Galaxy Formation in Standard and Modified Gravity

Piotr Oleśkiewicz

A Thesis presented for the degree of
Doctor of Philosophy

Institute for Computational Cosmology
Department of Physics
University of Durham
England

July 19, 2020

Dedicated to

My teachers

Galaxy Formation in Standard and Modified Gravity

Piotr Oleśkiewicz

Submitted for the degree of Doctor of Philosophy
July 2020

Abstract

This first part of this thesis examines the impact of changes to gravity on the formation of dark matter haloes and galaxies. We use two N-body simulations, one assuming general relativity (GR) and the other the Hu-Sawicki form of $f(R)$ gravity, to investigate the concentration-formation time relation of dark matter haloes. At fixed mass, haloes in modified gravity are more concentrated than those in GR, especially at low masses and redshifts. We find a clear difference between halo concentrations and assembly histories in modified gravity and GR. We incorporate these changes to the properties of dark matter haloes, along with halo merger histories from a N-body simulation of $f(R)$ gravity to build, for the first time, a partial modified gravity version of GALFORM. We concentrate on a model prediction that should, in principle, show a clear trace of modified gravity, the halo occupancy of emission line galaxies. In the second part of the thesis we present the first application of a variance-based sensitivity analysis (SA) to a galaxy formation model. We perform a multi-parameter exploration of GALFORM to compute how sensitive the present-day K-band luminosity function is to varying different model parameters. We first demonstrate the usefulness of the SA approach by varying just two model parameters, one which controls supernova feedback and the other the heating of gas by AGN. The SA analysis matches our physical intuition regarding how these parameters affect the predictions for different parts of the galaxy luminosity function. Our study marks a much needed step away from the traditional "one-at-a-time" parameter variation, often used in this area, and improves the transparency of semi-analytical models.

Declaration

The work in this thesis is based on research carried out at the Institute for Computational Cosmology, the Department of Physics of Durham University, England, under the supervision of Prof. Carlton Baugh and Dr Baojiu Li. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Chapters 3 & 5 have been published in the form of refereed papers:

- Oleśkiewicz, Piotr; Baugh, Carlton M.; Ludlow, Aaron D. **The connection between halo concentrations and assembly histories: a probe of gravity?** MNRAS (2019) **489**: 4658-4668.

- Oleśkiewicz, Piotr; Baugh, Carlton M. **Sensitivity analysis of a galaxy formation model** MNRAS, in press. (received on the 24th November 2019, accepted on the 3rd December 2019)

Acknowledgements

First and foremost, I would like to thank my supervisors, Carlton Baugh and Baojiu Li, for the excellent direction, perpetual support and endless patience they had for me during the last four years. This work could not have been done without them. I am also grateful to my collaborators, especially Aaron Ludlow, John Helly and Christian Arnold, for their assistance, contributions to this work, and readiness to share their knowledge. Many thanks to ICC IT administrators, in particular Lydia Heck, Alan Lotts and Alastair Basden, whose craft and vigilance keeps COSMA in great shape despite our efforts. I extend my thanks to all of ICC staff, all of whom have offered me nothing but support.

It is hard to express how much my Durham friends mean to me. Fabio, Kristian, Stefan, Andrew, Jaime, Stuart, Oliver, Behzad, Jack, Jacob, Jan, Jaime, Stuart, Daniel, Andrew, Ruari, Greg, Sownak, Alex, Paddy, Josh, Chris, Cristian, Louise, Anna, Steve, Will, Ra'ad, thank you for the wonderful time we had together and putting up with my many quirks. I also thank my friends from Oxford and Lublin: Ricky, Rob, Pete, Ramsay, Adam, Pauline, Raph, Cal, Ben, Amelia, Anna, Courtney, Nelli, Rosanna, Klaudia, Mateusz, Przemek, Kuba, Pawel, Kamil, Michal, Katarzyna, Daria, Krzysztof, Ola, Maciek and Lukasz, and all those who stay in touch as we grow older.

I would also like to thank the White Rose Aikikai, especially Shane, Tony, Keith, Gary, Lee, Amy, Nick, Nathan, Mercury, Andrea, Jack and Jarek, for many hours on the mats. You have made my time in Durham very precious, and aikido you have shown me will remain a very cherished memory.

Finally, a thank you to my parents, grandparents and my wife Honee, for always being there for me, not only during a PhD but throughout my life. Knowing that I can always rely on you is the greatest comfort.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Science advances on the fundamental assumption that the Universe is knowable and understandable by a human mind, and cosmology in particular has done more to justify this belief than any other area of research. Cosmology is the study of the Universe as a whole, attempting to explain its origin and evolution. As such, its roots can be traced to the earliest human mythologies, which makes it all the more fascinating that it became a quantitative discipline only a little over one hundred years ago, with Einstein's theory of gravity.

In the same way that the focus of biological sciences was transformed from classifying organisms to explaining their origins by Charles Darwin, the astrophysical sciences were elevated from a catalogue of celestial objects to building mathematical models of entire universes by Albert Einstein's General Relativity (GR) theory of gravity (Einstein et al., 1916). Einstein's equations were solved independently by Friedman (1922) and Lemaître (1927), with the solution being an expanding universe which started from an initial singularity, an infinitely small and dense point at $t = 0$. In the meantime, Slipher (1915) observed that spectral lines of the extragalactc objects are redshifted, and subsequently Hubble (1929) proposed a law that the redshift increases in proportion to the distance of the galaxy away from us, which provided evidence in favour of the expanding universe theory.

The century that followed saw rapid developments in cosmology, in both observations and theory. Zwicky (1933) estimated the velocity dispersion of galaxies in clusters, and argued that the clusters did not contain enough luminous matter to

hold them together. He postulated the existence of the "dark matter" to explain this discrepancy.

His claims were supported on a smaller scale by the work of Rubin & W. Kent (1970) on the rotation curves of the galaxies – independent measurements that both suggested the presence of matter that could not be observed directly, through its effects on the dynamics of matter that could be seen. Rubin showed that the rotation curves of spiral galaxies are flat, which implies that the mass in a galaxy increases in proportion to the distance from the centre, whereas the light in the galaxy is concentrated towards the centre. Moreover, the measurements of the type Ia supernovae in distant galaxies (Riess et al., 1998; Perlmutter et al., 1999) have shown that the rate of the expansion is accelerating. This inspired the postulate of "dark energy", another unknown component of the Universe, which can overcome the gravitational attraction over large scales in space and time.

1.1 ΛCDM

To identify the components of the Universe, we start by formulating GR in terms of Einstein-Hilbert action:

$$S = \int \left[\frac{1}{16\pi G}(R - 2\Lambda) + \mathcal{L}_m \right] \sqrt{-g}\ \mathrm{d}^4x, \tag{1.1}$$

where g is the metric, G is Newton's gravitational constant, R is the Ricci curvature scalar, Λ is a constant and $\mathcal{L}_m$ is a term describing matter field. The actin can be minimised with respect to metric g to obtain the Einstein Field Equations (EFE):

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.2}$$

where $R_{\mu\nu}$ is the Ricci curvature tensor, R is the Ricci curvature scalar, $g_{\mu\nu}$ is the metric tensor, Λ is the cosmological constant, G is Newton's gravitational constant, c is the speed of light in vacuum, and $T_{\mu\nu}$ is the stress – energy tensor.

EFE contain terms for time (1 dimension) and space (3 dimensions), but since all tensors are symmetric, the equation above denotes only 10, not 16 independent relations. Moreover, we can invoke the cosmological principle to simplify it even further. The cosmological principle states that the distribution of energy in the

Universe is homogeneous and isotropic on large scales. Formally, this is expressed by requiring the EFE to remain symmetric under rotations and translations. This requirement was used by Friedman (1922); Lemaître (1927) to derive a solution, known as the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric:

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2\sin^2(\theta)d\phi^2\right], \tag{1.3}$$

where a is expansion factor, size of the Universe normalised to 1 at present time, and k is the space-time curvature.

The FLRW metric can be inserted to EFE. On-diagonal terms of Ricci and stress-energy tensors give the Friedman equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{kc^3}{a^2} + \frac{\Lambda}{3} \tag{1.4}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3}. \tag{1.5}$$

For convenience, the Hubble parameter can be defined as $H^2(t) = (\dot{a}/a)^2$. ρ can also be split into matter (ρ_m) and radiation (ρ_r) parts (since they scale differently with expanding / contracting space-time). Then, ρ_i components can be normalised by the curvature term to $\Omega_i = \rho_i/\rho_c$, where

$$\rho_c = \frac{3H^2(t)}{8\pi G}, \tag{1.6}$$

is the critical density of the Universe at a given time t. Finally, the Ω_i terms can be normalised by their present-day values $\Omega_{i,0}$. Putting all the steps together, we can re-write the first Friedman equation as

$$H^2 = H_0^2\left(\frac{\Omega_{m,0}}{a^3} + \frac{\Omega_{r,0}}{a^4} + \Omega_{\Lambda,0} + \frac{\Omega_{k,0}}{a^2}\right), \tag{1.7}$$

where $\Omega_{m,0}$ $\Omega_{r,0}$ $\Omega_{\Lambda,0}$ $\Omega_{k,0}$ are the energy densities of matter, radiation, cosmological constant and curvature at the present time.

ΛCDM has so far been extremely successful in making cosmological predictions, passing observational tests on large scales and over a wide range of timescales such as the fluctuations in the Cosmic Microwave Background (CMB) radiation, galaxy cluster abundances, and hierarchical structure formation (Springel et al., 2005; Hinshaw et al., 2013; Planck Collaboration et al., 2016a). The small scale predictions, on the

other hand, remain controversial (Weinberg et al., 2015; Bullock & Boylan-Kolchin, 2017) – in particular the observed shape of the innermost regions of dark matter haloes and the number and properties of the satellite galaxies of the Milky Way. It has, however, been pointed out that criticism might be unfounded on statistical and numerical grounds (Genina et al., 2017), or that it might point to stronger impact of baryonic effects (such as a more efficient supernova feedback) or complex physics in the dark sector.

The concordance model emerging from the most recent observations is the Universe which is composed of $\approx 4\%$ visible matter (later referred to as "baryons"), $\approx 26\%$ dark matter and $\approx 70\%$ dark energy. In this model, the dark energy is the energy of the vacuum, called cosmological constant Λ. The most likely candidate for the dark matter are relatively heavy, with a rest mass at the order of GeV, weakly interacting particles called WIMPs (Peebles, 1982).

1.1.1 Cosmological Constant

Adding a cosmological constant or Λ term to the EFE offers one explanation for the accelerating cosmic expansion, but this comes at the price of introducing theoretical problems.

The Fine-Tuning Problem arises when we compare the value of the vacuum energy density, as predicted from quantum field theory, to the observationally inferred value of Λ. The quantum field theory prediction is at least 10^{60} times larger than the observed Λ, and so if these two values are indeed the same physical quantity, reconciling them would require adding some mechanism to the Standard Model of particle physics which could fine-tune the vacuum energy density to the correct level.

The Coincidence Problem comes from an observation that at the present day, the contributions of the Λ and matter energy components to the energy budget of the Universe are comparable. Eq (1.7) predicts that the history of the Universe has distinct epochs, each dominated by different component, and that at the present day, the $(a = 1)$ Λ and matter terms are comparable. For low values of a, after the epoch of radiation domination, the matter density dominates over Λ, while for large values of a it is the other way round. The question the Coincidence Problem asks is

how likely is it that our existence coincides with this special time in the history of the Universe, when both of these terms are within one order of magnitude?

Many modified gravity theories cite the Fine-Tuning Problem as their motivation, and they often provide a physical mechanism justifying the values of Λ (and sometimes evolution of Ω_Λ in relation to Ω_m) as natural by modifying Einstein-Hilbert action, or adding another term to it.

