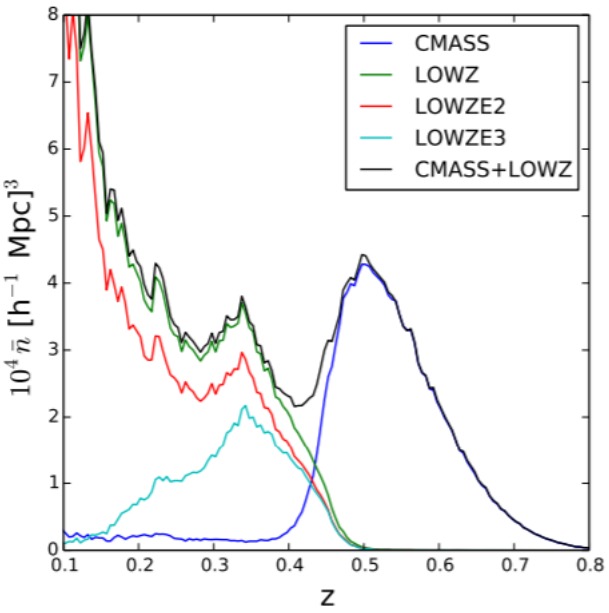

Figure 3.2: Number density as a function of redshift z for the LOWZ sample (green) from BOSS, part of SDSS-III, as well as three other samples. Figure reproduced from Reid et al. (2016).

applied in the final stages of the thesis. In addition to achieving this target number density, we must ensure that both the number density and projected clustering of F5, F6, and GR match as closely as possible to each other. The latter process is discussed in Section 3.4.

3.2.1 Finding the Optimal $M_{\rm min}$

In order to attain this target number density, we must determine the optimal values for the five basic HOD parameters outlined in Table 3.1. For ease of computation, and because we will later vary all five parameters, at this stage we choose to vary just one parameter, $M_{\rm min}$. Accordingly, we assign fixed values to the four parameters $\sigma_{\log M}$, α, fM_0, and fM_1. The latter two terms fM_0, and fM_1 are individual dimensionless parameters that can be defined as the ratios in Equation 3.2 below:

$$fM_0 = \frac{M_0}{M_{\rm min}} \tag{3.2a}$$

$$fM_1 = \frac{M_1}{M_{\rm min}} \tag{3.2b}$$

The fixed values of these four HOD parameters, as previously determined by Manera et al. (2013), are listed in Table 3.2. Because of how we have defined fM_0, and fM_1 in Equation 3.2, by simply varying $M_{\rm min}$, we are able to effectively test different values of M_0 and M_1 as well. In order to ascertain the optimal value of $M_{\rm min}$, we create a possible parameter space with 5000 possible points ranging from 10^{11} $\rm M_\odot$ to 10^{15} $\rm M_\odot$. For each of these possible solutions, we determine the number of central and satellite galaxies

HOD Parameter	Fixed Value
σ	0.596
α	1.0127
fM_0	0.97
fM_1	8.13

Table 3.2:: The remaining four HOD parameters, fixed at the values determined by Manera et al. (2014) for F5, F6, and GR.

Simulation	$\log(M_{min})$	n_{gal} (10^{-4} galaxies/(Mpc/h)^3)
F5	12.67	8.702
F6	12.60	8.708
GR	12.57	8.707

Table 3.3:: The optimal M_{min} HOD parameter for F5, F6, and GR which most closely matches the target number density $n_{target} = 8.8 \times 10^{-4}$ galaxies/(Mpc/h)^3. Fixing the other HOD parameters at the values determined by Manera et al. (2014) gives the above number densities.

which reside in each dark matter halo. Finally, we use interpolation to select the optimal value of $M_{\min}$ that will achieve the desired target number density. Carrying out this optimization procedure for F5, F6, and GR allows us to obtain the corresponding values for $M_{\min}$, the log of which appears in Table 3.3.

This table also includes the resulting galaxy number densities of each simulation, which, once we determined these optimal $M_{\min}$ values, we were easily able to compute using Equation 3.1. As desired, the results are extremely close to the desired target number density of 8.8×10^{-4} galaxies/$(\mathrm{Mpc}/h)^3$, and are even closer to each other, as displayed by Figure 3.3. Here, we have plotted the number of central, satellite, and total galaxies found in each dark matter halo as a function of its mass. The 16th and 84th percentiles of each of these three curves are also plotted for F5, F6, and GR.

As expected, the curve for central galaxies begins at very low values of approximately zero, and eventually rises to achieve the maximum value of one central galaxy per halo. Similarly, the curve for the number of central galaxies begins at very low values of approximately zero, and progressively rises. Note that no halos which host satellite galaxies are smaller in mass than $M_{\min}$, as it is impossible for a dark matter halo to host a satellite galaxy if it cannot even host a central galaxy. This cutoff point can be more clearly seen in Figure 3.4. Examining these plots thus makes it abundantly clear that the number of galaxies are extremely close for each of these three simulations, as they all fall within the 16th and 84th percentiles of each curve. Consequently, our goal to match the target number density across all three mock galaxy surveys is achieved at this intermediary step.

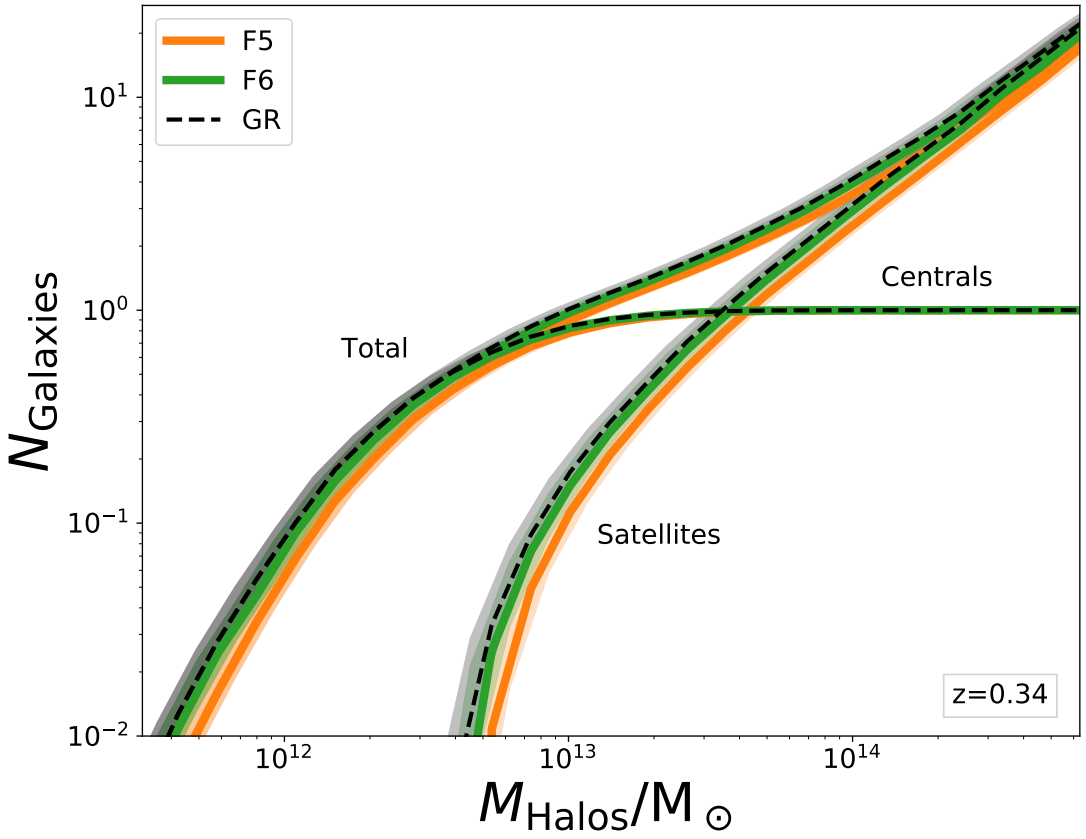

Figure 3.3: The number of galaxies as a function of the mass of each halo in F5 (orange), F6 (green), and GR (black). The furthermost left curve depicts all galaxies, the middle curve satellite galaxies, and the furthermost right curve central galaxies. The 16th and 84th percentiles of the number of galaxies are shaded in for each curve.

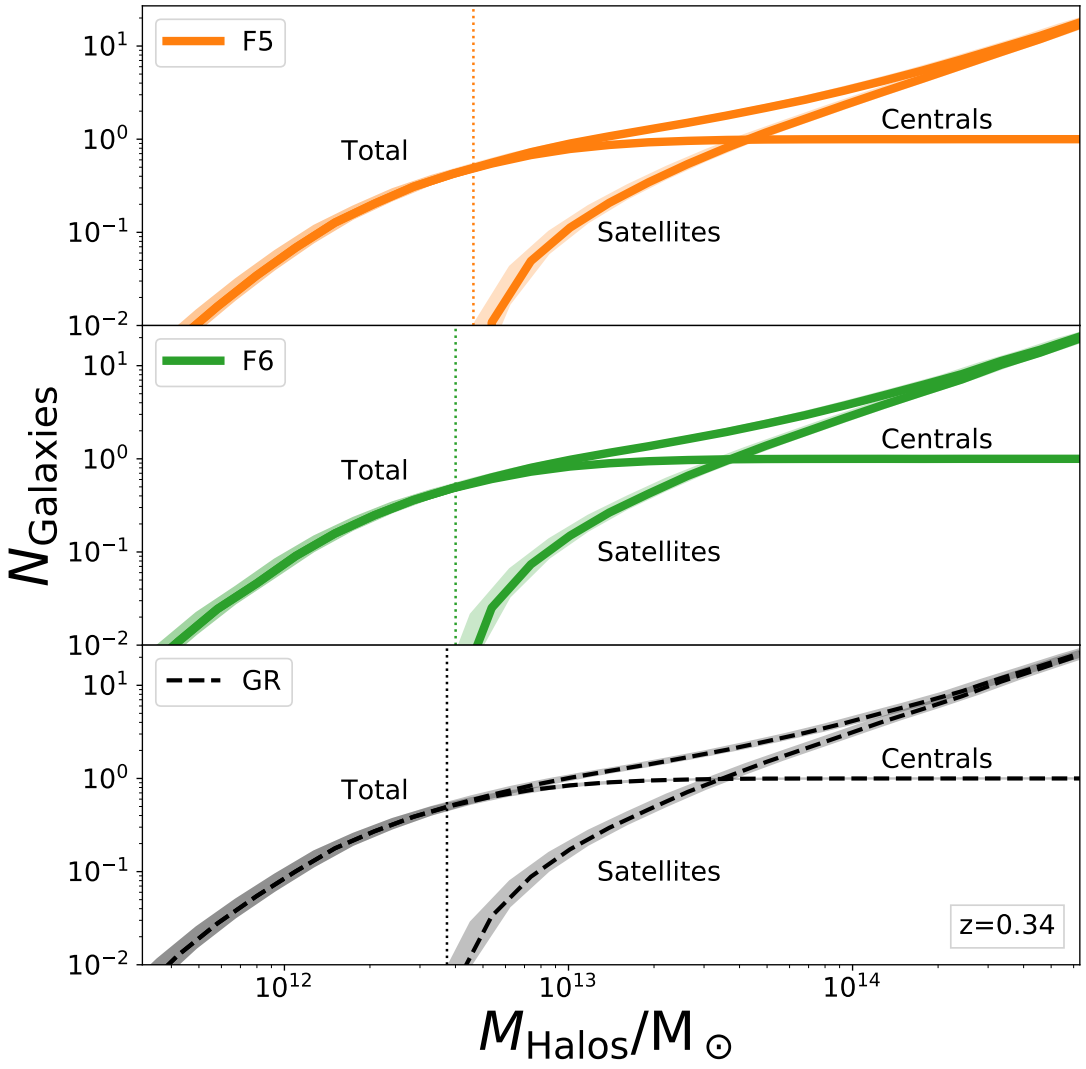

Figure 3.4: The number of galaxies as a function of the mass of each halo in F5 (orange), F6 (green), and GR (black). The furthermost left curve depicts all galaxies, the middle curve satellite galaxies, and the furthermost right curve central galaxies. The optimal value of M_{min} which achieves the target number density is plotted as a vertical dotted line for each simulation, and the 16th and 84th percentiles are shaded around each curve.

3.3 Assigning the Spatial Distribution of HOD Satellite Galaxies

Now that the optimal values of $M_{\min}$ are determined for F5, F6, and GR, and the dark matter halos are populated with the correct number of galaxies to obtain the desired target number density, we must assign positions to each of the mock galaxies. This will enable us to compute the two-point 2D projected correlation function $w_p(r)$, which we later hold as fixed for GR, and match with F5 and F6 by varying their HOD parameters in Section 3.4.