1.1.2 Cold Dark Matter

There is circumstantial evidence in favour of the dark matter hypothesis which goes beyond the rotation curves of the galaxies. Galactic clusters observations using both X-ray images of the hot gas, and gravitational lensing of matter have shown that the missing matter outweighs the luminous matter by a factor of five (Allen et al., 2011). Both observations (Markevitch et al., 2004; Massey et al., 2015) and simulations (Robertson et al., 2016) of galaxy cluster collisions, such as 1E 0657-558 (also known as the Bullet Cluster), have also provided what is largely considered the most direct evidence of dark matter. Further studies of the CMB anisotropies, Supernovae Type Ia distance measurements, baryon acoustic oscillations (BAO), redshift-space distortions and gravitational lensing maps (Refregier, 2003) are all overwhelmingly in favour of the missing matter hypothesis.

Much of the evidence cited in the previous paragraph points at another property of the dark matter – its temperature. To match observations, dark matter needs to be cold, with a low thermal velocity dispersion to preserve density fluctuations on small scales. Possible cold dark matter candidates include axions (Peccei & Quinn, 1977) and WIMPs (Peebles, 1982), but despite extensive searches by observational astronomers and particle physicists, there has not yet been a conclusive detection.

Nonetheless, theoretical research has found more indirect evidence in favour of the ΛCDM cosmology. In ΛCDM , structure grows hierarchically (bottom-up), with small objects collapsing first, and merging to form larger and more massive objects. The predictions of the cold dark matter model are in better agreement with large-scale structure observations than other mainstream alternatives (Springel et al., 2005; De Lucia & Blaizot, 2007).

1.2 Computational Cosmology

Over the past few decades, many areas of science have seen their two modes of research, observations and theory, supplemented by the third one – computation. Nowadays, numerical modelling is one of the main methods in which cosmologists can test their ideas regarding structure formation. Crucially, advancements in computing have allowed solving increasingly complex equations, and on an increasingly vast range of scales.

1.2.1 Initial Conditions

Modern cosmology posits that the large-scale structure formed from nearly-uniform random fluctuations in the primordial hot matter distribution. The linear perturbation power spectrum can be computed using a Boltzmann solver code such as CLASS (Blas et al., 2011). This is then used to generate the initial Gaussian density field for the simulation, using a code such as PANPHASIA (Jenkins, 2013). The particles are then displaced from their initial distribution as prescribed by a second-order Lagrangian Perturbation Theory algorithm (Jenkins, 2010), so that their distribution matches the input matter power spectrum.

1.2.2 Numerical Gravity Solvers

If we assume that these random over- and under-densities in the primordial matter distribution were the seeds of structure formation, we can construct a numerical simulation of the Universe by starting with a taking volume of space, populating it with N gravitationally interacting point particles, and solving the Newtonian equations of motion until we arrive at the present day ($z = 0$) state.

For simplicity, the N-body simulations used in this thesis follow only collisionless particles. This ignores collisional, radiative and pressure-driven exchanges of energy and momentum, and only models the gravitational interactions. We make use of the results of many gravity solvers: ECOSMOG (Li et al., 2012) based on RAMSES (Teyssier, 2002), AREPO (Springel, 2010) and GADGET (Springel, 2005).

1.2.3 Dark Matter Structures

Instead of operating on the entire phase space of billions of particles, it is usually sufficient to describe large-scale structure of the simulation using bulk properties. Therefore, results of the dark matter-only simulations are used to produce halo and subhalo catalogues.

Dark matter haloes are locally overdense regions of dark matter (Davis et al., 1985). As explained in (Mo et al., 2010, Chapters 5 and 7), the gravitational collapse of collisionless dark matter approximately follows the spherical collapse model (Gunn & Gott, 1972; Peebles, 1980), in which a spherical top-hat function of dark matter particles is evolved using Newtonian equations of motion. The linear collapse model predicts that after the overdensities collapse, they have an average equilibrium density of approximately 178 times greater than the critical density. Therefore, for simplicity, dark matter haloes are conventionally identified as isolated regions of space with mean enclosed density 200 larger than the critical one.

The algorithm commonly used to identify such objects is a percolation algorithm called Friends-of-Friends (Davis et al., 1985, FoF), which connects particles located within a linking length b of each other. The linking is done recursively until it forms a closed group, within which distances between particles are less than b, and hence the effective density is above some density threshold. Usually, FoF haloes only take positions into account, which means that gravitationally unbound particles can belong to a FoF halo.

Dark matter subhaloes are locally overdense regions within haloes, which can be used to identify remnants of previous mergers and build halo histories by tracing particles which belong to the same structures at different snapshots.

As subhaloes are typically required to be gravitationally bound their definition is more restrictive, and the subhalo finding algorithms are usually correspondingly more complex. There exist many subhalo finders, and in this thesis we use two algorithms, (Springel et al., 2001, hereafter SUBFIND) and (Han et al., 2012, 2018, Hierarchical Bound-Tracing, hereafter HBT+). Regardless of the algorithm used, the subhaloes we use are always composed of gravitationally bound particles only.

Merger trees are structures used to record the formation histories of haloes. It

records how much they grow in mass and size from snapshot to snapshot, as well as which haloes from the preceding snapshot had merged into the halo in the current one. Since structure identification for halo and subhalo finders typically takes place at individual snapshots in an isolated manner, an additional step is required to link these structures across the time of the simulation.

HBT+ (Han et al., 2012, 2018) is exceptional in this regard, as it constructs the merger trees and identifies substructures at the same time. It starts from the earliest snapshots and builds the tree up as it progresses towards $z = 0$. On the contrary, SUBFIND is run independently, and does not trace the evolution of structure or how dark matter particle membership changes from halo to halo across snapshots. Merger trees are then obtained using the algorithm developed by Jiang et al. (2014, DHALO). DHALO works by tracking particles between snapshots to determine the descendant of each subhalo, and linking these descendants together to produce a merger tree.

Chapter 2

Theory

This chapter outlines the theoretical background of modified gravity, introduces the $f(R)$ modified gravity theory, and describes semi-analytic model GALFORM in detail.

2.1 Modified Gravity

Arguably the first modification to GR was made by Albert Einstein himself when he inserted the cosmological constant, Λ, into Eq (1.1). While his original motivation did not stand the test of time[1], the Λ term turned out to be a crucial component in modern cosmological theory and opened the door to further changes to the original theory.

Most modern modified gravity theories, such as $f(R)$ (Carroll et al., 2004) which this thesis focuses on, cite the fine-tuning problem as the motivation behind their modifications. However, for many theories, once they fit Solar system observational constraints, the original appeal often disappears as the modification is no longer sufficient to explain the original Λ discrepancy. Nevertheless, the upcoming EUCLID (Laureijs et al., 2011) and DESI (Levi et al., 2013) surveys will attempt to extend the constraints on the theory of gravity from the local astrophysical tests to larger, cosmological scales, building on the work of galaxy surveys like 2dFGRS, SDSS, BOSS & eBOSS.

[1]Einstein was famously unhappy about the expanding Universe, believing that it ought to be constant in size

Of particular interest are phenomena which are not predicted by the Λ term in , such as the time-dependence of the dark energy component or the corrections to the gravitational interactions (usually called "the fifth force"). Observing either of these effects would be the "smoking gun" of the modified gravity effects, and will pose a great challenge to the ΛCDM cosmology. Such potential discovery will change our understanding of gravity and cosmology, and it is therefore of great importance that we do not miss it.

There is an additional benefit to modified gravity research, namely the decoupling of the astrophysical and cosmological models from the paradigm of the dominant gravitational theory. For instance, when publishing gravitational lensing data, the light deflection measurement is independent from the gravity theory, but the projected matter distribution is not. Therefore, tests of GR can only be made using data sets that have been processed using self-consistent assumptions. In this way, the very existence of modified gravity research keeps the astrophysical measurements accountable, by preventing assumptions originating in currently dominant model, ΛCDM , from being incorporated into the early stages of experiment design or data analysis.

Unlike the theories like MOND, modified gravity theories such as $f(R)$ attempt to address the dark energy problem, not the dark matter problem. As such, they do not attempt to account for the invisible dark matter, but instead aim to provide a theoretical explanation for the observed accelerated expansion of the Universe (Buchdahl, 1970).

The class of theories we focus on in this work are called $f(R)$ theories. This class of theories introduces an arbitrary function of the Ricci curvature scalar that is added to the Einstein-Hilbert action:

$$S = \int \left[\frac{1}{16\pi G} \left(R + f(R) - 2\Lambda \right) + \mathcal{L}_m \right] \sqrt{-g} \, \mathrm{d}^4 x. \tag{2.1}$$

The $f(R)$ term, after which the theory takes its name, causes an increase in the strength of the gravitational force compared to GR. In order to satisfy astrophysical constraints on the strength of gravity (Lombriser, 2014; Cataneo et al., 2015; Nunes et al., 2017), the theory contains a chameleon screening mechanism (Khoury & Weltman, 2004). This ensures that the gravitational attraction is restored to GR level in dense environments.

From Eq (2.1) we can derive the Poisson equation for modified gravity

$$\frac{1}{a^2}\vec{\nabla}^2\phi = \frac{16\pi G}{3}\left(\rho_m - \bar{\rho}_m\right) + \frac{1}{6}\left(R\left(f_R\right) - \bar{R}\right), \tag{2.2}$$

where $f_R = \mathrm{d}f/\mathrm{d}R$ and bars on top of variables signify background values. The equation remains valid for $|f(R)| \ll |\bar{R}|$ and $|f_R| \ll 1$, both of which hold for the model we are investigating. Evidently, the only difference with respect to the Newton-Poisson equation depends solely on f_R, the derivative of f with respect to R. The magnitude of f_R relative to the classical Newtonian potential, ϕ, splits the equation into two regimes:

1. $|f_R| \ll |\phi|$: gravity is to a good approximation described by GR, with no increased strength; these regions are called "screened".

2. $|f_R| \geq |\phi|$: the Poisson equation is enhanced by a factor of 1/3; in these regions screening is ineffective.

Hence, in $f(R)$ models the strength of gravity is always between 1 and 4/3 times the GR value. The $f(R)$ function does not determine the gravity enhancement, but instead controls the transition between the screened and the unscreened regions. On the other hand, the effectiveness of the screening mechanism is determined by the magnitude of its derivative $|f_R|$.

When $|\bar{R}| \gg M^2$, f_R simplifies to

$$\bar{f}_R \approx -n\frac{c_1}{c_2^2}\left(\frac{M^2}{-R}\right)^{n+1}, \tag{2.3}$$

and $|\bar{f}_R(z)|$ evolves with redshift as given by

$$\frac{|\bar{f}_R(z)|}{|\bar{f}_{R0}|} = \left[\left(1 + 4\frac{\Omega_\Lambda}{\Omega_M}\right) / \left((z+1)^3 + 4\frac{\Omega_\Lambda}{\Omega_M}\right)\right]^{n+1}. \tag{2.4}$$

Hence, the choice of n, c_1 and c_2 is sufficient to uniquely determine the value and evolution of the fifth force: n controls the gradient of the evolution of the screening mechanism, and the ratio c_1/c_2^2 determines its effectiveness.

Even though f is an arbitrary function of the Ricci curvature scalar, there are physical considerations which we can use to postulate its form. Specifically, since $f(R)$ theory sets out to find a mechanism behind the accelerated expansion, f should

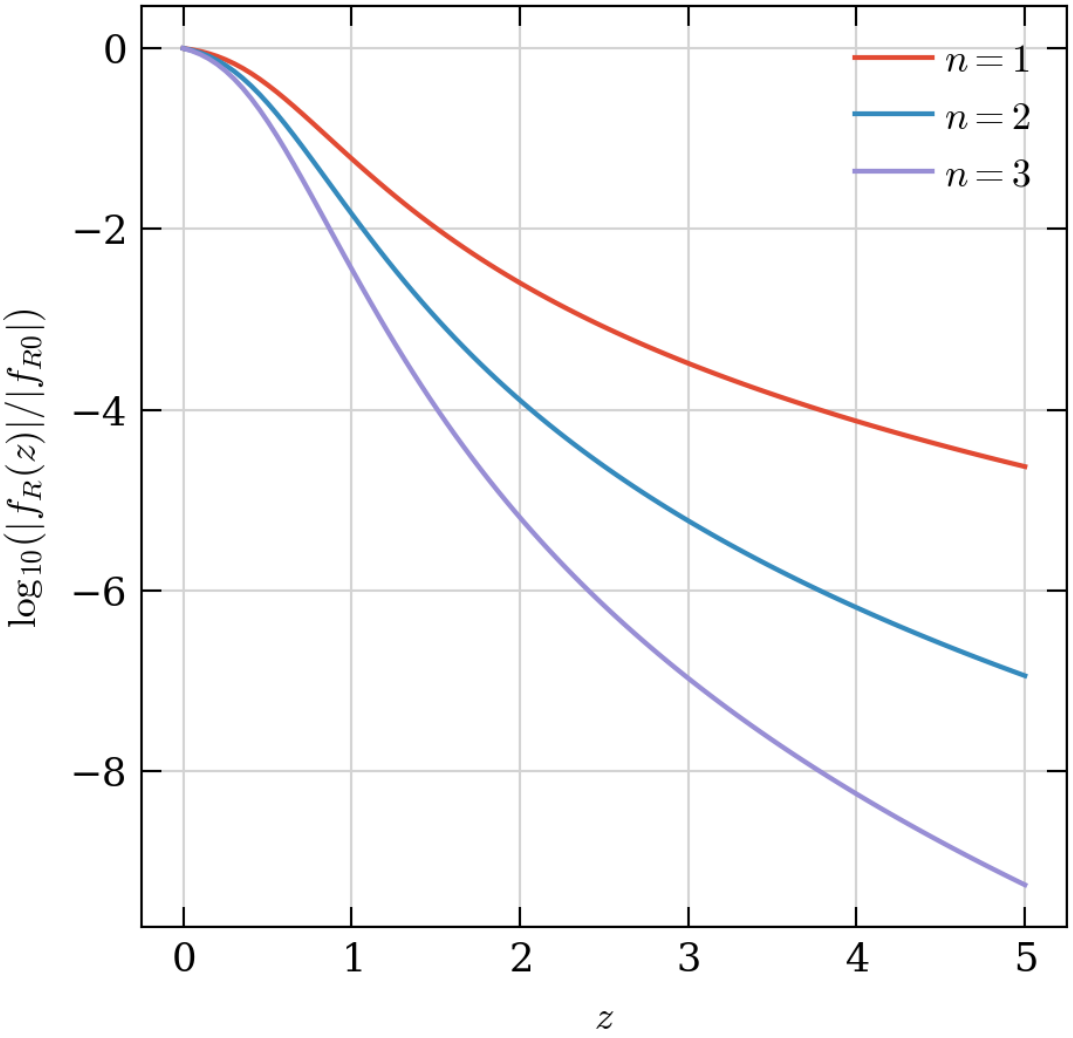

Figure 2.1: Evolution of the relative effectiveness of screening as a function of redshift (Eq (2.4)) for 3 different versions of the Hu & Sawicki (2007) $f(R)$ model with different values of n.

increase with decreasing R (in the low curvature regime). Hu & Sawicki (2007, HS) propose an empirical form:

$$f(R) = -M^2 \frac{c_1 \left(-R/M^2\right)^n}{c_2 \left(-R/M^2\right)^n + 1}, \tag{2.5}$$

where c_1 and c_2 are dimensionless constants which control the screening threshold, and $M = H_0^2/\Omega_\mathrm{m}$ is determined by the cosmology through its dependence on the Hubble constant, H_0, and the matter density parameter, Ω_m.

Since the effectiveness of the screening mechanism can be characterised by $|\bar{f}_R|$ only, we can denote theories with a shorter notation, e.g., F6 will mean a theory with $|\bar{f}_R| = 10^{-6}$. Astrophysical constraints limit the choices of the present day background value of $|\bar{f}_R|$. Supernovae (Upadhye & Steffen, 2013), X-ray (Terukina et al., 2014) and Solar System (Berry & Gair, 2011; Lombriser et al., 2014) observations already rule out models with $|\bar{f}_R| > 10^{-5}$ (F5, F4, etc.). On the contrary, cosmologies with $|\bar{f}_R| \leq 10^{-7}$ show negligible differences to GR in terms of structure formation.

2.2 Model of Galaxy Formation and Evolution

In most models, dark matter behaves different to baryons. Dark matter in simulations follows relatively simple physics, as it is typically assumed to be collisionless, pressureless, and interacts only gravitationally. Baryonic calculations, on the other hand, are considerably more complex. Nonetheless, since galaxies are thought to form in dark matter haloes, it is imperative that we develop a framework for modelling the growth of baryonic structures.

The simplest mechanisms for providing verifiable constraints for observational astronomy avoid modelling underlying baryonic processes in favour of developing a rule-based mapping between dark matter haloes and galaxies. Two examples of such an approach are halo occupation distribution (HOD) modelling (Berlind et al., 2003; Zheng et al., 2005) and subhalo abundance matching (SHAM) (Vale & Ostriker, 2004, 2006) models.

Naturally, such approaches are over-simplified, and can only answer the most basic questions regarding the distribution of galaxies in the Universe. Modelling

galaxy formation in a cosmological setting presents a far greater challenge, both from the physics perspective – there are many distinct physical processes involved, some of which are poorly understood – and from the computational one – the physical processes take place over a wide range of length and time scales, which make them difficult to simulate.

A common approach to this problem are large-scale hydrodynamic solvers which delegate small-scale physics to simplified sub-grid models (Vogelsberger et al., 2014; Schaye et al., 2014). Such hydrodynamic simulation of baryons is significantly more computationally expensive than the simulation of a dark matter only universe of a similar size and resolution, since it involves solving the hydro- and thermo-dynamic equations simultaneously with the Newton-Poisson equation. Subject to the limitations of the computer memory (and hence resolution), the hydrodynamic simulations have recently been very successful in reproducing many of the observed properties of the galaxies. However, they come at a very high computing cost, which restricts their size and resolution, and these calculations still rely on sub-grid models of supermassive black hole (SMBH) growth, supernovae (SNe) feedback, and star formation.

Semi-Analytic Modelling (SAM) builds galaxy catalogues by solving approximate, physically-motivated equations that describe individual baryonic processes, set in the context of the "scaffolding" of the dark matter merger trees (Cole et al., 2000; Baugh, 2006; Benson, 2010; Lacey et al., 2016). Since the results of the SAM do not change the underlying merger tree, it is not possible to use SAMs to predict the effect that the presence of baryons has on dark matter haloes. Moreover, SAM galaxies depend only on the properties of the host haloes, which sacrifices accuracy of the properties of individual galaxies (since some effects of the local environment may not be included). With this in mind, SAM offers an alternative trade-off between the speed of execution and the accuracy of predictions that presented by hydrodynamic simulations. SAMs can be used not only to generate galaxy catalogues for very large cosmological volumes (Merson et al., 2019), but also to explore the parameter space of the galaxy formation model. Since SAM parameters are physically motivated, when comparing results obtained from SAM with the ones from the observational data, it is straightforward to assess whether the model is qualitatively reproducing

the system, or if it needs additional physical process. Similarly, any other analysis carried out using SAM, such as the one in Chapter 5, can be easily interpreted in physical terms.

In this work we focus on GALFORM, a semi-analytic model developed mostly in Durham. It includes prescriptions for the following galaxy formation and evolution processes:

1. merger histories of dark matter haloes
2. heating and cooling of gas inside dark matter haloes
3. quiescent star formation
4. bursts of star formation
5. feedback driven by supernovae
6. heating by an active galactic nucleus (AGN),
7. galaxy mergers,
8. disc instabilities,
9. stellar evolution,
10. dust extinction,

which are discussed in detail in the following subsections.

2.2.1 Merger Histories of Dark Matter Haloes

Fundamentally, GALFORM starts with the assembly histories, density profiles and angular momenta of the dark matter haloes, solving analytic prescriptions which describe how galaxies form inside them, and computing various galaxy properties. Here, we focus on the dark matter halo catalogue extracted from an N-body simulation; many earlier studies focused on merger trees generated using a based on the Extended Press-Schechter theory, using a Monte Carlo method to accurately build

the complete merger histories of dark matter haloes (Cole et al., 2000; Parkinson et al., 2007).

There exist analytical models which can fit dark matter halo density profiles using a small number of free parameters, most notably the the Einasto profile (Einasto, 1965) and the NFW profile (Navarro et al., 1996). The description that follows adopts the NFW prescription, but GALFORM models can be configured and calibrated to use either profile.

The NFW halo density profile is formulated as

$$\frac{\rho\left(r\right)}{\rho_{\mathrm{crit}}}=\frac{\delta_{\mathrm{c}}}{\left(r/r_{-2}\right)\left(1+r/r_{-2}\right)^{2}}, \tag{2.6}$$

where r_{-2} is the characteristic scale radius (the radius at which the slope of logarithmic density profile, $\mathrm{d}\ln\left(\rho\right)/\mathrm{d}\ln\left(r\right)|_{r_{-2}}=-2$), and δ_{c} is a characteristic overdensity. At fixed mass M_{200}, δ_c is given by

$$\delta_{c}=\frac{200}{3}\frac{c^{3}}{\ln\left(1+c\right)-c/\left(1+c\right)}, \tag{2.7}$$

where c is a dimensionless concentration parameter, defined as the ratio of r_{200} over the scale radius of a halo: $c = r_{200}/r_{-2}$. Halo concentration changes when that halo changes mass, for instance during a halo merger or after accreting more mass. In GALFORM, following Press & Schechter (1974); Navarro et al. (1997); Cole et al. (2000), the halo formation time is defined as the snapshot at which a halo has exceeded the mass from the previous formation time by a factor 2 in mass. (The factor of 2 is a parameter choice.) At this snapshot, the density profile, the velocity profile and the angular momentum values are updated. Therefore, even though halo size can change between each snapshot, halo concentration, circular velocity and spin remain constant until the mass threshold is exceeded. Note that the halo spin is drawn from a distribution rather than using the value measured from the simulation. Bett et al. (2007) showed that a large number of particles is needed to get a robust measurement of the halo spin, and most galaxies reside in haloes below the threshold for direct spin measurements.

2.2.2 Heating and Cooling of Gas

In order to form a galaxy, the dark matter halo must gravitationally attract hot gas, which then cools, collapses, and starts the process of star formation. In GALFORM, host haloes are assigned three separate gas reservoirs: hot gas ($M_{\rm hot}$), gas ejected from the halo by SNe and AGN feedback processes ($M_{\rm res}$), and cold gas ($M_{\rm cold}$). Cold gas can be either accreted by a SMBH, used for star formation, or ejected by a feedback process. Ejected gas can be returned to the hot gas reservoir, after some specified timescale. Then, the hot gas can be accreted by the SMBH, or cool down (and move to the cold gas reservoir). The stars in a galaxy belong to the fourth reservoir, M_*.

The mass of these reservoirs evolves according to the following system of equations:

$$\begin{aligned}
\dot{M}_{\rm hot} &= -\dot{M}_{\rm acc} + \alpha_{\rm ret}\frac{M_{\rm res}}{\tau_{\rm dyn,halo}} \\
\dot{M}_{\rm cold} &= \dot{M}_{\rm acc} - (1 - R + \beta)\psi \\
\dot{M}_{*} &= (1 - R)\psi \\
\dot{M}_{\rm res} &= \beta\psi - \alpha_{\rm ret}\frac{M_{\rm res}}{\tau_{\rm dyn,halo}}
\end{aligned} \tag{2.8}$$

where $\dot{M}_{\rm acc}$ is the accretion rate, $\tau_{\rm dyn,halo} = R_{\rm vir}/V_{\rm vir}$ is the halo dynamic timescale, ψ is the star formation rate (SFR), β and $\alpha_{\rm ret}$ are free parameters, and R is the returned mass fraction which is derived from the initial mass function (IMF).