3.3.1 The Navarro–Frenk–White Profile

Assigning positions to central galaxies is extremely straightforward, as they will simply be located at the center of their host halo. Assigning positions to each mock satellite galaxy created is a little more complicated, however. In order to accomplish this, we use the Navarro-Frenk-White (NFW) profile, which describes the equilibrium density profile of cold dark matter halos in N-body cosmological simulations (Navarro et al. 1997). This density profile can be described as a function of radius r by Equation 3.3 below:

$$\rho(r) = \frac{\rho_0}{\frac{r}{R_s}\left(1 + \frac{r}{R_s}\right)^2} \tag{3.3}$$

Here, $\rho(r)$ refers to the mean interior matter density of a cold dark matter halo and R_s refers to the scale radius of the halo. Similarly, ρ_0 is a halo-specific density parameter related to the characteristic dimensionless density δ_c and the critical density of the

universe $\rho_c(z)$, which is a function of redshift z, as described by Equation 3.4 below:

$$\rho_0 = \delta_c \, \rho_c \tag{3.4a}$$

$$\rho_c = \frac{3H^2(z)}{8\pi G} \tag{3.4b}$$

Here, $H(z)$ refers to the Hubble parameter and G to the Newtonian constant of gravitation. Plotting the NFW profile using Equation 3.3 produces Figure 3.5, which depicts the characteristic matter density profile ρ of a dark matter halo as a function of q. Here, q refers to the fraction of radius r from the halo's "virial" radius R_{200}. Because the density of an equilibrium halo is proportional to the density of the universe at the time of its formation, the virial radius R_{200} of a halo is defined as the radius which achieves an interior mean matter density of $200\rho_c$ (Kravtsov 2013). As expected, the density profile depicts a gradually-changing logarithmic slope which tends towards r^{-3} for $r \gg R_s$.

3.3.2 Creating a Mock Galaxy Catalogue

In order to assign positions to all of the mock galaxies and create a cohesive catalogue, we assume that the NFW profile perfectly describes their location in terms of the radius r, which spans from the center of each host halo to the center of a mock satellite galaxy. However, we must first rewrite Equation 3.4 in terms of the known parameters listed in Table 2.2 in the Appendix. To accomplish this, we first determine the total mass M of a dark matter halo enclosed within some radius r using Equation 3.5 below:

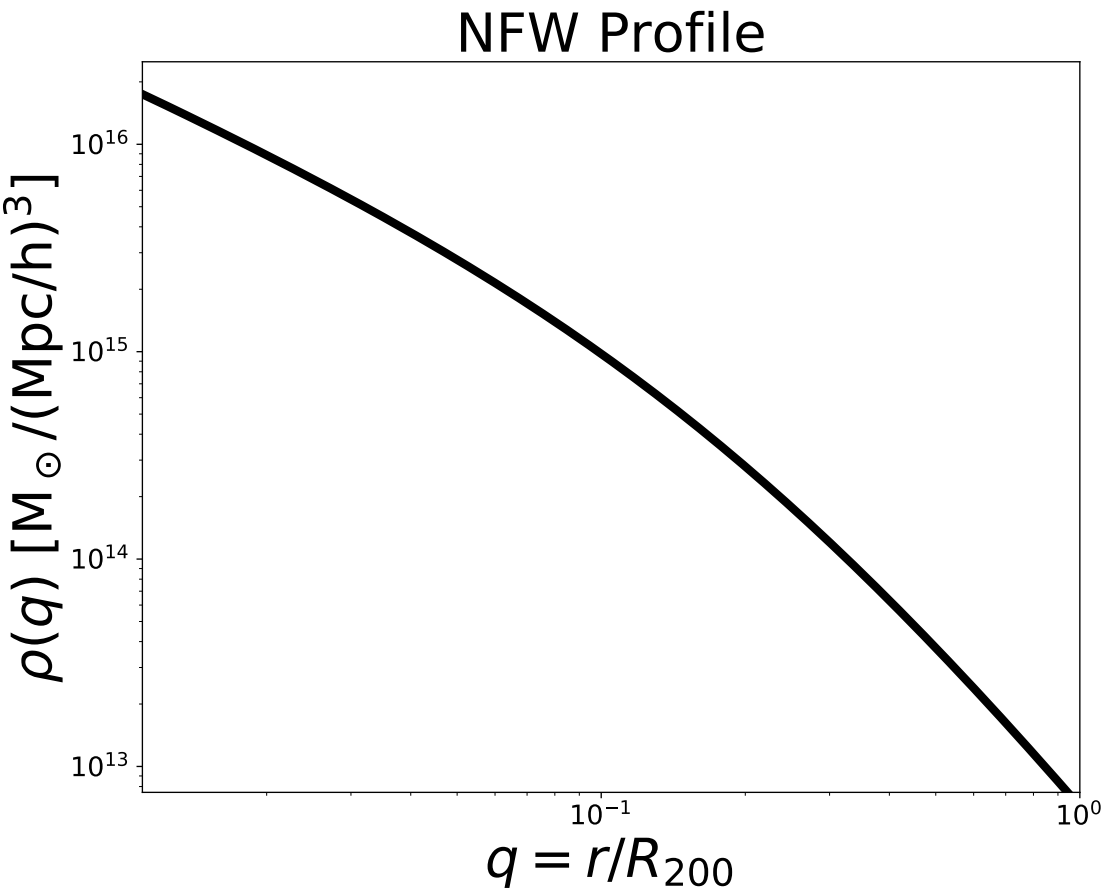

Figure 3.5: The Navarro-Frenk-White profile, which descibes the equilibrium density profile of cold dark matter halos in N-body cosmological simulations. Here it is plotted as a function of a fraction q of the virial radius R_{200}. This figure depicts the expected density profile for the halo in GR with the largest mass $M_{200} = 1.479 \times 10^{15}$ $M_\odot$ at $z = 0.34$, as well as a concentration parameter $c_{\rm NFW} = 4.753114$ and scale radius $R_s = 0.501447$.

$$M = \int_V \rho dV = \int_0^{R_{\max}} 4\pi r^2 \rho(r) dr \tag{3.5a}$$

$$M = 4\pi\rho_0 R_s^3 \left[\ln\left(\frac{R_s + R_{\max}}{R_s}\right) - \frac{R_{\max}}{R_s + R_{\max}} \right] \tag{3.5b}$$

Here, dV is the differential volume element and R_{max} is the maximum point of integration along the radial axis, referring to the radius r at which the satellite galaxy in question is located. Consequently, $\rho(r)$ is identical to the density profile described by Equation 3.3, and refers to the mean enclosed matter density of the halo within some radius r. We are thus able to use this equation to solve for the density parameter ρ_0, using the fact that the virial radius R_{200} can be related to the scale radius R_s by some "concentration parameter" $c_{\rm NFW}$, as described by Equation 3.6 below:

$$R_{200} = c_{\rm NFW} R_s \tag{3.6}$$

The concentration parameter $c_{\rm NFW}$, like the scale radius R_s, is unique to each individual dark matter halo. Using this definition allows us to solve for the halo-specific density parameter ρ_0 in Equation 3.5 by first substituting R_{200} for $R_{\max}$, M_{200} for M, and $c_{\rm NFW}$ for R_{200}/R_s. This results in the following equation, described by Equation 3.7 below:

$$\rho_0 = \frac{M_{200}}{4\pi {R_s}^3 \big(\log(1 + c_{\rm NFW}) - \frac{c_{\rm NFW}}{1 + c_{\rm NFW}} \big)} \tag{3.7}$$

With this expression for the density parameter ρ_0, we can now compute the density profile $\rho(r)$ in Equation 3.3 in terms of the virial mass M_{200}, the scale radius R_s, and

the concentration parameter $c_{\rm NFW}$, all of which are well-defined constants for each halo in all the simulations, as displayed by Table 2.2 in the Appendix. Furthermore, because the virial radius refers to the radius at which the mean enclosed matter density is $200\,\rho_c$, the virial mass M_{200} can be defined in terms of the critical density of the Universe ρ_c and the virial radius R_{200}, as described by Equation 3.8 below:

$$M = \rho\, V \tag{3.8a}$$

$$V = \frac{4\pi R^3}{3} \tag{3.8b}$$

$$M_{200} = 200\rho_c(z)\,\frac{4\pi {R_{200}}^3}{3} \tag{3.8c}$$

With this new definition of the density parameter ρ_0 described by Equation 3.7 and the above definition of M_{200} in Equation 3.8, we can solve for the characteristic dimensionless density δ_c in Equation 3.4. The resulting expression is defined by Equation 3.9 below:

$$\delta_c = \frac{200}{3}\,\frac{c_{\rm NFW}^3}{\left(\log(1+c_{\rm NFW}) - \frac{c_{\rm NFW}}{1+c_{\rm NFW}}\right)} \tag{3.9}$$

3.3.3 Assigning Final Positions and Velocities to Mock Galaxies

With the halo-specific density parameter ρ_0 defined in Equation 3.7, we are finally able to use the NFW profile in Equation 3.3 to assign positions to all of the mock galaxies in each simulation. To do so, we generate an array of possible distances from the center

of the halo, ranging from $10^{-5}\,R_{200}$ to $10\,R_{200}$. We then generate a random mass that falls within M_{200} of each halo, and use interpolation to obtain the corresponding radial position along the NFW profile. Finally, we are generate random angular coordinates (θ, ϕ) for each galaxy within every halo, and use the newly obtained radial position to convert to Cartesian coordinates (X, Y, Z). In doing so, we shift the positions of the satellite galaxies so that they are offset by the location of their host halo. Because this offset may push a mock satellite galaxy beyond the confines of the simulation box length, we then correct these stray positions by taking advantage of the periodic nature of the box to "wrap" the satellite galaxy back to the other end of the simulation.

With this, the process of assigning positions to mock galaxies is complete. Now, we can assign velocities to the mock galaxies. Just as central galaxies share the same positions as their host halo, as described in Section 3.3.1 and depicted in Figure 3.1, determining the velocities of central galaxies is similarly straightforward. To do this, we simply assign mock central galaxies the same velocity as their host halo.

In parallel, just as assigning the positions of satellite galaxies is more complicated than assigning the positions of central galaxies, assigning the velocities of satellite galaxies is slightly more complex than doing the same for central galaxies. To accomplish this, we drew 2^{10} random samples from a normal Gaussian distribution that were centered around the velocity of the host halo, and had a standard deviation equivalent to the root mean square velocity V_{RMS} of the host halo. Therefore, for each component V_X, V_Y, and V_Z of the total velocity of each satellite galaxy, the standard deviation is $V_{RMS}/\sqrt{3}$. Finally, from these 2^{10} randomly sampled velocities, we then randomly selected exactly enough velocities to assign to the number of satellite galaxies in each host halo, which are determined according to the optimal HOD parameters calculated in

Section 3.4.3.

3.3.4 Intermediary Results

Using this method to assign positions and velocities to mock galaxies in F5, F6, and GR at redshift $z = 0.34$ produces the plot in Figure 3.6. Here, the dotted line depicts the expected NFW profile $\rho(r)$ for the most massive halo in each simulation, while the solid line in each panel displays the actual produced density profile. This graph makes clear that the resulting data does indeed follow the curve of the characteristic density profile $\rho(r)$ outlined by Equation 3.3, as originally intended.

Although these mock galaxy positions only achieve the intermediary step of matching the target number density of 8.8×10^{-4} galaxies/(Mpc/h)^3 as closely as possible, we have chosen to compute the projected clustering for comparison: Figure 3.7 depicts the two-point 2D projected correlation function $r^2 w_p(r)$ for F5, F6, and GR at redshift $z = 0.34$. In Section 3.4, we take into account the projected clustering as well as number density, finalize the mock galaxy positions, and examine the resulting correlation function.