The IMF $\Phi(m)$ gives the distribution of the masses of stars at the time of formation of a stellar population, normalised to unity between lower (m_L) and upper (m_U) mass bounds:

$$\int_{m_L}^{m_U} m\Phi(m) d\ln m = 1. \tag{2.9}$$

In the simplest case, we assume that $\Phi(m)$ follows a power-law, such that

$$\Phi(m) = \frac{dN}{d\ln m} \propto m^{-x}, \tag{2.10}$$

where x is the slope parameter (for instance, Salpeter (1955) take $x = 1.35$).

The returned fraction R is defined as the mass returned to the interstellar medium after a generation of stars is formed. We adopt an instantaneous recycling

approximation, as the mass return is dominated by short-lived Type II supernovae. The returned fraction can be expressed as

$$R = \int_{m_L}^{m_U} (m - m_{\rm rem}(m))\, \Phi(m) d\ln m, \tag{2.11}$$

where $m_{\rm rem}(m)$ is the remnant mass of a star of mass m.

2.2.3 Star formation

Stars in galaxies M_* are assumed to only form directly from the mass in the cold gas reservoir. Each galaxy has two components that contain stars: the disc and spheroid. When the host halo accretes gas, it is assigned to the disc by default, and only moved to the spheroid through galaxy mergers and disc instabilities. Stars form differently in discs than they do in spheroids.

In discs, stars are being formed in a quiescent mode based on the Blitz & Rosolowsky (2006) model, originally implemented in (Lagos et al., 2011). Star formation surface density is defined as

$$\Sigma_{\rm SFR,disc} = \nu_{\rm SF} \times f_{\rm mol} \times \Sigma_{\rm gas}, \tag{2.12}$$

where $\nu_{\rm SF}$ is the inverse of the star formation time-scale, $f_{\rm mol}$ is the ratio of the surface densities of the molecular and total gas masses, $\Sigma_{\rm mol}/\Sigma_{\rm gas}$, and $\Sigma_{\rm gas}$ is the surface density of gas. $\nu_{\rm SF}$ is the inverse of the star formation time-scale and a free parameter of the model, allowed to vary within 1σ, or 0.24 dex, from the best-fit value of $0.43\,{\rm Gyr}^{-1}$, based on measurements of local galaxies from Bigiel et al. (2011).

In spheroids, $f_{\rm mol}$ is assumed to be 1, but the star formation time-scale is calculated in terms of bulge size $r_{\rm bulge}$,

$$\tau_{*\rm burst} = f_{\rm dyn} \frac{r_{\rm bulge}}{V_c(r_{\rm bulge})}, \tag{2.13}$$

but it has a minimum value $\tau_{*\rm burst,min}$ it does not drop below. This minimum value and $f_{\rm dyn}$ are both free parameters.

2.2.4 Supernova feedback

The first feedback mechanism in GALFORM, which regulates star formation and suppresses the luminosity function at the low-mass end of the halo mass function, is driven by the supernovae (White & Rees, 1978). It is modelled as a process which ejects cold gas from a galaxy to a reservoir of mass $m_{\rm res}$, at a rate of

$$\dot{m}_{\rm out} = \beta\psi, \tag{2.14}$$

where ψ is the star formation rate and β is a mass loading factor defined as

$$\beta = \left(\frac{V_{\rm c}}{V_{hot}}\right)^{-\gamma_{\rm SN}}. \tag{2.15}$$

Here $V_{\rm hot}$ and $\gamma_{\rm SN}$ are model parameters and $V_{\rm c}$ is the effective circular velocity of the disc or bulge (for starbursts) at the half mass radius. Note that these equations are applied to quiescent and burst star formation. Different values can be adopted for the $V_{\rm hot}$ parameters for the disc and burst contributions to star formation.

Gas is returned from this reservoir to the hot gas halo at the rate of

$$\dot{m}_{\rm ret} = \alpha_{\rm ret} \times \frac{m_{\rm res}}{\tau_{\rm dyn}}, \tag{2.16}$$

which is controlled by the free parameter $\alpha_{\rm ret}$; $\tau_{\rm dyn} = r_{\rm vir}/V_{\rm vir}$ is the dynamical time of the halo, where $r_{\rm vir}$ is the virial radius of the halo and $V_{\rm vir}$ is the effective circular velocity at this radius.

2.2.5 AGN feedback

Another feedback mechanism, regulating star formation at the high galaxy mass end, is the AGN feedback (Bower et al., 2006). An SMBH in the centre of the galaxy releases energy through accretion of gas, which produces the observational signature of an AGN. This energy can balance the cooling flow in massive haloes, shutting down gas cooling and hence star formation. SMBHs in GALFORM grow in three modes: through accretion of gas during burst star formation, through hot gas accretion, and through black hole mergers.

Supermassive black holes (SMBHs) grow in the centres of galaxies, and inject energy into the gas reservoir of a halo following accretion, which disrupts the cooling process (see (Fanidakis et al., 2011; Griffin et al., 2019) for descriptions of the

treatment of SMBHs in GALFORM). In GALFORM AGN heating occurs when two conditions are met: (i) the hot gas halo is in quasi-hydrostatic equilibrium, defined in terms of the ratio of the cooling time, $\tau_{\rm cool}$, to the free-fall time, $\tau_{\rm ff}$:

$$\frac{\tau_{\rm cool}(r_{\rm cool})}{\tau_{\rm ff}(r_{\rm cool})} > \frac{1}{\alpha_{\rm cool}}, \tag{2.17}$$

where $\alpha_{\rm cool}$ is a parameter, and (ii) the AGN power required to balance the radiative cooling luminosity $L_{\rm cool}$ is below a fraction $f_{\rm Edd}$ of the Eddington luminosity $L_{\rm Edd}$ of the SMBH of mass $M_{\rm BH}$:

$$L_{\rm cool} < f_{\rm Edd} L_{\rm Edd}\left(M_{\rm BH}\right). \tag{2.18}$$

2.2.6 Galaxy Mergers

Galaxy mergers are one of the processes which regulate the movement of cold gas from the disc to bulge, burst star formation in the spheroid, and contribute to the SMBH growth rate.

Similar to how dark matter haloes can be subdivided into subhaloes, the *central galaxies* of host haloes have *satellite galaxies* associated with them. The central galaxy is chosen during each dark matter halo merger event, as recorded in the merger tree. The galaxy associated with the most massive halo becomes the central galaxy, while the other galaxies become the satellites. Subsequently, the central galaxy grows in size through the accretion of cooling gas and mergers with the satellite galaxies.

2.2.7 Disc instabilities

Galaxies can also undergo morphological transformations and starbursts as a result of disc instabilities. Galaxy discs which are dominated by rotational motions are unstable to bar formation when they are sufficiently self-gravitating. In particular, we can relate the rotational and the self-gravitating energies as

$$F_{\rm disc} = \frac{v_c(r_{\rm disc})}{\sqrt{1.68 G M_{\rm disc}/r_{\rm disc}}}, \tag{2.19}$$

where $M_{\rm disc}$ is the total disc mass (ie stars plus cold gas), $r_{\rm disc}$ is the disc half-mass radius, and the factor 1.68 relates this to the exponential scale length of the disc.

The quantity $F_{\rm disc}$ measures the contribution of disc self-gravity to its circular velocity, with larger values corresponding to less self-gravity and greater disc stability. Efstathiou et al. (1982) found a stability threshold $F_{\rm stab} \approx 1.1$ for a family of exponential stellar disc models. Note that a completely self-gravitating stellar disc would have $F_{\rm disc} = 0.61$, which is therefore the minimum value allowed for this parameter. If $F_{\rm disc}$ drops below $F_{\rm stab}$, disc becomes unstable, forms a bar and then a spheroid. The instability then triggers a burst of star formation which consumes all cold gas from the (former) disc component.

2.2.8 Stellar Evolution & Dust Extinction

GALFORM tracks the chemical evolution of the hot gas, cold gas and stellar components, using a similar set of equations to those above for the mass transfer (Eq (2.8)). The yield of metals produced by star formation depends on the form of the IMF. The cold gas metallicity is used to compute the cooling rate and the stellar metallicity is used to build the composite stellar population of each galaxy.

For each galaxy, the model calculates a complete star formation and metallicity history. This is combined with a stellar population synthesis (SPS) model based on stellar evolution models to calculate the luminosity and spectral energy distribution (SED) of the stellar population. For historical reasons, GALFORM tracks the galaxy luminsity in filters that are specified prior to run time, rather than computing the full SED.

Some of the light emitted by stars in galaxies is absorbed by dust, and re-emitted at IR and sub-mm wavelengths. In GALFORM, we assume that the dust is in thermal equilibrium – the energy the dust emits is equal to the energy it absorbs from starlight.

The mass of dust in each galaxy is calculated from the mass of the cold gas as

$$\frac{M_{\rm dust}}{M_{\rm cold}} = 6.7 \times 10^{-3} \left(\frac{Z_{\rm cold}}{0.02} \right), \tag{2.20}$$

where $Z_{\rm cold}$ is the metallicity of the cold gas component, and 0.02 is metallicity normalisation (Silva et al., 1998; Cole et al., 2000).

The dust is assumed to reside in two components, dust clouds and diffuse medium.

The dust medium has the same half-mass radius as the disc if the galaxy forms stars in a quiescent mode, and the same half-mass radius as the stellar bulge if the stars are formed in bursts. Both components emit a modified blackbody radiation:

$$L_{\lambda}^{\mathrm{dust}} \propto M_{\mathrm{dust}} \kappa(\lambda) B_{\lambda}(\lambda, T_{\mathrm{dust}}), \tag{2.21}$$

where M_{dust} is dust mass, $\kappa(\lambda)$ is dust opacity, $B_{\lambda}(\lambda, T_{\mathrm{dust}})$ is Planck's function, and T_{dust} is dust temperature. The emission from the two components is integrated separately and summed to obtain the full spectrum for a given galaxy.

Dust opacity is modelled as a broken power law:

$$\kappa\left(\lambda\right) = \begin{cases} \lambda^{-2}, & \lambda < \lambda_{\mathrm{b}} \\ \lambda^{-\beta_{\mathrm{b}}}, & \lambda > \lambda_{\mathrm{b}}, \end{cases} \tag{2.22}$$

where β_{b} and λ_{b} are free parameters, fit to match the local observations.

This model makes a few simplifying approximations, but it performs well at far-IR and sub-mm wavelengths (Lacey et al., 2016). For a more accurate prediction a more sophisticated model, such as GRASIL, can be used.

Chapter 3

Haloes in ΛCDM and $f(R)$

We compare the concentration–formation time relation of dark matter haloes in two high resolution N-body simulations, one assuming general relativity and the other the Hu-Sawicki form of $f(R)$ gravity with $|\bar{f}_R| = 10^{-6}$. We assign haloes to logarithmically spaced mass bins, and fit median density profiles and extract median formation times in each bin. At fixed mass, haloes in modified gravity are more concentrated than those in GR, especially at low masses and at low redshift, and do not follow the concentration–formation time relation seen in GR. We assess the sensitivity of the relation to how concentration and formation time are defined, as well as to the segregation of the halo population by the amount of gravitational screening. We find a clear difference between halo concentrations and assembly histories displayed in modified gravity and those in GR, and conclude that existing models for the mass–concentration–redshift relation that have gained success in cold and warm dark matter models require revision in $f(R)$ gravity.