3.4 Matching the Projected Clustering

Thus far, the positions represented in Figure 3.6 have only satisfied one criterion of the final mock galaxy catalogue, in that they achieve the target number density 8.8×10^{-4} galaxies/(Mpc/h)^3. This section discusses the second goal of also achieving an equivalent projected correlation function across F5, F6, and GR. To do this, we hold the HOD

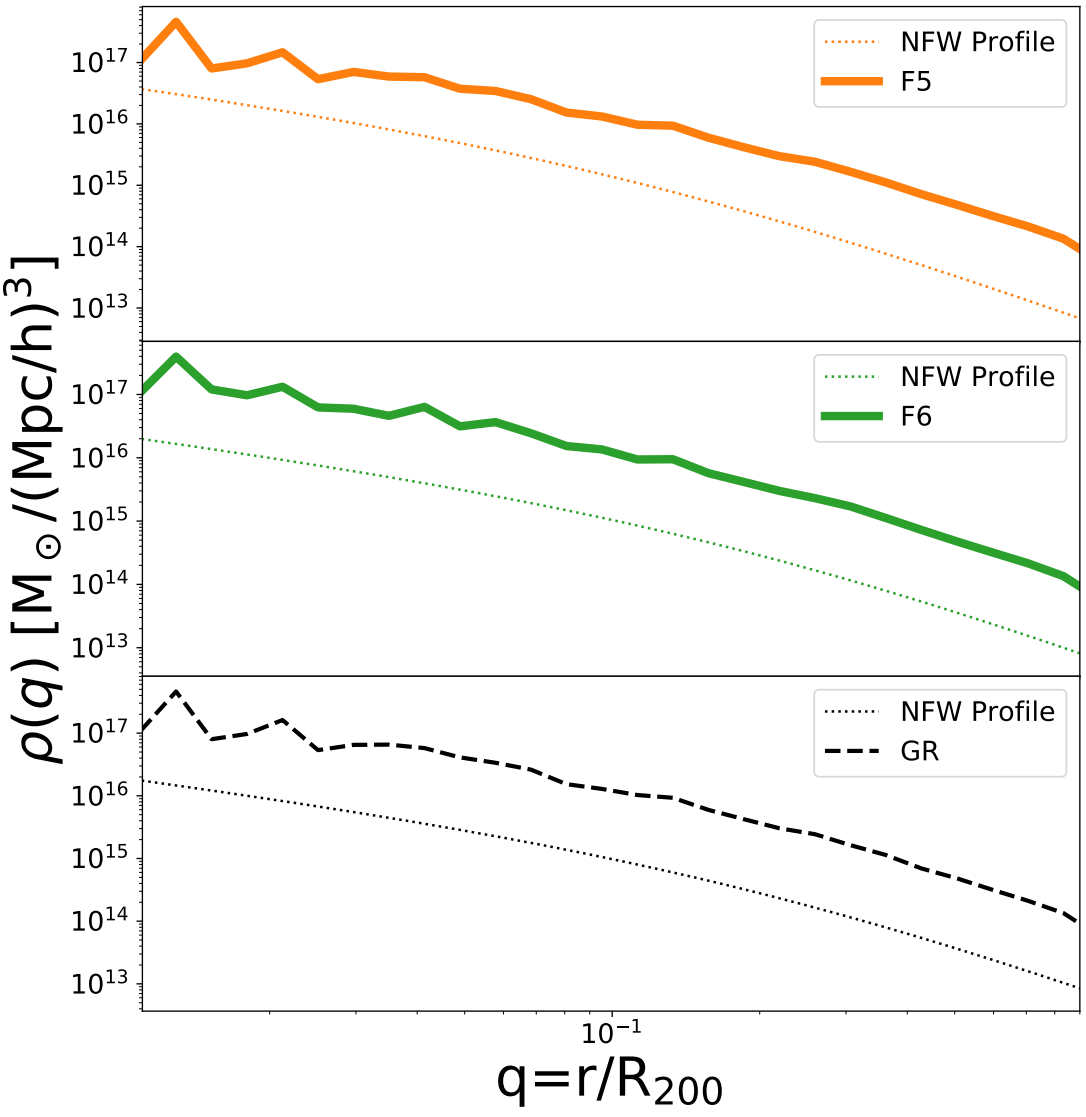

Figure 3.6: The Navarro-Frenk-White profile, plotted on the dotted line as a function of a fraction q of the virial radius R_{200}. The assigned positions of each galaxy in this selected dark matter halo, the most massive halo in the simulation, is also plotted for F5 (orange), F6 (green), and GR (black).

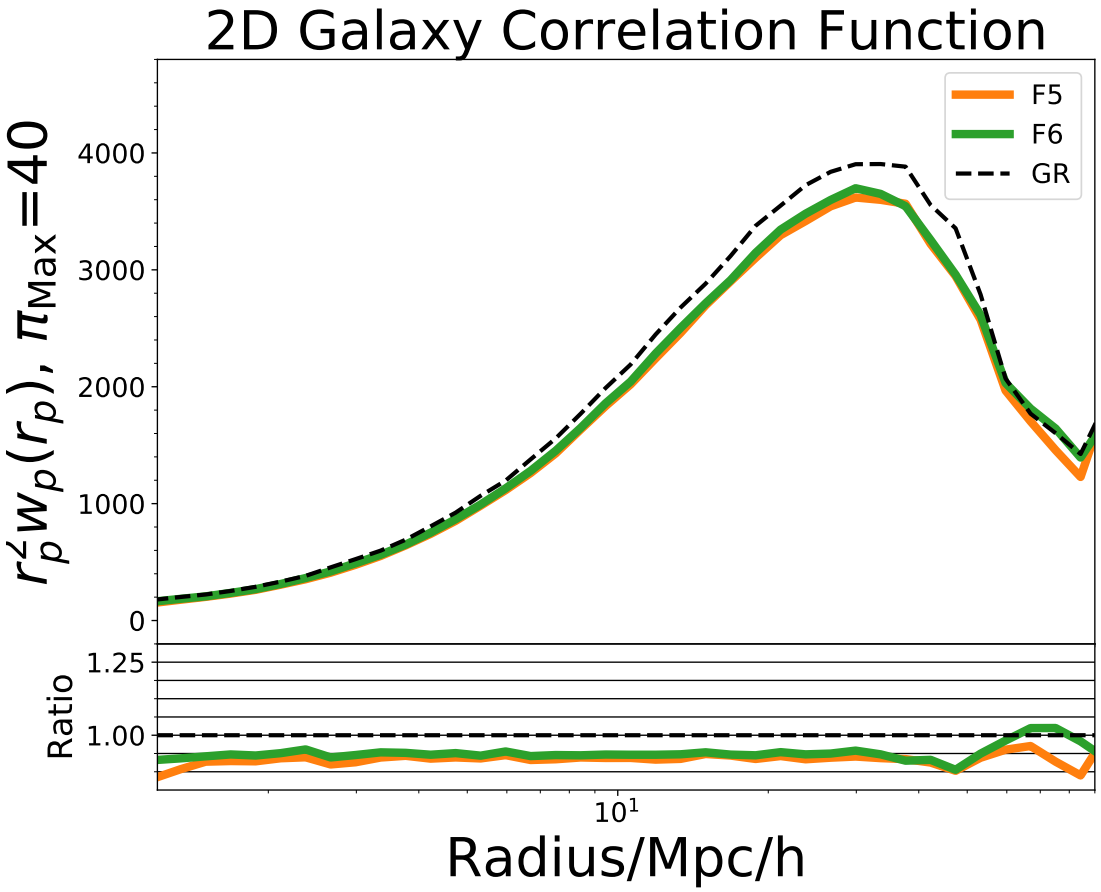

Figure 3.7: The pre-minimization two-point 2D projected correlation function $r^2 w_p(r)$ at $z = 0.34$, calculated at fixed abundance for central and satellite galaxies in F5 (orange), F6 (green), and GR (black). F5: 934,362 galaxies in a (1024 Mpc/h)^3 volume box make a number density of 8.702×10^{-4} galaxies/(Mpc/h)^3. F6: 935,002 galaxies make a number density of 8.708×10^{-4} galaxies/(Mpc/h)^3. GR: 934,918 galaxies make a number density of 8.707×10^{-4} galaxies/(Mpc/h)^3. *Bottom panel*: The ratio of $r^2 w_p(r)$ of F5 and F6 to GR

parameters represented in Tables 3.2 and 3.3 as fixed, for GR only. We then vary all five HOD parameters for F5 and F6, until they are as close as possible to both the number density and the two-point 2D projected correlation function of the mock galaxies populated in GR.

3.4.1 Defining the Error

In order to determine exactly how close the F5 and F6 are to GR, we use a form of the chi-square statistic χ_c, described by Equation 3.10 below:

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{3.10}$$

Here, c refers to the degrees of freedom, O_i to the observed value of the ith point in an experiment, and E_i to the expected value of the ith point. In this case, the observed value O_i represents the mock galaxies in either F5 or F6, and the expected value E_i represents the mock galaxies in GR. Thus, χ_c^2 sums over the difference between the observed and expected values for each point in the data. In this thesis, we adapt this basic model into Equation 3.11 below:

$$\chi_c^2 = \delta w_p(r) + \delta n_{\text{gal}} \tag{3.11}$$

Here, chi-square χ_c^2 can be defined as the sum of the error in clustering $\delta w_p(r)$ and the error in number density δn_{gal}. The former can be computed as the square root of the original chi-square χ_c^2 definition, as described by Equation 3.12.

$$\delta w_p(r) = \sqrt{\sum_{i=1}^{\mathrm{N_{bins}}} \frac{(w_{p,\,\mathrm{f(R)}}(r_i) - w_{p,\,\mathrm{GR}}(r_i))^2}{\mathrm{N_{bins}}}} \tag{3.12}$$

Here, $w_{p,\,\mathrm{f(R)}}(r_i)$ refers to the ith value of the two-point 2D projected correlation function for the $f(R)$ model in question, and is akin to the observed value O_i in Equation 3.10. Similarly, $w_{p,\,\mathrm{GR}}(r_i)$ refers to the ith value of the two-point 2D projected correlation function for GR, and is akin to the expected value E_i. This is identical to the χ_c^2 definition found in Equation 3.10, except the square of the difference is divided by the number of bins $\mathrm{N_{bins}}$, rather than the expected value E_i. Similarly, the error in number density δn_{gal} is defined in Equation 3.13 below:

$$\delta n_{\mathrm{gal}} = 8[(n_{\mathrm{f(R)}} - n_{\mathrm{GR}}) \times 10^4]^2 \tag{3.13}$$

Here, n_{gal} refers to the number density of galaxies in the $f(R)$ model, and n_{GR} to the number density of galaxies in GR, which we are holding as fixed. We multiply the difference in number density $n_{\mathrm{f(R)}} - n_{\mathrm{GR}}$ by 10^4 because all number densities are of order 10^{-4}, and we would like to bring them to unity in order to weight the error in number density δn_{gal} in Equation 3.11 accordingly. Thus, we choose 8 as the weighting factor to make the error in number density δn_{gal} of the same magnitude as the error in clustering $\delta w_p(r)$, determined in Equation 3.12.

3.4.2 Minimizing the Error

At this stage, we have the five HOD parameters M_{min}, $\sigma_{\log M}$, M_0, M_1, and α, which are described in Table 3.1, fixed for GR. Now, we must optimize these parameters for

F5 and F6. To accomplish this, we minimize the chi-square error χ_c^2 between the $f(R)$ gravity models and GR using the definition outlined in Equations 3.11 through 3.13. This process, which we will refer to as "minimization" throughout this thesis, is carried out using the `optimize.minimize` Python package, which is part of the `SciPy` library (Virtanen et al. 2020).

In this minimization process, we begin with an initial guess of these five HOD parameters for F5 and F6. For this initial guess, we use the values that we determined most closely match the target number density 8.8×10^{-4} galaxies/$(\mathrm{Mpc}/h)^3$, as listed in Tables 3.2 and 3.3. As described in Section 3.1, we then populate the dark matter halos in each $f(R)$ gravity model by plugging these initial values into the HOD model. This allows us obtain the number of central and satellite galaxies in each halo, as well as compute the number density of galaxies. Simultaneously, we use the NFW profile to assign real positions to each satellite galaxy within the host halo, as described in Section 3.3.3. Because central galaxies share the same location as their host halo, we can then easily compute the two-point 2D projected correlation function $w_p(r)$ of all the mock galaxies. Finally, we use these values to determine the chi-square χ_c^2 between the $f(R)$ gravity model and GR. The number density and two-point 2D projected correlation function $w_p(r)$ for galaxies in GR is fixed, since the HOD parameters have remained fixed throughout this process.

This is the first loop involved in the minimization process. For numerous iterations thereafter, `scipy.optimize.minimize` repeats this process by testing new HOD parameters until the chi-square error χ_c^2 reaches as low a value as possible. At this point, the minimization process is considered complete, and the last set of guesses for the HOD parameters becomes the optimal values we use for producing the final mock

galaxy catalogues in this thesis. In running `scipy.optimize.minimize`, we also set the tolerance for termination to 10^{-3} and utilized the Nelder-Mead algorithm (Nelder & Mead 1965; Wright 1996). We chose to use adaptive algorithm parameters, which adjust themselves according to the dimensionality of the problem (Gao & Han 2012). This is extremely useful for high-dimensional minimization such as this, as the HOD model involves five free parameters.

3.4.3 Final Results

The final HOD parameters which optimally minimize the chi-square error χ_c^2 using the method outlined in Section 3.4.2 are listed in Table 3.4. Here, we can see the result of the minimization process, and the values of the five HOD parameters ($M_{\rm min}$, $\sigma_{\log M}$, M_0, M_1, and α) ultimately used to populate the mock galaxies throughout dark matter halos in F5 and F6. The number density of galaxies which are listed in this table are obtained using this method.

From this table, it is evident these mock galaxy catalogues have achieved a final number density which is relatively close to the target number density of 8.8×10^{-4}

Simulation	$\rm M_{min}$	σ	α	fM_0	fM_1	$n_{\rm gal}$ (10^{-4} galaxies/(Mpc/h)^3)
F5	12.677	0.58690	1.1053	1.0205	7.4127	8.6553
F6	12.609	0.58856	1.0654	1.0060	7.8379	8.6596

Table 3.4:: The optimal HOD parameters for F5 and F6 which most closely match both the number density and two-point 2D projected galaxy correlation function $w_p(r)$ of GR. Plugging these values into the HOD model results in the above galaxy number densities.

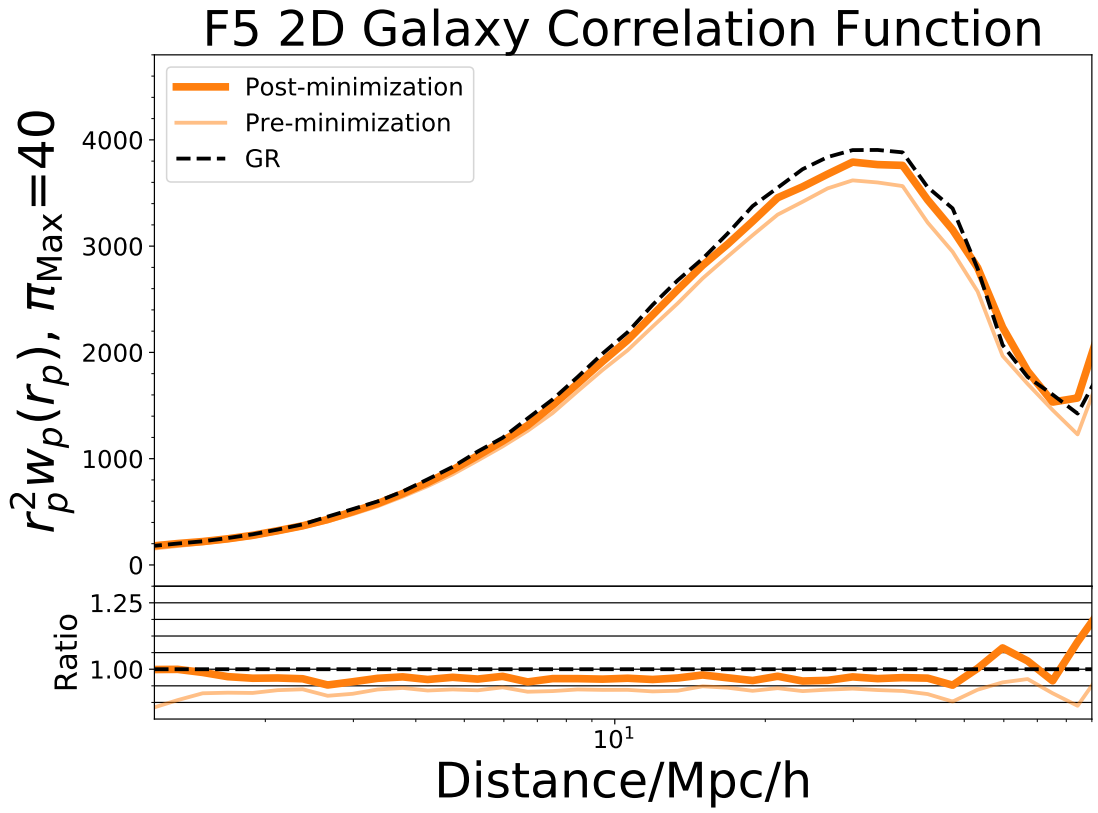

Figure 3.8: *Top Panel*: The two-point 2D projected correlation function $r^2 w_p(r)$ at $z = 0.34$, calculated at fixed abundance for central and satellite galaxies in F5 (orange) and GR (black). F5, post-minimization: 929,821 galaxies in a (1024 Mpc/h)^3 volume box make a number density of 8.6553×10^{-4} galaxies/(Mpc/h)^3. F5, pre-minimization: 934,362 galaxies make a number density of 8.702×10^{-4} galaxies/(Mpc/h)^3. GR: 934,918 galaxies make a number density of 8.707×10^{-4} galaxies/(Mpc/h)^3. *Bottom panel*: The ratio of $r^2 w_p(r)$ of F5 to GR, before and after the minimization process.

galaxies/(Mpc/h)^3, and even closer to the final galaxy number density achieved in GR, which is listed in Table 3.3 as 8.707×10^{-4} galaxies/(Mpc/h)^3. At first glance, it may seem quite alarming that they seem to have strayed further away than the original galaxy number density, which were achieved during the intermediary stage described in Section 3.2. As listed in Table 3.3, these achieved galaxy number densities are within 0.1% of each other. While this is indeed true, it is also very important to note that this final calculation of galaxy number density utilize HOD parameters which were optimized to minimize the error in both number density and the two-point 2D projected galaxy correlation function $w_p(r)$, as Equations 3.10 through 3.13 describe. Thus, any widening gaps in the achieved and target number density are compensated for by simultaneously closing the gap in projected clustering between the $f(R)$ gravity model and GR.

This is illustrated by Table 3.5, which lists the values of the chi-square statistic χ_c in F5 and F6. The column titled "Pre-minimization" refers to the chi-square statistic χ_c had it been calculated prior to the minimization process described in Section 3.4.2, and instead using the HOD parameters listed in Table 3.2 and the optimal $M_{\rm min}$ values shown in Table 3.3. The column titled "Post-minimization" thus refers to the chi-square statistic χ_c calculated after this minimization process, using the HOD parameters which appear in Table 3.4. It is therefore apparent that the minimization process has aided in achieving a uniform galaxy number density and projected clustering among the mock galaxy catalogues in F5, F6, and GR.

This adjustment can be seen very clearly in Figures 3.8 and 3.9, which plot the two-point 2D projected correlation function $r^2 w_p(r)$ at $z = 0.34$, calculated at fixed abundance for central and satellite galaxies in F5 and F6 respectively. The curves titled "Pre-minimization" and "Post-minimization" thus correspond with the

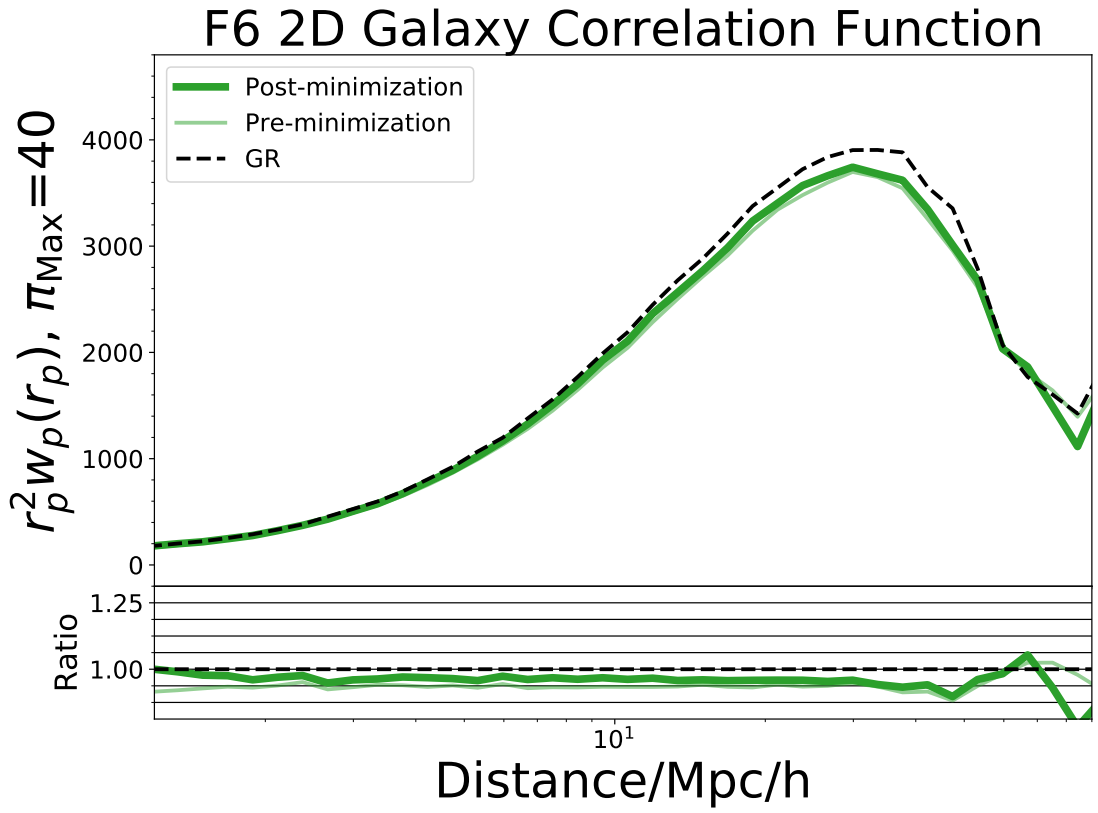

Figure 3.9: *Top Panel*: The two-point 2D projected correlation function $r^2 w_p(r)$ at $z = 0.34$, calculated at fixed abundance for central and satellite galaxies in F6 (green) and GR (black). F6, post-minimization: 929,821 galaxies in a (1024 Mpc/h)^3 volume box make a number density of 8.6596e-4 galaxies/(Mpc/h)^3. F6, pre-minimization: 935,002 galaxies make a number density of 8.708e-4 galaxies/(Mpc/h)^3. GR: 934,918 galaxies make a number density of 8.707e-4 galaxies/(Mpc/h)^3. *Bottom panel*: The ratio of $r^2 w_p(r)$ of F6 to GR, before and after the minimization process.

"Pre-minimization" and "Post-minimization" columns in Table 3.5, which list the final chi-square statistic χ_c. Between "Pre-minimization" and "Post-minimization", the ratio of the projected clustering in F5 to that in GR increased by $\sim 6.6\%$. Similarly, the ratio of the projected clustering in F6 to that in GR increased by $\sim 2.0\%$, thus showing that the minimization process successfully preserved the target number density of galaxies, while matching the projected clustering among the three simulations.

From Figures 3.8 and 3.9, we can see that the mock galaxies in F6 originally dispayed a more similar 2D correlation function to GR than F5 did in the "Pre-minimization" phase. This is to be expected, as F6 has a more efficient screening mechanism than F5 does, and should be clustered in a fashion more similar to GR than F5. However, after the minimization process described in Section 3.4.2 is complete, the mock galaxies in F5 actually counter-intuitively display a more similar 2D correlation function to GR than F6 does. For completeness, we also include Figure 3.7, which plots the two-point 2D projected correlation function $r^2 w_p(r)$ for F5, F6, and GR on the same graph.

This shift in the 2D projected clustering is reflected in the chi-square χ_c values displayed in Table 3.5. Here we can see that F5 originally began with a larger deviation from GR than F6 did, but after the minimization process is complete, F5 actually has a slightly lower chi-square χ_c than F6. Thus, on the basis of comparing the galaxy number density and projected clustering between the $f(R)$ gravity models and GR, the F5 mock galaxy catalogue is actually closer to GR than F6 is, even though it has a more efficient screening mechanism. However, the two still have a chi-square χ_c which is very close in value to each other, putting us in a good position to then apply realism cuts to the galaxy catalogue, a process which is discussed in Chapter 4.

χ^2_{Min}	Pre-minimization	Post-minimization
F5	5.7826	1.3217
F6	3.7710	1.3403

Table 3.5:: The values of the chi-square statistic χ_c in F5 and F6, calculated before and after the final χ^2 minimization process. This procedure calculates the optimal HOD parameters, taking into account the error from the two-point 2D projected galaxy correlation function $w_p(r)$, in addition to the already included error from the galaxy number density.

Chapter 4

Creating a Realistic Mock Galaxy Survey

4.1 Adding Survey Realism

In this chapter, we discuss the process of adding survey selections to the galaxy catalogues described in Chapter 3 in order to create a mock galaxy survey. These survey selections originate from the LOWZ sample parameters of the 2015 BOSS Data Release, which is part of SDSS-III (Reid et al. 2016). The cut includes a redshift z selection of $0.15 < z < 0.43$, a number density cut which mimics number density plot in Figure 3.2, and a projection of the simulation box onto a 2D sky. To make this transformation, we use the publicly accessible MAKE_SURVEY C code (White et al. 2014).

4.1.1 Volume Remapping

In order to utilize MAKE_SURVEY to transform the mock galaxy catalogue into a mock galaxy survey, we simply input the survey parameters from the LOWZ sample for both the north and south Galactic caps of the sky. However, it is important to note that the $f(R)$ and GR simulations are in the shape of a box, while true galaxy surveys cover an irregular volume with geometrical constraints determined by observations. As a result, the MAKE_SURVEY code must "remap" the simulation cube into a cuboid, or non-cubical box shape, using remap vectors. To determine the remap vectors which optimally transform this periodic box, we rely on a remapping technique developed by Carlson & White (2010). This technique embeds the shape of the target galaxy survey inside the simulation cube, while limiting the amount of wasted volume and preserving the volume and internal structure of the mock galaxy catalogue.