3.1 Introduction

N-body simulations have driven astounding progress in improving our understanding of gravitational collapse and its role in the formation of cosmic structure and galaxy evolution. For example, simulations have demonstrated that the mass distribution inside dark matter haloes follows an approximately universal form that can be specified by only two parameters (Navarro et al., 1996, 1997, hereafter NFW

collectively):

$$\frac{\rho\left(r\right)}{\rho_{\rm crit}} = \frac{\delta_{\rm c}}{\left(r/r_{-2}\right)\left(1+r/r_{-2}\right)^2}, \tag{3.1}$$

where r_{-2} is a scale radius (at which the logarithmic slope of the density profile is equal to -2), and $\delta_{\rm c}$ is a characteristic overdensity. It is common to recast these into other forms, such as halo virial[1] mass, M_{200}, and concentration, $c = r_{200}/r_{-2}$ (the ratio of the virial and scale radii). At fixed M_{200}, δ_c is given by

$$\delta_c = \frac{200}{3}\frac{c^3}{\ln\left(1+c\right) - c/\left(1+c\right)}, \tag{3.2}$$

such that higher concentration implies higher characteristic density.

Simulations of structure growth in the cold dark matter model (CDM) have also revealed a well-defined, redshift-dependent correlation between these parameters: at fixed redshift concentrations decrease with increasing mass, and at fixed mass decrease with increasing redshift (see, e.g., Bullock et al., 2001; Gao et al., 2008). These trends betray an simpler relation between the characteristic density of a halo and its formation time, $z_{\rm f}$: haloes that form early have, on average, higher δ_c than late-forming ones, reflecting the higher background density at that time (e.g., Neto et al., 2007; Ludlow et al., 2013). This fact has been used to construct a number of empirical models for the concentration-mass-redshift relation (hereafter $c(M,z)$, for short) that appeal to various definitions of formation time to predict characteristic densities, and hence concentrations (e.g., NFW; Bullock et al., 2001; Wechsler et al., 2002; Zhao et al., 2003; Macciò et al., 2008; Zhao et al., 2009; Ludlow et al., 2014a; Correa et al., 2015; Ludlow et al., 2016).

Various models have met with varied success, plausibly due to diverse definitions of collapse time (see, e.g., Neto et al., 2007; Gao et al., 2008; Ludlow et al., 2016, for details). Several studies define the formation time of a halo as the point at which some fraction F of its final virial mass had first assembled, either into one *main* progenitor or accumulated over many small progenitors. However, as first discussed in Ludlow et al. (2013), better agreement with simulation results can be

[1]We define the virial mass, $M_{200} = (800/3)\,\pi\, r_{200}^3\,\rho_{\rm crit}$, and corresponding virial radius, r_{200}, as that of a sphere (centred on the particle with the minimum potential energy) whose mean density is equal to 200 times the critical density, $200 \times \rho_{\rm crit}$.

obtained by defining $z_{\rm f}$ in terms of the halo's *characteristic* mass, $M_{-2} = M(< r_{-2})$, rather than M_{200}. This is because mass assembly histories and mass profiles of dark matter haloes are self-similar, and can both be modelled by the NFW (or Einasto) profiles; moreover, best-fit parameters of both profiles have a strong linear dependence. The origin of this relation is currently unknown, but we elaborate on this point in Section 3.2.2. This has inspired a number of empirical models that successfully reproduce the $c(M, z)$ relation in both cold (Ludlow et al., 2014a; Correa et al., 2015) and warm dark matter cosmologies (Ludlow et al., 2016)

As a result, there exists an increasingly well-described relation between halo mass and concentration (Duffy et al., 2008; Prada et al., 2012; Angel et al., 2016; Klypin et al., 2016; Diemer & Kravtsov, 2015; Diemer & Joyce, 2019) – the two parameters that are needed to specify the density profile of a relaxed dark matter halo–and how they evolve with time. Further, both analytic and empirical models have been shown to describe reasonably well the $c(M, z)$ relation for a variety of cosmological parameters and power spectra. Our objective here is to investigate whether the relation between concentration and formation time–upon which many of these models are based–is sensitive to the gravitational force law, as stark differences could be used to probe departures from general relativity.

Proposals for modifications to general relativity (GR) were originally motivated by trying to solve one of the biggest remaining problems with the concordance ΛCDM: the origin of the accelerated cosmic expansion. ΛCDM achieves this by invoking a cosmological constant, Λ, but the required value is difficult to justify from a theoretical viewpoint (Carroll et al., 2004). Many alternatives have been proposed to the standard ΛCDM model: the accelerated expansion could be driven by as-of-yet unknown physics in the dark sector (Zuntz et al., 2010) or by a modification to GR itself (Koyama, 2016). Among the alternatives to GR, one of the most widely studied is $f(R)$ gravity – an umbrella term referring to modified gravity models which change the Ricci scalar in Einstein-Hilbert action (Buchdahl, 1970; Clifton et al., 2012; Joyce et al., 2015). Current versions of the theory are fine-tuned to match the expansion history in ΛCDM, which removes some of the model's original appeal. Nevertheless, $f(R)$ gravity remains a workable alternative to GR with inter-

esting phenomenology. While the parameter space of $f(R)$ models is already tightly constrained by observations (Lombriser, 2014), there still exists a range of models which may display measurable differences from GR (see, for example, He et al., 2018; Hernández-Aguayo et al., 2018).

Our study uses the merger histories of dark matter haloes traced back to progenitors that are two orders of magnitude less massive than the final halo mass. Hence, high resolution simulations are necessary (see Table 3.1). We therefore use the LIMINALITY simulations of Shi et al. (2015), a suite of very high resolution dark-matter-only runs including examples of the Hu & Sawicki (2007, HS) parametrisation of $f(R)$ gravity. Two simulations are compared: one of GR and another $f(R)$ modified gravity model that is compatible with current observational constraints.

This Chapter is structured as follows. The theoretical background is given in Section 3.2: the details of the N-body simulation are presented in Section 3.2.1, a description of the $c(M, z)$ model of Ludlow et al. (2016) in Section 3.2.2; the methods for building halo catalogues and merger trees are described in Section 3.2.3 and Section 3.2.4, respectively. Our results are presented in Section 3.3. Halo selection is outlined in Section 3.3.1, and the processing (fitting density profiles and estimating formation times) is covered in Sections 3.3.2 and 3.3.4. The concentration – formation time relation obtained from the processed simulation data is presented in Section 3.3.5. We explore the sensitivity of the model predictions to the parameter choices that specify the model in Section 3.3.6, and to the segregation of the halo population by the effectiveness of the screening of the gravity fifth force in Section 3.3.7. Finally, in Section 3.4, we present our conclusions. Results obtained by fitting Einasto (1965) (rather than NFW) profiles to determine halo structural parameters are discussed in Section 3.3.3.

3.2 Theory

3.2.1 N-body Simulation

As the equations describing the modifications to standard gravity are non-linear, modified gravity simulations are more demanding of computational resources than

Table 3.1: Relevant parameters of the LIMINALITY N-body simulations from Shi et al. (2015).

$\Omega_{\rm m}$	(matter density)	0.281
Ω_Λ	(dark energy density)	0.719
Ω_b	(baryon density)	0.046
σ_8	(power spectrum amplitude)	0.820
n_s	(spectral index)	0.971
h	$(H_0/[100\,{\rm km\,s^{-1}\,Mpc^{-1}}])$	0.697
L	(box side)	$64h^{-1}$Mpc
M_p	(particle mass)	$1.523 \times 10^8 h^{-1}{\rm M}_\odot$
N_p	(particle number)	512^3
$z_{\rm final}$	(final redshift)	0.0
z_0	(initial redshift)	49.0
$N_{\rm out}$	(number of outputs)	122

their standard gravity counterparts of the same size and resolution. However, significant progress has been made recently in numerical techniques designed specifically for this class of theories (Li et al., 2012; Bose et al., 2015). We focus our analysis on the LIMINALITY simulation (Shi et al., 2015), a high-resolution, N-body simulation of HS F6 modified gravity. For comparison, a GR simulation with otherwise identical cosmology is also studied. The cosmological parameters of both runs (Table 3.1) have been tuned to match the cold dark matter, WMAP-9 cosmology (Hinshaw et al., 2013). This is in contrast to the simulations in Ludlow et al. (2016), who used WMAP-1, WMAP-7 and Planck cosmologies, for both warm and cold dark matter versions. However, as demonstrated in **?**Ludlow et al. (2016), cosmology differences are not significant, and should not have a significant effect; moreover, simulation results we used are state-of-the-art in terms of resolution, which is critical in observing modified gravity effects in individual haloes.

3.2.2 Mass-Concentration-Redshift relation

The $c(M, z)$ model tested here, first described in Ludlow et al. (2016), uses the extended Press-Schechter (EPS) formalism to approximate the gravitational collapse of collisionless DM haloes (Bond et al., 1991; Mo et al., 2010). In EPS, the *collapsed mass history* (hereafter CMH), $M(z)$, of a dark matter halo (i.e. the sum of progenitor masses at redshift z exceeding $f \times M_{200}(z_0)$) identified at redshift z_0 is given by

$$\frac{M(z)}{M_0} = \mathrm{erfc}\left(\frac{\delta_{\rm sc}(z) - \delta_{\rm sc}(z_0)}{\sqrt{2(\sigma^2(f \times M_0) - \sigma^2(M_0))}}\right). \tag{3.3}$$

Here $M_0 = M_{200}(z_0)$ is mass at the identification redshift, $\sigma^2(m)$ is the variance of the density field smoothed with a standard choice of a window function, a spherical top-hat window containing mass m, and $\delta_{\rm sc}(z) \approx 1.686/D(z)$ is the redshift-dependent spherical collapse threshold, with $D(z)$ the linear growth factor.

One difference between the EPS theory and the Ludlow et al. (2016) scheme is the definition of halo formation time: in EPS, a common definition of a formation redshift, z_f, is the one at which the sum of progenitor masses more massive than $f \times M_{200}$ first exceeds a fraction $F \times M_{200}$, where typically $f = 0.01, F = 0.5$ (e.g. Lacey & Cole, 1993; Navarro et al., 1996). In Ludlow et al. (2016), F is not a parameter, but varies between the haloes and can be calculated from their concentration:

$$F = \frac{M_{-2}}{M_{200}} = \frac{\ln(2) - 1/2}{\ln(1+c) - c/(1+c)}, \tag{3.4}$$

where the right-most equation is strictly valid for an NFW profile. For each halo, z_f therefore corresponds to the redshift at which a fraction M_{-2}/M_{200} of the halo's final mass had first assembled into progenitors more massive than $f \times M_{200}$ (where $f = 0.02$). Ludlow et al. (2016) referred to this redshift as z_{-2}, to annotate its explicit dependence on the characteristic mass, M_{-2}.

The CMH is scale invariant in both CDM and warm dark matter (WDM) models, and can be used to estimate z_{-2} and the corresponding critical density, $\rho_{\rm crit}(z_{-2})$. The $c(M, z)$ model advocated by Ludlow et al. (2016) exploits the strong, linear correlation between $\rho_{\rm crit}(z_{-2})$ and $\langle\rho_{-2}\rangle$, the mean density within r_{-2}. Empirically, they found $\langle\rho_{-2}\rangle = A \times \rho_{\rm crit}(z_{-2})$, with $A \approx 400$. Once the CMH is known, this

expression can be used to compute $\langle\rho_{-2}\rangle$, and hence infer the halo mass profile.

The model accurately reproduces the concentrations of dark matter haloes in both CDM and WDM cosmologies. This may appear surprising at first as dark matter haloes in WDM simulations have been found to display different concentrations and formation times than in CDM (Macciò et al., 2013; Bose et al., 2016). However, these changes act to preserve the $\langle\rho_{-2}\rangle - \rho_{\rm crit}(z_{-2})$ relation seen in CDM.

It has been shown that haloes in $f(R)$ cosmologies follow NFW density profiles (Lombriser, 2014) like their GR counterparts, but with systematically higher concentrations. Their assembly histories also differ, but only slightly (Shi et al., 2015). Hence, it might be expected that the relation discovered by Ludlow et al. (2016) for CDM and WDM haloes in standard gravity might hold for $f(R)$ haloes only under certain conditions: (i) for small values of $|f_{R0}|$, and (ii) for all haloes *except* low-mass objects at low redshifts, due to screening. It is therefore plausible that the above concentration – formation time relation will not be applicable to the full population of haloes in $f(R)$ gravity, and this is the hypothesis that we test here. This breakdown could potentially be circumvented by either re-parametrising the model or segregating haloes to reflect the influence of the fifth force, which we explore later.