A visual of this can be seen in Figure 4.1, which depicts the remapped cuboid, the original simulation cube with box length $L_{\rm box} = 1024$ Mpc/h, and the target survey volume with comoving radial distance D_C, RA and DEC spans $\theta_{\rm RA}$ and $\theta_{\rm DEC}$, and RA and DEC arc lengths $s_{\rm RA}$ and $s_{\rm DEC}$. The latter can be calculated using the classic formula for arc length, where s represents the arc length along a wedge, r represents the radius of the encompassing circle, θ represents the angle that the wedge spans. To calculate the RA arc lengths $s_{\rm RA}$ and $s_{\rm DEC}$, we substitute for the radius r the radial comoving distance D_C. Similarly, the angle spanned θ is equivalent to the RA and DEC spans $\theta_{\rm RA}$ and $\theta_{\rm DEC}$ in units of radians. This relationship is described by Equation 4.1 below:

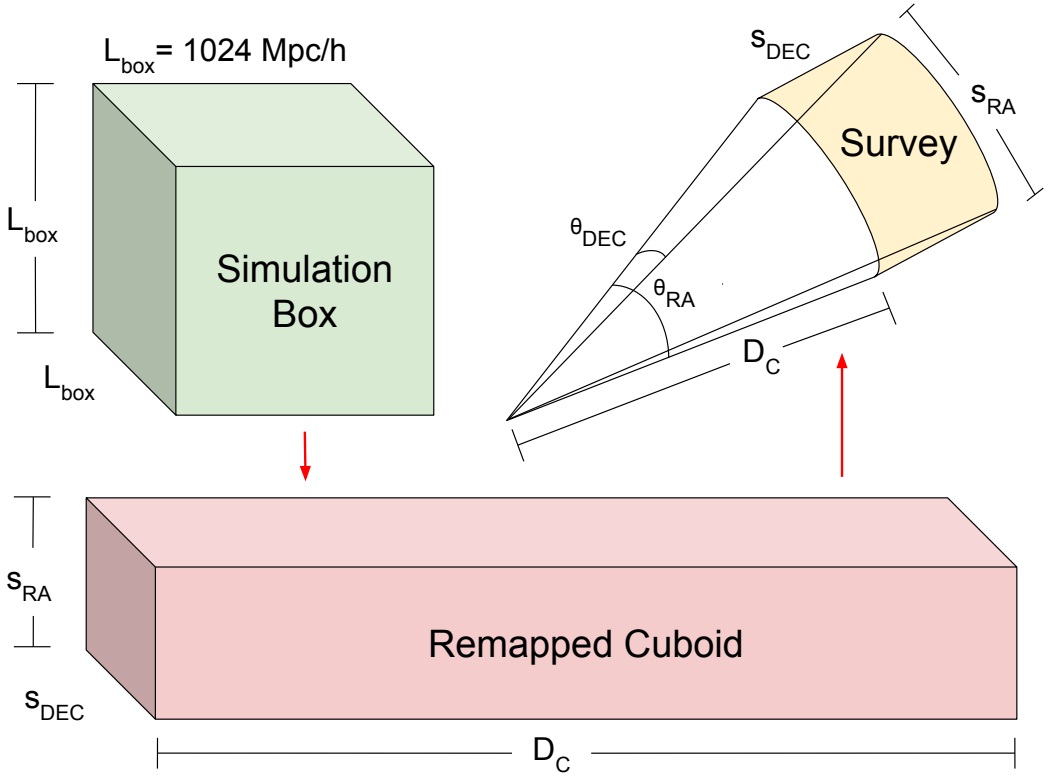

Figure 4.1: The simulation box, with side length $L_{\rm box} = 1024$ Mpc/h, as well as the remapped cuboid it will transform into. The dimensions of this cuboid are determined by the survey volume, which can be described by the comoving radial distance D_C, the RA and DEC spans $\theta_{\rm RA}$ and $\theta_{\rm DEC}$, and the RA and DEC arc lengths $s_{\rm RA}$ and $s_{\rm DEC}$. The dimensions of the remapped cuboid are thus D_C, $\theta_{\rm RA}$, and $\theta_{\rm DEC}$ so that it can encompass the entirety of the survey volume.

$$s = r\theta \tag{4.1a}$$

$$s_{\mathrm{RA}} = D_C \theta_{\mathrm{RA}} \tag{4.1b}$$

$$s_{\mathrm{DEC}} = D_C \theta_{\mathrm{DEC}} \tag{4.1c}$$

As Figure 4.1 shows, because the survey should be able to fit inside the remapped cuboid, it must have, at minimum, physical dimensions of s_{RA}, s_{DEC}, and D_C. From these dimensions, we are able to calculate the corresponding remap vectors, which consist of a 9-element matrix made up of remap vectors u_{11}, u_{12}, u_{13}, u_{21}, u_{22}, u_{23}, u_{31}, u_{32}, and u_{33} as shown below. This remap matrix is then input to MAKE_SURVEY and accordingly remaps the original simulation box. Therefore, in order to determine the proper remap vectors, we must first calculate the physical dimensions of the remapped cuboid, s_{RA}, s_{DEC}, and D_C.

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{bmatrix}$$

4.1.2 Determining the Radial Comoving Distance D_C

Because the arc lengths s_{RA} and s_{DEC} are dependent upon it, as illustrated by Figure 4.1 and Equation 4.1 above, we must first begin by calculating radial comoving distance D_C. To do so, we use the cosmology calculator created by Wright (2006), which uses the definition of radial comoving distance D_C described by Equation 4.2 below (Hogg 1999):

$$D_C = D_H \int_0^z \frac{dz'}{E\left(z'\right)} \tag{4.2a}$$

$$E(z) \equiv \sqrt{\Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda} \tag{4.2b}$$

Here, $D_H = c/H_0$ refers to the Hubble Distance, while $E(z) = H(z)/H_0$ can be described by a function in which Q_M, Q_k, and Q_Λ refer to, respectively, the current fractional density of non-relativistic matter, relativistic matter, and the cosmological constant (Peebles 1993). As a result, in order to determine the radial comoving distance D_C using Wright (2006)'s cosmology calculator, we need only input values for the current Hubble constant H_0, the current fractional density of non-relativistic matter Ω_M, the current fractional density of the cosmological constant Ω_Λ, and the redshift z. It should be noted that this equation can be applied to the $f(R)$ simulation as well as the GR simulation because the exact cosmological background expansion history of ΛCDM is preserved in the $f(R)$ gravity model, as described in Section 2.1.2.

The input parameters are listed in Table 4.1, as well as the resulting radial comoving distance $D_C = 1165.1$ Mpc. We have chosen redshift $z = 0.43$ in order to encompass the entirety of the redshift range $0.15 < z > 0.43$ of the galaxies in the LOWZ catalogue. Furthermore, the values of the current fractional density of non-relativistic matter Ω_M and current fractional density of the cosmological constant Ω_Λ have been chosen according to the values assigned in the ELEPHANT simulations, which are listed in Table 2.1. These are the same values used in the MAKE_SURVEY code. Additionally, the Hubble constant H_0 is assumed to be 100 km/s/Mpc simply for ease of converting the radial comoving distance D_C from units of Mpc to Mpc/h. The relationship between

the Hubble constant H_0 and the dimensionless h-parameter is described by Equation 4.3 below:

$$H_0 = 100\,h\,\mathrm{kms}^{-1}\,\mathrm{Mpc}^{-1} \tag{4.3}$$

Selecting a Hubble constant $H_0 = 100$ km/s/Mpc ensures that $h = 1$. As a result, the comoving radial distance D_C, determined by the cosmology calculator to be 1165.1 Mpc, is equivalent to 1165.1 Mpc/h.

4.1.3 Determining the Remap Vectors

Now that the radial comoving distance is determined, we can obtain the the RA and DEC arc lengths s_{RA} and s_{DEC}. To do so, we introduce L_1, L_2, and L_3, which are dimensionless parameters describing the ratio of the new cuboid lengths to the original simulation box length $L_{\mathrm{box}} = 1024$ Mpc/h. Ideally, the dimensions of this cuboid should be such that they encompass the shape of the LOWZ survey, as is illustrated by Figure 4.1. In order for the survey, and thus all of the mock galaxies, to fit inside the final remapped cuboid, each dimension of the cuboid must, in principle, be at minimum equal to the RA and DEC arc lengths s_{RA} and s_{DEC} spanned by the LOWZ survey, as well as the radial comoving distance D_C calculated in Section 4.1.2. In terms of L_1, L_2, and L_3, this relation can be described according to Equation 4.4 below:

$$L_1 \geq \frac{s_{\mathrm{RA}}}{L_{\mathrm{box}}} \tag{4.4a}$$

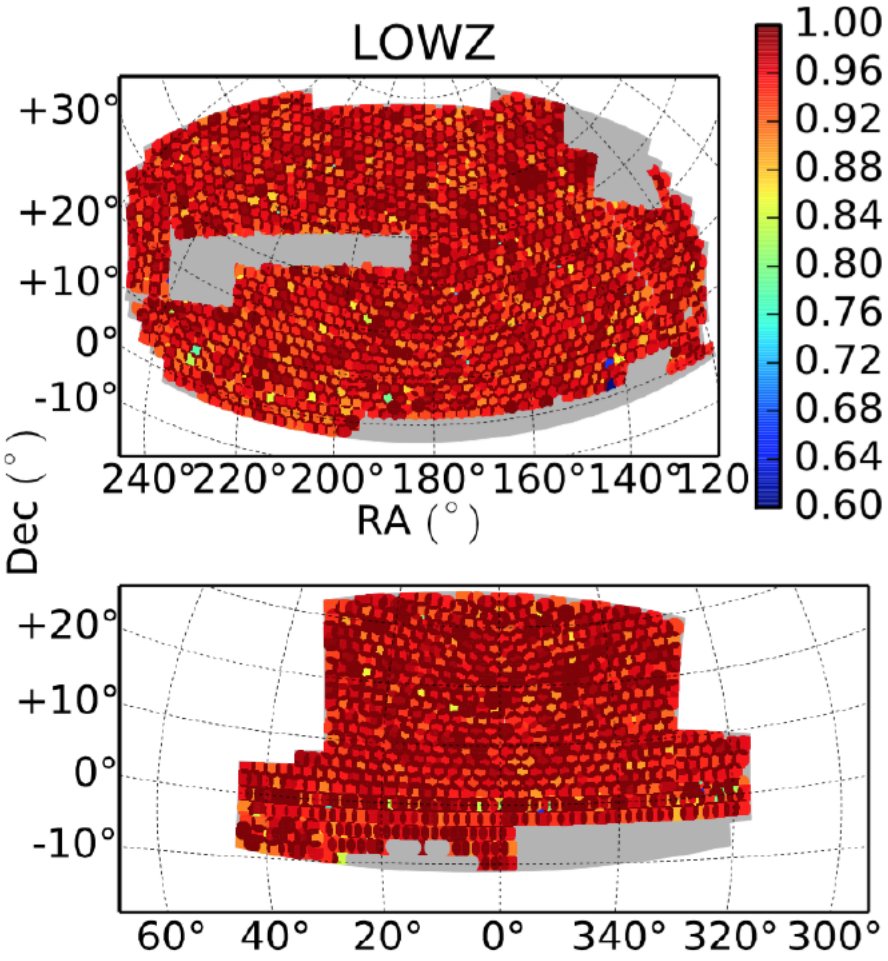

Figure 4.2: Completeness maps in the north and south Galactic caps for the LOWZ catalogue selected from the BOSS sample, part of SDSS-III. Each patch of colour represents an individual plate, and the color determines the completeness, or weight, of that plate, according to the colorbar on the right. The gray shaded regions represent the initial target area on the sky, and any missing patches of color are regions where plates did not reach the minimum required completeness. Figure reproduced from Reid et al. (2016).

$$L_2 \geq \frac{s_{\mathrm{DEC}}}{L_{\mathrm{box}}} \tag{4.4b}$$

$$L_3 \geq \frac{D_C}{L_{\mathrm{box}}} \tag{4.4c}$$

Equation 4.4b should, in theory, describe the minimum physical dimensions of the remapped cuboid which fully encompass the shape of the LOWZ catalogue. However, in reality, the final remapped cuboid is limited by the volume V of the initial $f(R)$ and GR N-body simulations, which is fixed at $V = {L_{\mathrm{box}}}^3$. As a result, once the RA arc length s_{RA} and radial comoving distance D_C is obtained, although the second dimension of the remapped cuboid L_2 should in principle correspond with the DEC arc length s_{DEC}, as defined in Equation 4.4b, it must instead be defined in terms of the remaining "true" arc length, which we denote as S_{DEC}. Since we have already determined the radial comoving distance D_C and the RA arc length s_{RA}, we therefore know that L_2 must be defined by Equation 4.5 below:

$$L_2 \geq \frac{S_{\mathrm{DEC}}}{L_{\mathrm{box}}} \tag{4.5a}$$

$$S_{\mathrm{DEC}} = \frac{V}{D_C \, s_{\mathrm{RA}}} \tag{4.5b}$$

With Equations 4.1 and 4.5, we can finally begin calculating the dimensions L_1, L_2, and L_3 of the remapped cuboid. First, we estimate the RA and DEC spans θ_{RA} and θ_{DEC} from the completeness maps of the north and south Galactic caps in the LOWZ catalogue. These maps are depicted in Figure 4.2, where each patch of colour represents an individual plate with a certain level of completeness. The gray shaded regions

Parameter	Value
H_0	100 km/s/Mpc
Ω_M	0.281
Ω_Λ	0.719
z	0.43
D_C	1165.1 Mpc

Table 4.1:: The input parameters H_0, Ω_M, Ω_Λ, and z for the comoving radial distance D_C determined by Wright (2006)'s cosmology calculator, assuming a flat cosmology where $\Omega_\Lambda = 1 - \Omega_M$.

Galactic cap	$\theta_{\rm RA}$ (°)	$s_{\rm RA}$ (Mpc/h)	$\theta_{\rm DEC}$ (°)	$s_{\rm DEC}$ (Mpc/h)	$S_{\rm DEC}$ (Mpc/h)
North	155	3151.9	75	292.4	1525.1
South	95	1931.8	50	477.1	1016.7

Table 4.2:: The RA and DEC spans $\theta_{\rm RA}$ and $\theta_{\rm DEC}$ which encompass the LOWZ catalogue, as well as the corresponding arc lengths $s_{\rm RA}$ and $s_{\rm DEC}$. The last parameter listed is $S_{\rm DEC}$, which represents the remaining dimension constrained by the vaues of the radial comoving distance D_C, the RA arc length $s_{\rm RA}$, and the volume V of the $f(R)$ and GR N-body simulations.

represent the initial target area of completeness on the sky, and any missing patches of color are regions where plates did not reach the minimum required completeness. The estimated RA and DEC spans $\theta_{\rm RA}$ and $\theta_{\rm DEC}$ are listed in Table 4.2.

We convert these estimated RA and DEC spans $\theta_{\rm RA}$ and $\theta_{\rm DEC}$ from degrees to radians and apply Equation 4.1 to obtain the RA and DEC arc lengths $s_{\rm RA}$ and $s_{\rm DEC}$ for both the north and south Galactic caps in the LOWZ sample. The final results are listed in Table 4.2, as well as the "true" DEC arc length $S_{\rm DEC}$ and the original RA and DEC spans $\theta_{\rm RA}$ and $\theta_{\rm DEC}$. Having determined the values of the radial comoving distance D_C, the RA arc length $s_{\rm RA}$, and the "true" DEC arc length $S_{\rm DEC}$, we simply divide each of these parameters by $L_{\rm box}$ to obtain the final remapped dimensions L_1, L_2, and L_3, which are listed in Table 4.3. From these remapped dimensions, we determine the closest possible side lengths L_1, L_2, and L_3 from a pre-calculated list developed by Carlson & White (2010). These new side lengths are listed in Table 4.4 for both the north and south Galactic caps. The final remap vectors for the north and south Galactic caps of the LOWZ catalogue, each consisting of a 9-element matrix, are written as the following, respectively:

$$\text{North} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\text{South} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

	L_1	L_2	L_3	$s_{\rm RA}$	$S_{\rm DEC}$	D_C
North	3.08	0.286	1.14	3151.9	1525.1	1165.1
South	1.89	0.466	1.14	1931.8	477.1	1165.1
		—			(Mpc/h)	

Table 4.3:: The final remap vectors L_1, L_2, and L_3 for both the north and south Galactic caps of the LOWZ catalogue. The corresponding physical dimensions of the remapped cuboid $s_{\rm RA}$, $S_{\rm DEC}$, and D_C are also listed for comparison.

	L_1	L_2	L_3
North	3.1623	0.2294	1.3784
South	2.2361	0.3333	1.3416

Table 4.4:: The final side lengths L_1, L_2, and L_3 which describe the geometry of the remapped cuboid and most closely match existing side lengths calculated by Carlson & White (2010). These side lengths L_1, L_2, and L_3 are calculated for both the north and south Galactic caps of the LOWZ catalogue.

With these matrices, we are able to input the remap vectors into the MAKE_SURVEY code and transform the mock galaxy catalogue obtained in Chapter 3 into a mock galaxy survey, which in turn mimics the parameters of the LOWZ survey.

4.2 The Final Mock Galaxy Surveys

The MAKE_SURVEY code produces a mock galaxy survey with information on the number density, redshift z, RA, DEC, and completenesss of each galaxy. With this information, we are able to reproduce Figure 3.2, which depicts the number density of the LOWZ catalogue as a function of redshift z. This number density plot is reproduced for each of the F5, F6, and GR mock galaxy surveys in Figure 4.3, where the number density of each galaxy in both the $f(R)$ gravity model and GR mock galaxy surveys is plotted as a function of redshift z. It should be noted that due to the redshift cut which restricts galaxies to a redshift range $0.15 < z < 0.43$, as well as other realism selections, the number density of galaxies has been down-sampled from the target number density of 8.8×10^{-4} galaxies/$(\mathrm{Mpc}/h)^3$ to an average galaxy number density of about 3×10^{-4} galaxies/$(\mathrm{Mpc}/h)^3$.

Figure 4.4 simply displays the sky coverage of the mock survey galaxies on an Aitoff projection. Figure 4.5 is similar, but instead mimics Figure 4.2 to plot the sky completeness maps for each of the F5, F6, and GR mock galaxy surveys in both the north and south Galactic caps. As in the original LOWZ survey, the colorbar distinguishes the different colors of each point, signifying which are of higher completeness, or weight, and therefore more accurate as data. The minimum required completeness for a point to be used is 0.60. These sky completeness plots seem to be successful in mimicking

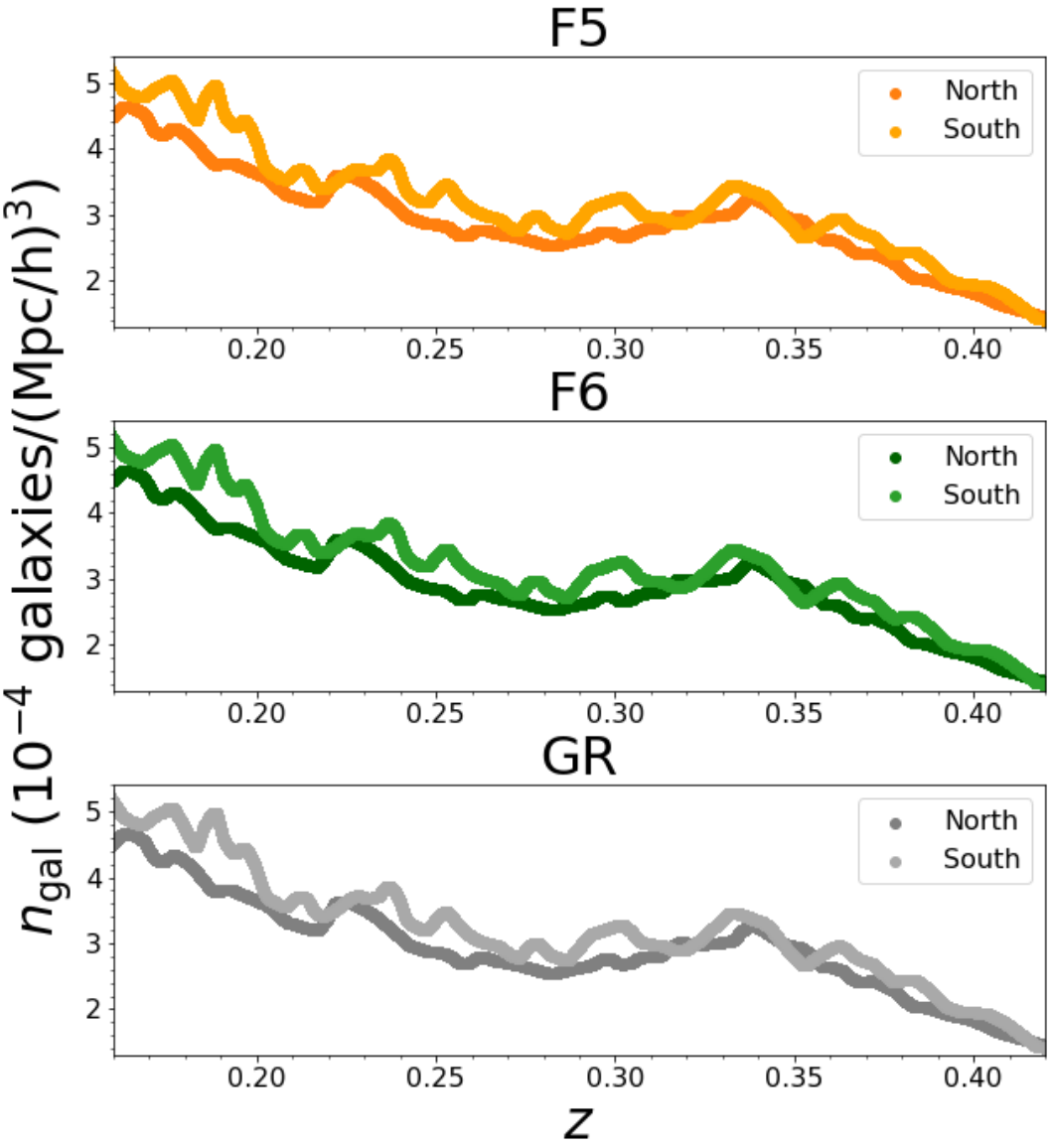

Figure 4.3: The galaxy number density n_{gal} of F5 (orange), F6 (green), and GR (gray), plotted as a function of redshift z. The appropriate galaxy survey realism cuts have already been applied, in accordance with the LOWZ parameters from the 2015 Baryon Oscillation Spectroscopic Survey, which is part of the Sloan Digital Sky Survey III.

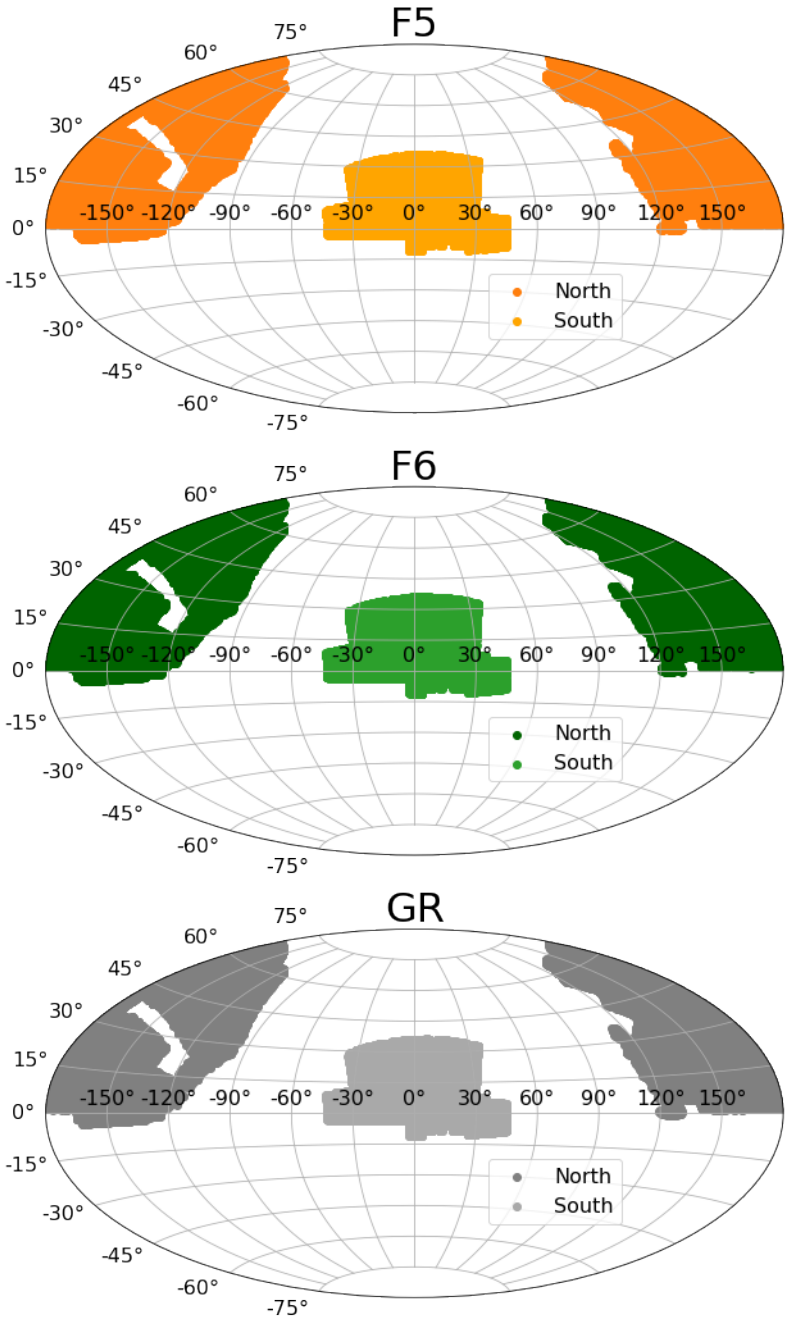

Figure 4.4: The sky coverage for F5 (orange), F6 (green), and GR (gray), plotted on an Aitoff geographic projection. The appropriate galaxy survey realism cuts have already been applied, in accordance with the LOWZ parameters from the 2015 Baryon Oscillation Spectroscopic Survey, which is part of the Sloan Digital Sky Survey III.

the parameters of the LOWZ catalogue, as identical patches in the sky are missing, and the area covered is preserved. One major difference is the unique "fuzziness" seen at the edges of the mock sky completeness plots. In part, this is due to the lowered transparency that the points are being plotted with, in an effort to make visible areas with a lower density of mock galaxies. The root cause, however, lies in the fact that the original $f(R)$ gravity model and GR N-body simulations are simply not not large enough in volume to contain all of the LOWZ survey space, as discussed in Section 4.1.3 and exemplified by the large difference between the ideal DEC arc length and the "true" DEC arc length listed in Table 4.2.