3.2.3 Halo identification

The gravitational collapse of collisionless CDM can be approximated by the spherical collapse model (Gunn & Gott (1972); Peebles (1980); but see Ludlow et al. (2014b)). In this model, overdensities collapse to form dark matter haloes, which are defined as isolated regions with an average matter density larger than a threshold $\Delta_{\rm vir} \approx 178\,(\approx 200)$ times the critical density (Mo et al., 2010, Ch. 5).

Because we are primarily concerned with the GR / $f(R)$ comparison, we have elected to use r_{200} to define halo virial radii and M_{200} for the corresponding masses. This convention follows that of Ludlow et al. (2016) and is based on the fact that, while r_{200} remains well-defined and is independent of the gravity model, the virial parameters vary systematically with the strength of gravity (Schmidt et al., 2009). The virial mass and radius therefore define a sphere (centred on the particle with the minimum potential energy) that encloses a mean density equal to 200 times the

critical density, $\rho_{\rm crit}(z)$, and are thus labelled with the subscript 200.

Subhaloes are locally overdense regions within haloes, and are the surviving remnants of past mergers. Haloes are initially identified using a friends-of-friends (FoF) algorithm (Davis et al., 1985). The halo catalogue is then processed using an upgraded version of HBT (Han et al., 2012, Hierarchical Bound-Tracing algorithm), HBT+ (Han et al., 2018), which identifies subhaloes and builds their merger trees.

HBT+ is a publicly available[2] merger tree code, which identifies subhaloes and follows them between simulation outputs, from the earliest snapshot at which they can be identified until the final one, building a merger tree from the catalogue on-the-fly. A list of gravitationally bound particles is created for each halo; these are used to identify a descendant (a halo at a lower redshift, sharing subhaloes), and are passed to the successive snapshot. Each halo can have one or more progenitors (haloes at a higher redshift, sharing subhaloes). If a halo has multiple progenitors, the most massive one is selected, and it becomes the "main" (i.e. most massive) subhalo. Other progenitors are mapped to the subhaloes which belong to the host halo. The host halo of a subhalo is the FoF halo containing its most bound particle.

3.2.4 Merger trees

The merger tree of a halo, visualised in Fig. 3.1, can be obtained from the HBT+ output by following the progenitors of a given halo, recording their host haloes, and repeating this process recursively until the earliest progenitors are reached in each branch. However, the trees produced by this procedure have two common defects[3]:

1. Re-mergers, such as the right-most halo in the second row in Fig. 3.1, happen when one of the subhaloes temporarily becomes gravitationally unbound and is identified as a separate halo for one or more snapshots; in a later snapshot it merges back into the original host halo, creating a "loop". The halo in the "loop" is retained as a progenitor halo and so re-mergers do not alter

[2] `https://github.com/Kambrian/HBTplus`

[3] Technically, these are not trees as they contain loops, and some nodes might have more than one parent node.

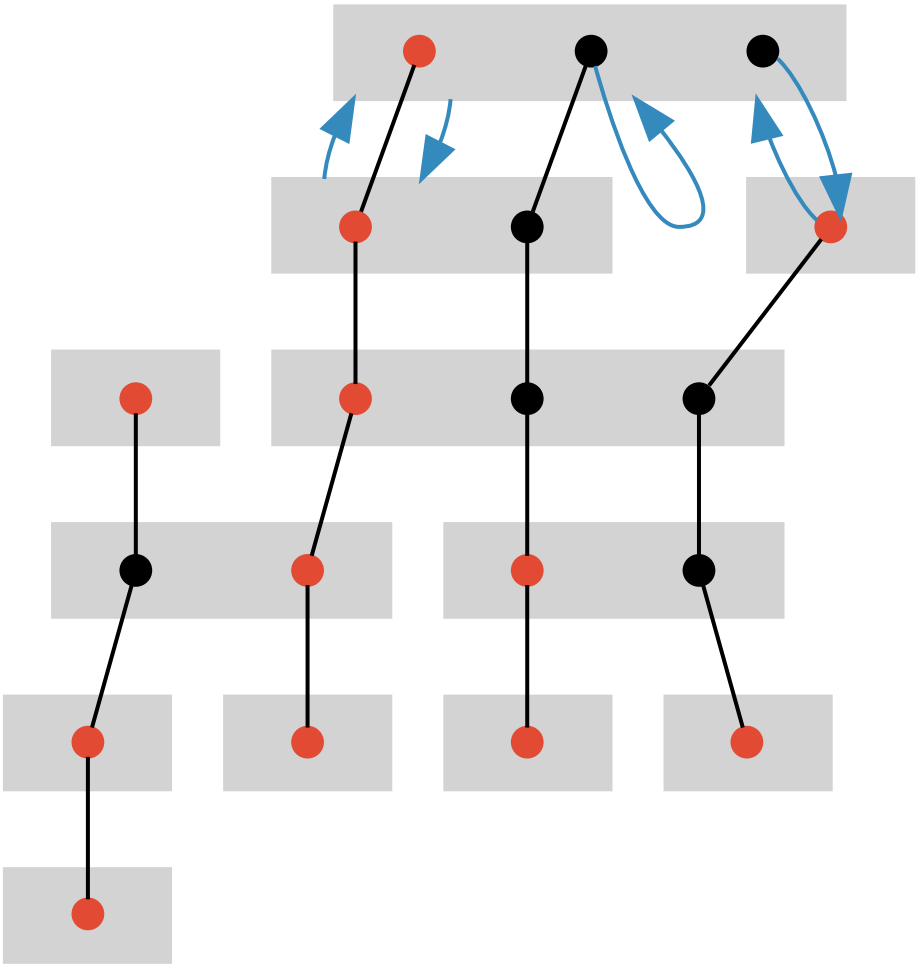

Figure 3.1: A schematic representation of a merger tree with two defects: a re-merger (halo in the second row down, on the right) and a fly-by (halo in the third row down on the left). Grey rectangles represent haloes, and dots subhaloes; every halo has one main subhalo, marked with a red dot; subhaloes are matched between snapshots (black lines) by following the most bound particles. The blue arrows indicate the relationships relevant in building merger trees, and represent (left to right): (i) halo descendant, (ii) halo progenitor, (iii) host of a subhalo, (iv) host of subhaloes' progenitor, (v) descendant of a subhalo. This plot can be compared with similar diagrams included in Thomas et al. (2015); Han et al. (2018).

the collapsed mass history (which sums over the masses of progenitors at any given snapshot, and as such is not affected by the order or the sequence of the mergers). This is similar to the scheme used to build merger trees by Jiang et al. (2014).

2. Fly-bys (e.g. the branch merging into, and then leaving, the left-most halo in the fourth row down in Fig. 3.1) happen when a subhalo is identified as a part of a FoF halo for one or more snapshots due to a temporary spatial overlap, but later becomes an isolated halo again. The presence of fly-bys pollutes the CMH, artificially inflating the mass at snapshots with extra subhaloes.

Both defects can be avoided by only keeping those haloes in the tree which merge as the main subhaloes of the host in the preceding snapshot (which would remove both example defects shown in Fig. 3.1). This is not equivalent to keeping only the main branch of the halo mass history – the full CMH is still used, but it is calculated from a pruned merger tree.

3.3 Results

Our goal is to determine the relation between halo concentration (or more specifically $\langle\rho_{-2}\rangle$) and the critical density at the formation time z_{-2} (namely $\rho_{\rm crit}(z_{-2})$) for haloes of different masses at different redshifts. For each mass bin we construct the median density profile and CMH, which are used to estimate median concentration and formation time. This approach has the benefit of producing smoother profiles, and in turn a smoother density-density relation, as is evident from comparing Fig. 3 from Ludlow et al. (2013) and Fig. 7 from Ludlow et al. (2016).

This section outlines the details of each step of our analysis. The source code used for the analysis is publicly available[4].

[4] `https://doi.org/10.5281/zenodo.2593623`

3.3.1 Filtering & binning

Our halo catalogues are obtained by filtering the HBT+ output and retaining objects with a minimum of 20 particles. Since we are interested in resolving the merger history of haloes down to progenitors with $f = 0.02$ times their final mass, this places a lower limit of $n_{200} = 10^3$ on the number of particles a halo must contain in order to be included in our analysis.

Haloes are divided into bins that are equally-spaced in $\log_{10}(M_{200}/[h^{-1}M_\odot])$, with $\Delta \log_{10}(M_{200}/[h^{-1}M_\odot]) = 0.162$. To identify potentially unrelaxed systems we use the centre-of-mass offset parameter,

$$d_{\rm off} = \frac{|\mathbf{r}_p - \mathbf{r}_{\rm CM}|}{r_{200}}, \tag{3.5}$$

where $\mathbf{r}_p$ is the centre of potential, and $\mathbf{r}_{\rm CM}$ the centre-of-mass (Thomas et al., 2001; Maccio et al., 2007; Neto et al., 2007). Only haloes with $d_{\rm off} < 0.07$ are retained for analysis.

The fitting of mass profiles (Section 3.3.2) and calculation of formation times (Section 3.3.4) is performed on the median mass profiles and CMHs, respectively, for each mass bin.

3.3.2 Fitting mass profiles

The cumulative mass profile is defined using all particles within r_{200}, and not only those deemed *bound* to the main halo or its subhaloes. These particles are assigned to logarithmically spaced radial bins, within which enclosed masses are computed. The mass profiles of haloes in each mass bin are assigned in this way, and their median is calculated. Finally, the median mass profile is normalised by the total median enclosed mass, $M_{200} = M\,(r < r_{200})$. The best-fitting value of the concentration, c, is obtained by minimising

$$\chi^2 = \sum_{i=0}^{20} \left[\log_{10}\left(M_i\right) - \log_{10}\left(M\left(r < r_i, c\right)\right)\right]^2, \tag{3.6}$$

where M_i is the mass measured within r_i, $M(< r, c)$ is the mass enclosed within radius r for an NFW profile with a concentration c (Eq (3.1)); quantities with subscript i refer to the $i^{\rm th}$ bin in $\log_{10}$ radius from the halo centre.