Figures 4.3 through 4.5 are therefore the final visualizations of our mock galaxy surveys for F5, F6, and GR. They are remarkable in how closely each simulation matches the other in terms of number density and sky completeness, as well as how closely they are to the original parameters of the LOWZ survey. This shows that we have successfully implemented a pipeline which transforms an N-body cosmological dark matter simulation into a realistic mock galaxy survey. With this pipeline, we are in a good position to begin investigating underlying differences between the mock galaxy survey. Because we have eliminated trivial differences that are due to the observer, such as a difference in redshift selection, sky completeness, number density, or projected clustering, we can now attribute any differences to the underlying theory itself. This question is further explored in Chapter 5.

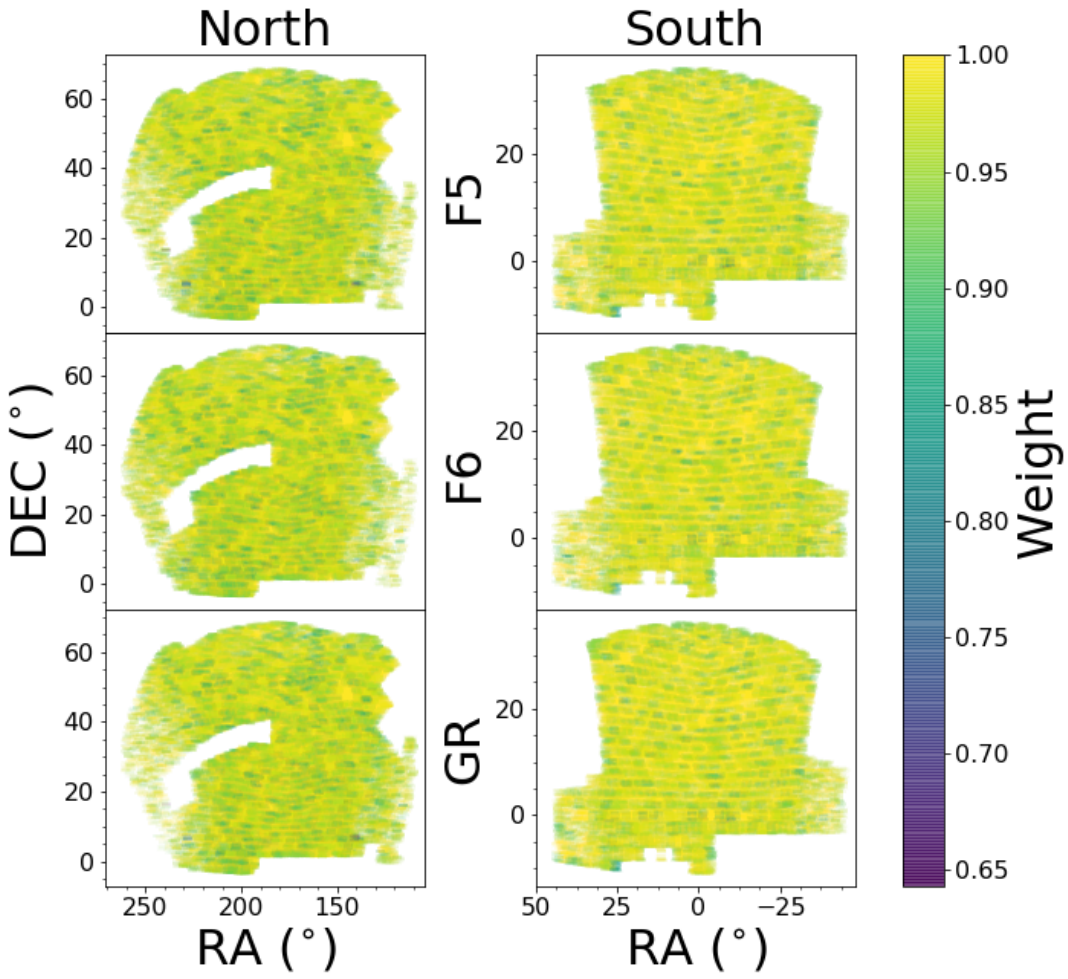

Figure 4.5: Completeness maps for F5, F6, and GR in the north and south Galactic caps, plotted as a function of RA and DEC. The minimum allowed completeness, or weight, of each "plate" is 0.6, as depicted by the colorbar.

Chapter 5

Conclusions

5.1 Discussion

In this thesis, we have compared the nature of gravity in $f(R)$ and GR N-body cosmological simulations, starting at the level of dark matter halos. In Chapter 2, we examined the halo mass function and two-point 3D real-space halo correlation function $r^2\xi(r)$ for F4, F5, F6, and GR at both redshifts $z = 0$ and $z = 0.34$. Here, we were able to see the full effect of the varying efficiency of the screening mechanism for the $f(R)$ gravity model, where it is least inefficient for F4 and most inefficient for F5. Although there is, in a sense, a resolution limit of ~ 12.5 $M_\odot$, this limit simply marks the point at which the number of low-mass halos begin decreasing in the simulations. That does not mean, however, that physical results for halos of this size are untrustworthy.

Having completed this analysis at the halo level, we then began the process of transforming the dark matter simulation into a mock galaxy catalogue by populating

each of the dark matter halos with mock galaxies. We do so by assigning the number of central and satellite using the HOD model and the positions of these galaxies using the NFW profile. Although there are certain limitations to the basic HOD model, such as an under-prediction of galaxies in halos, these errors occur at such small scales, they are at a precision beyond those of upcoming galaxy surveys (Hadzhiyska et al. 2020).

Furthermore, in order to be able to analyze the resulting mock galaxy surveys for differences due to the underlying gravity, we must eliminate any causes of observational variation, such as galaxy number density and projected clustering. To accomplish this, we develop a pipeline which determines the optimal HOD parameters that match the galaxy number density and projected clustering most closely among the $f(R)$ and GR N-body cosmological simulations. Visualizing the mock galaxy survey as a sky completeness map and number density plot, and comparing to the original LOWZ catalogue confirms that this goal has been accomplished. As a result, we have successfully created a pipeline for transforming an N-body dark matter particle cosmological simulation into a realistic mock galaxy survey.

5.2 Future Work

With this new pipeline as a tool, moving forward it is finally possible to begin analyzing the underlying differences between the $f(R)$ gravity model and GR. Because the ELEPHANT simulations used were too small in volume, as evidenced by the "fuzziness" seen in the completeness maps of Figure 4.5, we would need to do so using a simulation with a large enough box length. This ensures that the target survey can fit inside of the remapped cuboid, which shares the same volume as the original simulation. Fortunately,

this is a straightforward process as we simply need follow the HOD pipeline developed in Chapter 3. Once this new simulation is transformed into a mock galaxy survey, we are able to begin testing for observable differences between the $f(R)$ gravity model and GR.

5.2.1 Redshift Space Distortions

Because we have made other factors such as galaxy number density and projected clustering essentially equivalent between the mock galaxy surveys for the $f(R)$ gravity models and GR simulations, we believe that the observable difference which may prove most useful likely lies in redshift space distortions (RSDs). RSDs are an observational phenomenon in which the observed redshift-space position $s(r)$ of a galaxy is different from its true real-space position r.

An example of this can be seen in Figure 5.1, which portrays two common effects of RSDs. On the left are two galaxies as they would be seen in real-space, and on the right are how they appear in redshift-space. The lower galaxy is located in an over-density, and as a result, it becomes "squashed" in redshift-space. In contrast, the upper galaxy is located in an under-dense environment, and in redshift-space it is distorted so that is stretches out along the line-of-sight of the observer.

Such RSDs are often referred to as the "Finger-of-God" (FOG), and can be seen in SDSS galaxies in Figure 5.2 (Tegmark et al. 2004). On the left, they are clearly distorted along some line-of-sight, and on the right they have assumed their original shapes in real-space, after the removal of the FOG effect. Ultimately, these RSDs are caused by the peculiar velocity $v_r(r)$ of a galaxy, or motion unique to an individual galaxy beyond its motion due to the Hubble flow, which is described using Hubble's Law in Equation

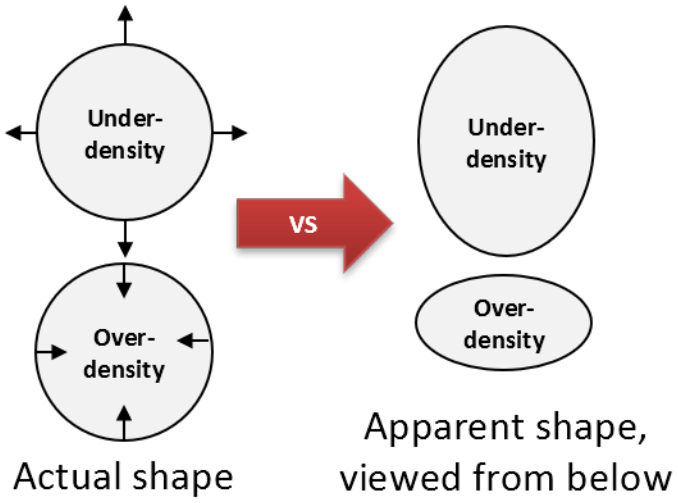

Figure 5.1: A diagram depicting the effects of Redshift Space Distortions (RSD) on observed galaxies. Due to the peculiar velocities V_p of galaxies, galaxies in an underdense region will become elongated along the observer's line of sight, and galaxies in an overdense region will become contracted along the observer's line of sight.

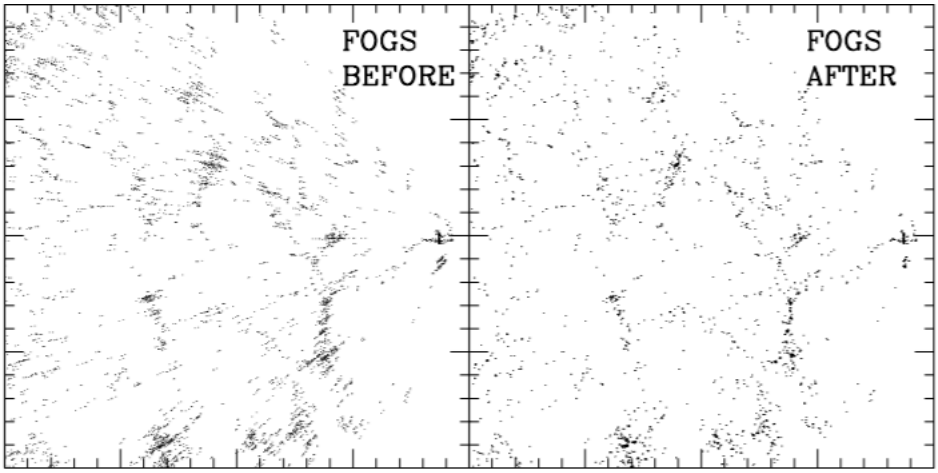

Figure 5.2: The measured large-scale real-space power spectrum $P(k)$ from SDSS. *Left Panel*: An example of a galaxy redshift survey where the initial data displays a strong FOG effect. *Left Panel*: The true real-space positions of the galaxies, after the FOG effect has been removed. Figure reproduced from Tegmark et al. (2004).

5.1 below:

$$V \simeq cz = H_0 r \tag{5.1}$$

Here, V is the recession velocity of a galaxy due to the Hubble Flow. However, Hubble's Law is only adequate for very low redshifts, and can only apply on cosmological scales in a completely homogeneous Universe. Our Universe, in contrast, contains numerous inhomogeneities which derive from galaxies and other structures. This inhomeneity introduces distortions into Hubble's Law in the form of a peculiar velocity $v(r)$, as can be seen in Equation 5.2 below:

$$V = H_0 r + v(r) \tag{5.2}$$

The peculiar velocity $v(r)$ can thus be used to determine the new position of a galaxy in redshift-space, or positions directly derived from the redshift of a galaxy. The exact relationship can be described using Equation 5.3 below, where $v_r(r)$ refers to the radial component of the peculiar velocity (Raccanelli et al. 2013):

$$s(r) = r + v_r(r)\hat{r} \tag{5.3}$$

5.2.2 Velocity Power Spectrum

In order to detect RSDs in a real galaxy survey, we can simply recognize that they are encoded in the observed large-scale clustering signal. This relationship is shown by a simple model for the power spectrum in Equation 5.4 below:

$$P^s_{gg}(k,\mu,z) = k^n T^2(k) G^2(z) \left[b(z,k) + f(z)\mu^2\right]^2 \tag{5.4}$$

Here, we see how observations of the linear power spectrum can be powerful measurements. For instance, a measurement of $P^s_{gg}(k,\mu,z)$ can lead to the full power spectrum, where k is the comoving wavenumber and μ is the cosine angle of the line-of-sight. On the right-hand side of the equation, measuring either $G^2(z)$ or $f(z)\mu^2$ can lead to RSDs. Lastly, $^2(k)$ and $b(z,k)$ are terms whose measurements, respectively, lead to information from the power spectrum shape and information from large-scale bias.