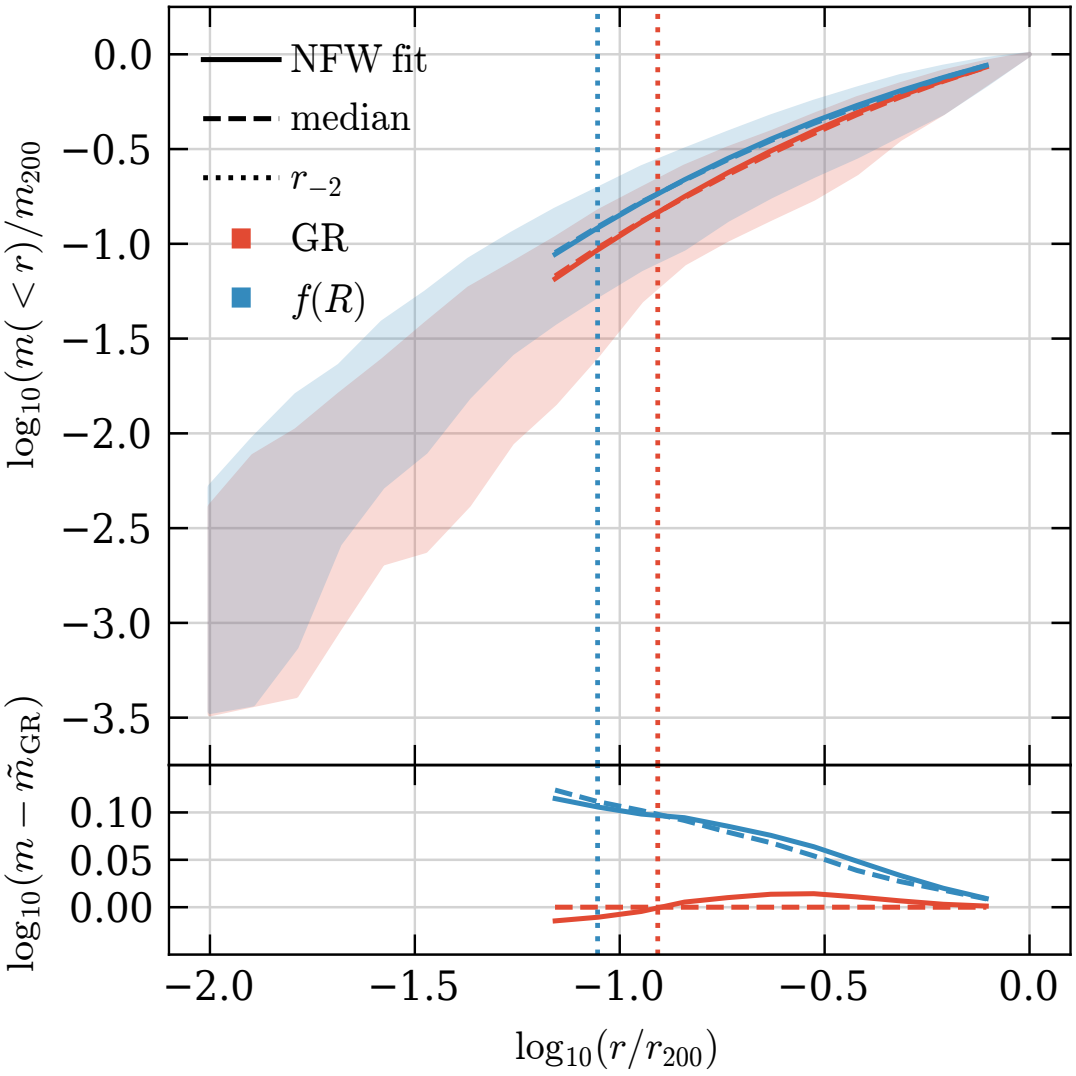

Figure 3.2: Radial enclosed mass profiles for haloes in the mass range $11.5 < \log_{10}(M_{200}/[h^{-1}M_\odot]) < 11.7$ at $z_0 = 0$. GR and $f(R)$ runs are shown using red and blue curves, respectively, as indicated in the legend; residuals from GR are shown in the lower panel. The faint shading shows the envelope of the individual mass profiles; dashed lines show median mass profiles; solid lines show the best fitting NFW profiles to the median mass profiles, for radii between $r_{\rm min} < r < r_{\rm max}$; vertical dotted lines show the characteristic scale r_{-2}. Residuals are taken from the median mass profile of GR haloes, $\tilde{m}_{\rm GR}$.

We have used both NFW and Einasto profiles in our analysis. Results for NFW profiles are provided in Figs. 3.2 and 3.5, and for Einasto profiles in the next subsection, in Figs. 3.3 and 3.4. In particular, Table 3.2 shows that the quality of fit does not improve sufficiently to warrant using the Einasto profile (which has an extra parameter) over NFW. We emphasise that the choice of analytic density profile does not change our results or conclusions.

Our fits to Eq (3.6) are minimised over the radial range $r_{\rm min} < r_i < r_{\rm max}$, where $r_{\rm min}$ is a minimum fit radius, and $r_{\rm max}$ is set to $0.8 \times r_{200}$ to exclude the unrelaxed outer edges of haloes (Ludlow et al., 2010). We consider two definitions of $r_{\rm min}$:

1. half of the mean particle separation within r_{200} (Moore et al., 1998),

$$r_{\rm min} = \frac{1}{2}\left(\frac{4\pi}{3n_{200}}\right)^{1/3} r_{200}, \tag{3.7}$$

   where n_{200} is the number of particles enclosed within r_{200}, and

2. the radius at which the two-body relaxation time is equal to the age of the universe, t_0 (Power et al., 2003; Ludlow et al., 2019), which can be approximated by the solution to

$$\frac{t_{\rm relax}\,(r)}{t_0} = \frac{\sqrt{200}}{8}\frac{n\,(<r)}{\ln\left(n\,(<r)\right)}\left(\frac{\langle\rho\,(<r)\rangle}{\rho_{\rm crit}}\right)^{-1/2}. \tag{3.8}$$

   Here $n\,(<r)$ is the number of particles enclosed by radius r and $\langle\rho\,(<r)\rangle$ is the mean enclosed density, $\langle\rho\,(<r)\rangle = 3m\,(<r)\,/4\pi r^3$.

Although we have considered both options, results are shown for the Moore et al. (1998) definition as it is typically more conservative than the alternative. Henceforth, all $r_{\rm min}$ values are calculated using Eq (3.7).

Once c is found, M_{-2} can be calculated from Eq (3.4); the characteristic density of the halo is then given by $\langle\rho_{-2}\rangle = 3\,M_{-2}/4\,\pi\,r_{-2}^3$.

3.3.3 Einasto profile

The Einasto density profile (Einasto, 1965) can be expressed as

$$\ln\left(\frac{\rho}{\rho_{-2}}\right) = -\frac{2}{\alpha}\left[\left(\frac{r}{r_{-2}}\right)^{\alpha} - 1\right], \tag{3.9}$$

where r_{-2} is a scale radius (at which where the logarithmic slope of the density profile is equal to -2), and α is a "shape" parameter. Einasto function is commonly used for modelling dark matter halo density profiles, and in many cases it has been shown to perform at least as well as the NFW profiles, even with one parameter fixed at $\alpha = 0.17$ (Merritt et al., 2006; Ludlow & Angulo, 2016).

Fits using both NFW and Einasto density profiles have been performed for comparison. We have computed and compared model selection criteria, called Akaike Information Criterion (AIC) and Bayesian information Criterion (BIC), as an objective way to determine if the additional parameter in the Einasto profile is justified in terms of improved fits to the simulation results (Akaike, 1974; Schwarz, 1978). The AIC and BIC measures take into account the χ^2 value of the fit and the number of free parameters. The fit with the smallest value of AIC or BIC is deemed to be the most appropriate one to use[5].

The Einasto density profiles for an illustrative mass bin at $z_0 = 0$ are shown in Fig. 3.3; the CMHs for the same mass bin at $z_0 = 0$, for values of F calculated from Einasto concentrations, are shown in Fig. 3.4. Table 3.2 shows values the values of the AIC and BIC statistics for the NFW fits from Section 3.3.2 and the Einasto fits from this section. Despite the fact that the Einasto profile produces a better fit, it has an extra free parameter, which yields higher values of the information criteria. Moreover, despite producing a higher quality fit, z_f values for GR and $f(R)$ haloes calculated from the Einatsto profile (see Fig. 3.4) are still indistinguishable, like the ones calculated from the NFW profile (see Fig. 3.5). This indicates that the NFW profile is the more justified choice.

3.3.4 Calculating halo formation times

The mass growth history of a dark matter halo, $m(z)$, can be defined in different ways. The mass assembly history (MAH) is the mass history of a halo obtained using

[5]There is a subtle difference between the AIC and BIC statistics. BIC introduces a higher penalty for more complicated models; however, this is only important if the criteria give conflicting results, which is not the case here

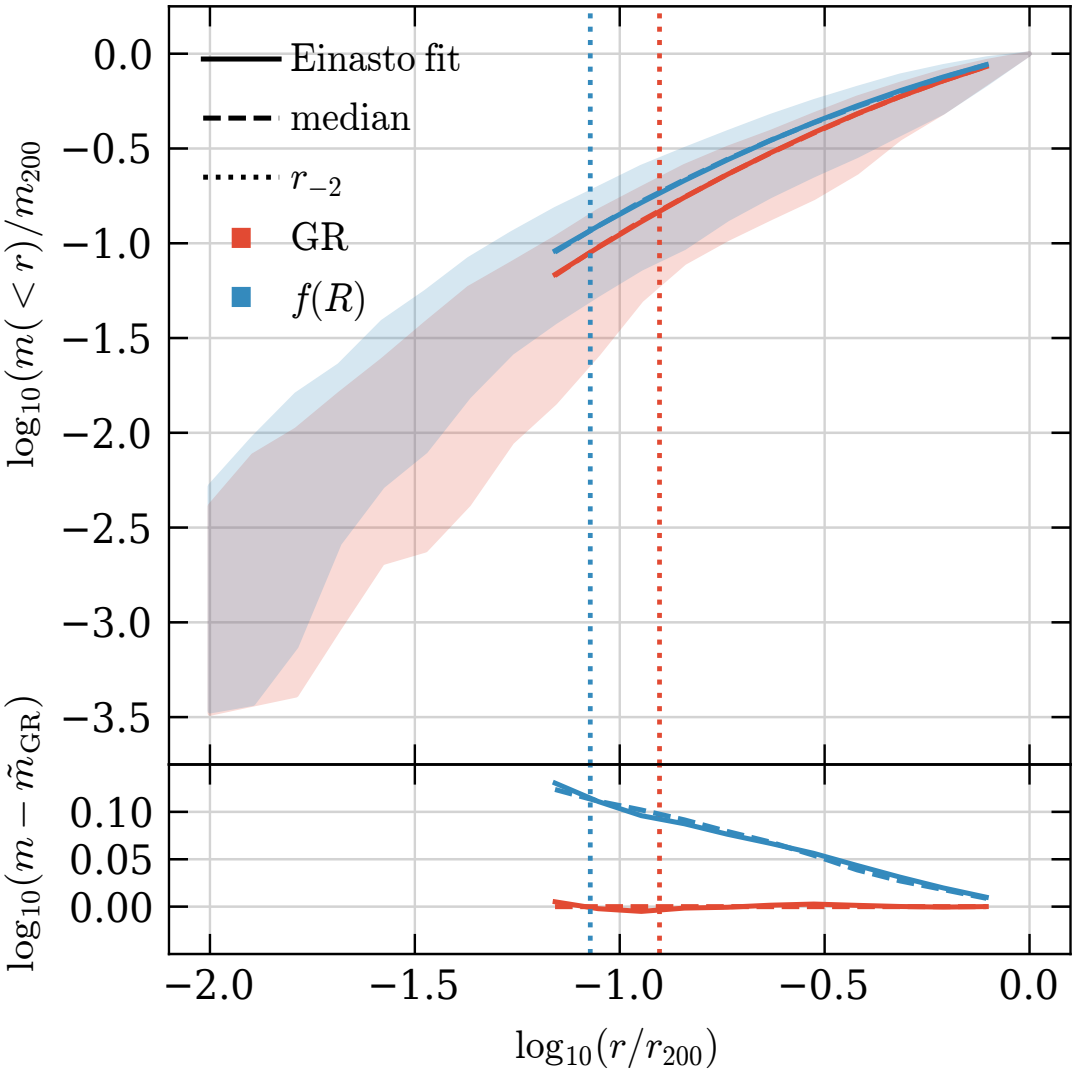

Figure 3.3: Like Fig. 3.2, but fit to the Einasto density profile.

Table 3.2: Goodness-of-fit comparison between the NFW and Einasto density profiles for haloes with masses in the range $11.5 < \log_{10}(M_{200}/[h^{-1}M_\odot]) < 11.7$ at $z_0 = 0$ for the GR run.