As a result, in our simulations, we can calculate the RSDs of the mock galaxies by examining the velocity power spectrum for the $f(R)$ gravity model and GR. This is possible because we have already calculated the velocity of the mock galaxies, along with their positions, as detailed in Section 3.3.3. Thus, we need only compute the monopole, quadrupole, and hexadecapole power spectra $P_{gg}(k)$, $P_{gv}(k)$, and $P_{vv}(k)$ in redshift-space (Tegmark et al. 2004).

5.3 Final Thoughts

Because of their dependency on the peculiar velocity $v(r)$, RSDs are a powerful tool for distinguishing between different theories of gravity (Kaiser 1987). It is significant that the peculiar velocity $v(r)$ depends upon the structure of galaxies, clusters, and superclusters, which in turn depends on the underlying theory of gravity. This link between inhomogeneities, large-scale structure, under- and over-densities, and the

peculiar velocity $v(r)$ of each individual galaxy means that RSDs should look distinct between mock galaxy surveys for the $f(R)$ gravity model and GR.

Figure 5.3 depicts a prediction of what these differences may look like between F5, F6, and GR in particular. Because F5 has the least efficient screening mechanism, it undergoes the strongest enhancement of gravity. As exemplified by Figure 5.1, galaxies in F5 will likely experience the strongest RSDs between the three simulations, due to its strong enhancement of gravity. It is thus clear that will indeed be differences between the three cosmologies F5, F6, and GR, and it will be particularly interesting to see whether the FOG effect or "squashed" effect is most dominant, as the efficiency of the screening mechanism, and thus the enhancement of gravity, depends on the density of the surrounding region.

As a result, the question we will ultimately need to investigate is whether the differences in RSDs are even visible to observers, as we are constrained by the precision of upcoming galaxy redshift surveys. However, given the automated pipeline we have developed in this thesis, as well as future surveys such as Euclid and DESI, we are poised to make great strides in testing different theories of gravity in the future.

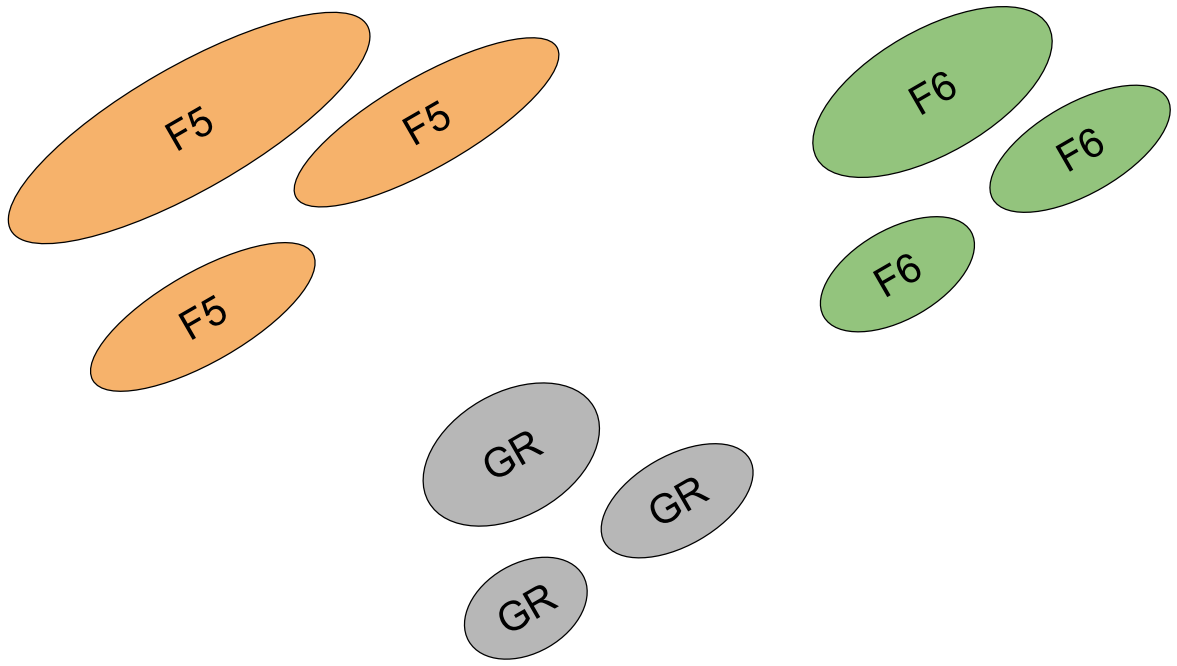

Figure 5.3: A diagram depicting some possible effects of Redshift Space Distortions (RSD) on mock galaxies in the $f(R)$ gravity models and GR. Because it has the least efficient screening mechanism, F5 will likely suffer the most distortion in redshift-space, while GR will likely suffer the least.

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102

Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, Living Reviews in Relativity, 21, 2

Barrow, J. D. 1989, QJRAS, 30, 163

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2012, The Astrophysical Journal, 762, 109

Berti, E., Barausse, E., Cardoso, V., et al. 2015, Classical and Quantum Gravity, 32, 243001

Bertone, G., & Hooper, D. 2018, Reviews of Modern Physics, 90, doi:10.1103/revmodphys.90.045002

Bertone, G., Hooper, D., & Silk, J. 2005, Physics Reports, 405, 279

Blake, C., Kazin, E. A., Beutler, F., et al. 2011a, MNRAS, 418, 1707

Blake, C., Brough, S., Colless, M., et al. 2011b, MNRAS, 415, 2876

—. 2012, MNRAS, 425, 405

Bose, S., Li, B., Barreira, A., et al. 2017, Journal of Cosmology and Astroparticle Physics, 2017, 050–050

Bucher, M., & Ni, W.-T. 2015, International Journal of Modern Physics D, 24, 1530030

Carlson, J., & White, M. 2010, ApJS, 190, 311

Cautun, M., Paillas, E., Cai, Y.-C., et al. 2018, MNRAS, 476, 3195

Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, Phys. Rep., 513, 1

de Swart, J. G., Bertone, G., & van Dongen, J. 2017, Nature Astronomy, 1, 0059

Debono, I., & Smoot, G. 2016, Universe, 2, 23

Delgaty, M. S. R., & Lake, K. 1998, Computer Physics Communications, 115, 395

DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints, arXiv:1611.00036

Dyson, F. W., Eddington, A. S., & Davidson, C. 1920, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 220, 291

Einstein, A. 1916, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.), 1916, 688

Einstein, A. 1916, Annalen der Physik, 354, 769

—. 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, 142

Einstein, A. 1918, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.), 1918, 154

Freedman, W. L., & Madore, B. F. 2010, Annual Review of Astronomy and Astrophysics, 48, 673

Friedman, A. 1922, Zeitschrift für Physik, 10, 377

Friedmann, A. 1922, Zeitschrift fur Physik, 10, 377

—. 1924, Zeitschrift fur Physik, 21, 326

Frieman, J. A., Turner, M. S., & Huterer, D. 2008, Annual Review of Astronomy and Astrophysics, 46, 385–432

Gao, F., & Han, L. 2012, Computational Optimization and Applications, 51, 259

Gravity Collaboration, Abuter, R., Amorim, A., et al. 2018, A&A, 615, L15

Hadzhiyska, B., Bose, S., Eisenstein, D., Hernquist, L., & Spergel, D. N. 2020, MNRAS, arXiv:1911.02610

Hogg, D. W. 1999, arXiv e-prints, astro

Hu, W., & Sawicki, I. 2007, Physical Review D, 76, doi:10.1103/physrevd.76.064004

Ishak, M. 2007, Foundations of Physics, 37, 1470

Ishak, M. 2019, Living Reviews in Relativity, 22, 1

Jackson, N. 2007, Living Reviews in Relativity, 10, 4

Janssen, M., Norton, J., Renn, J., Sauer, T., & Stachel, J. 2007, The Genesis of General Relativity, Vol. 250 (Springer, Dordrecht), doi:10.1007/978-1-4020-4000-9

Joyce, A., Lombriser, L., & Schmidt, F. 2016, Annual Review of Nuclear and Particle Science, 66, 95

Kaiser, N. 1987, MNRAS, 227, 1

Khoury, J., & Weltman, A. 2004, Phys. Rev. D, 69, 044026

Kravtsov, A. V. 2013, ApJ, 764, L31

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, The Astrophysical Journal, 609, 35–49

Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64

Le Verrier, U. J. 1859, Annales de l'Observatoire de Paris, 5, 1

Lemaître, A. G., & Eddington, A. S. 1931, Monthly Notices of the Royal Astronomical Society, 91, 490

Li, B., Zhao, G.-B., Teyssier, R., & Koyama, K. 2012, Journal of Cosmology and Astroparticle Physics, 2012, 051

Lombriser, L. 2014, Annalen der Physik, 264, 259

Lovelock, D. 1971, Journal of Mathematical Physics, 12, 498

Manera, M., Scoccimarro, R., Percival, W. J., et al. 2013, MNRAS, 428, 1036

Manera, M., Samushia, L., Tojeiro, R., et al. 2014, Monthly Notices of the Royal Astronomical Society, 447, 437–445

Michelson, A. A., & Morley, E. W. 1887, Sidereal Messenger, 6, 306

Milgrom, M. 1983, ApJ, 270, 365

Miralda-Escudé, J. 2003, Science, 300, 1904

Montgomery, C., Orchiston, W., & Whittingham, I. 2009, Journal of Astronomical History and Heritage, 12, 90

Mota, D. F., & Shaw, D. J. 2007, Phys. Rev. D, 75, 063501

Navarro, I., & Acoleyen, K. V. 2007, Journal of Cosmology and Astroparticle Physics, 2007, 022–022

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Nelder, J. A., & Mead, R. 1965, The Computer Journal, 7, 308

Nurmi, P., Heinämäki, P., Sepp, T., et al. 2013, MNRAS, 436, 380

Ostriker, J., Peebles, P., & Yahil, A. 1974, The Astrophysical Journal, 193, L1

Peebles, P. J. E. 1968, ApJ, 153, 1

—. 1993, Principles of Physical Cosmology

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, The Astrophysical Journal, 517, 565

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13

Raccanelli, A., Bertacca, D., Pietrobon, D., et al. 2013, MNRAS, 436, 89

Reid, B., Ho, S., Padmanabhan, N., et al. 2016, MNRAS, 455, 1553

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, The Astronomical Journal, 116, 1009–1038

Robertson, H. P. 1935, ApJ, 82, 284

Rubin, V. C., Thonnard, N., & Ford, Jr., W. K. 1980, Astrophys. J., 238, 471

Ryden, B. 2003, Introduction to cosmology

Seager, S., Sasselov, D. D., & Scott, D. 1999, The Astrophysical Journal, 523, L1

Sinha, M., & Garrison, L. H. 2020, MNRAS, 491, 3022

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Stephani, H., Kramer, D., MacCallum, M., Hoenselaers, C., & Herlt, E. 2003, Exact Solutions of Einstein's Field Equations, 2nd edn. (Cambridge University Press, Cambridge, UK)

Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, ApJ, 606, 702

Teyssier, R. 2002, A&A, 385, 337

Trimble, V. 1987, Annual Review of Astronomy and Astrophysics, 25, 425

v. Eötvös, R., Pekár, D., & Fekete, E. 1922, Annalen der Physik, 373, 11

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Walker, A. G. 1937, Proceedings of the London Mathematical Society, s2-42, 90

Weinberg, S. 1989, Rev. Mod. Phys., 61, 1

White, M., Tinker, J. L., & McBride, C. K. 2014, MNRAS, 437, 2594

Will, C. M. 2014, Living Reviews in Relativity, 17, 4

Wright, E. L. 2006, PASP, 118, 1711

Wright, M. 1996, Direct search methods: Once scorned, now respectable, ed. D. Griffiths & G. Watson (Addison-Wesley), 191–208

Zaroubi, S. 2012, Astrophysics and Space Science Library, 45–101

Zeldovich, Y. B., Kurt, V. G., & Syunyaev, R. A. 1968, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 55, 278

Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791

Zwicky, F. 1933, Phys. Rev., 43, 147

Zwicky, F. 1937, ApJ, 86, 217