	NFW	Einasto
number of parameters	1	2
AIC	2.002	2.399
BIC	4.002	4.797

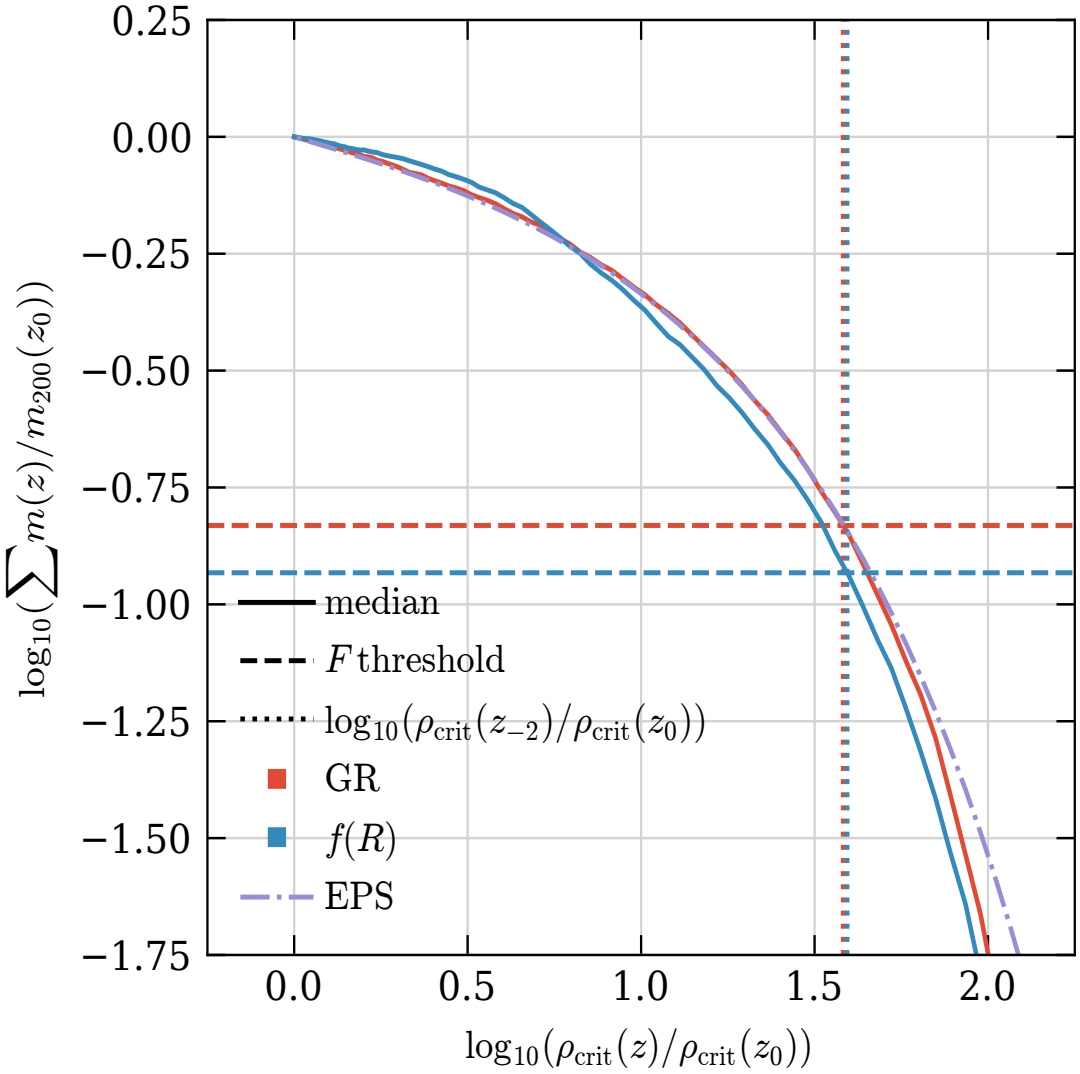

Figure 3.4: Like Fig. 3.5, but for the values of F calculated from Einasto concentrations from Fig. 3.3.

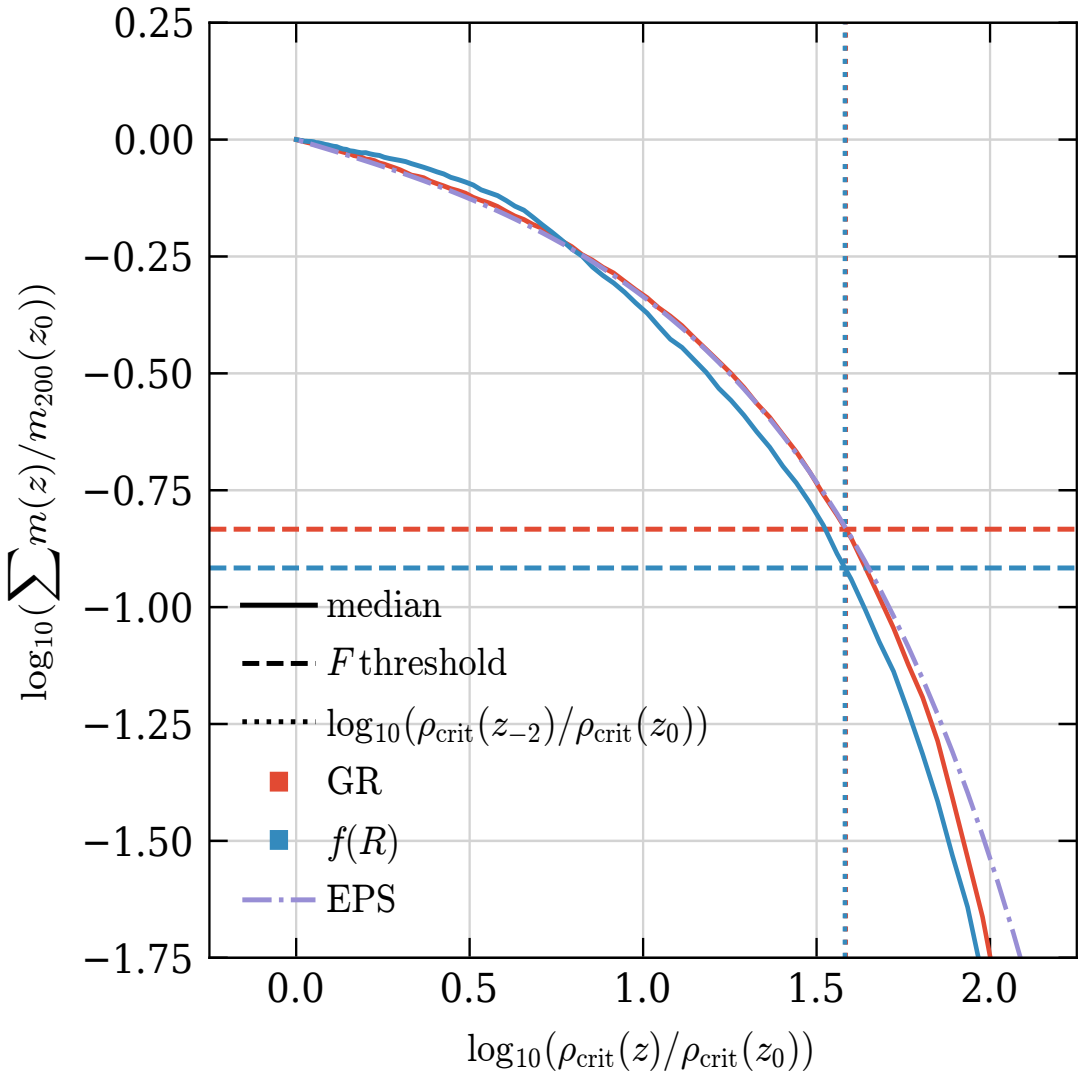

Figure 3.5: Median collapsed mass histories (CMHs) for haloes in the mass range $11.5 < \log_{10}(M_{200}/[h^{-1}M_\odot]) < 11.7$ at $z_0 = 0$. As with Fig. 3.2, GR and $f(R)$ runs are shown using red and blue lines, respectively. Solid lines show the median collapsed mass histories; two dotted vertical lines indicate the formation times, z_{-2}, at which the CMHs drop below a fraction $F = M_{-2}/M_{200}$ of the virial mass at z_0, plotted as two horizontal dashed lines. Note that despite the difference in F, derived from concentrations shown in Fig. 3.2, z_{-2} values are indistinguishable. The purple dashed-dot line shows the EPS prediction from Eq (3.3) for M_0 equal to median mass in this bin.

a "greedy" algorithm, by following the most massive (or main) progenitor through all snapshots and storing its M_{200}. As discussed previously, the collapsed mass history (CMH) is defined as the sum of the masses $M_{200,i}$ of every progenitor i whose virial mass exceeds $f \times M_0$, where f is a model parameter (we use 0.02 as our default value, but consider alternatives as well), and M_0 is M_{200} of the root halo. The CMH therefore takes into account all branches of the merger tree at a given snapshot.

The CMH can be obtained by gathering all progenitors of a halo from the merger tree, and summing over the masses of the progenitors at each snapshot. However, for performance reasons, the step of building a tree can be skipped in favour of searching for all progenitors of a root halo at each preceding snapshot. In other words, since the halo masses are summed over, it is not the structure of the merger tree that matters but its members.

Once the median CMH is calculated for each mass bin, it is normalised by the final mass M_0 at redshift z_0. For each mass bin, a formation time z_{-2} can then be calculated. This is defined as the time at which the CMH first exceeds a fraction $F = M_{-2}/M_0$ (calculated for a given mass bin from Eq (3.4)) of the final mass, M_0:

$$z_{-2} = z \ni \frac{M(z)}{M(z_0)} = F. \tag{3.10}$$

The formation time may be ill-defined for non-monotonic assembly histories. The monotonic behaviour of the CMH, while difficult to guarantee for individual haloes, is in practice obtained by considering the median CMH of all haloes in each mass bin (as opposed to calculating formation times of individual haloes, binning the haloes and then taking the median in each mass bin). As simulations have a finite number of outputs, and hence finite time resolution, the value of the formation time is obtained using linear interpolation between the snapshots which are immediately before and after the crossing of the formation threshold fraction.

Examples of the median CMHs for $z = 0$ haloes in a narrow bin of M_{200} are shown in Fig. 3.5. Solid red curves correspond to our GR simulation, and blue to $f(R)$. An analytic prediction from Eq (3.3), as discussed in Ludlow et al. (2016), is plotted in a purple dashed-dot line; the result agrees quite well with the CMHs obtained from *both* simulations. For example, the formation times, z_{-2} (vertical

dotted lines of corresponding color), agree with one another to $\approx 5\%$. Nevertheless, despite similarities in CMHs, these haloes *do not* have similar concentrations. The horizontal dashed lines correspond to M_{-2}/M_{200}, which show clear differences; indeed, concentration is 30% larger in $f(R)$ than in GR.

3.3.5 The density–density relation

The above analysis was carried out at $z_0 = 0$, 0.5, 1, 2 and 3. At each snapshot, haloes were filtered as described in Section 3.3.1, and binned into 20 logarithmically spaced mass bins spanning the range $\log_{10}(M_0/[h^{-1}M_\odot]) = 11.18$ to 14.42. Median mass profiles and CMHs of haloes, normalised by M_0, were used to calculate the concentration, c, and formation time, z_{-2}, for each M_0 and z_0. These were then converted to their equivalent values in "density space": c expressed in terms of the characteristic density $\langle\rho_{-2}\rangle$ (following Eq (3.1)), and z_{-2} in terms of the critical density, $\rho_{\rm crit}\,(z_{-2})$; both are then normalised by $\rho_{\rm crit}\,(z_0)$.

As shown in Figs. 3.6 and 3.7, the $\langle\rho_{-2}\rangle - \rho_{\rm crit}(z_{-2})$ relation for F6 haloes is similar to that in GR for most densities, but displays a steepening at high formation redshifts where $\langle\rho_{-2}\rangle$ increases more rapidly than $\rho_{\rm crit}(z_{-2})$. This effect is most apparent at lower redshifts (Fig. 3.6) and for lower masses (Fig. 3.7). For instance, only $f(R)$ halo mass bins with $\log_{10}(M_{200}/[h^{-1}M_\odot]) \lesssim 11.9$ at $z_0 = 0.5$, and with $\log_{10}(M_{200}/[h^{-1}M_\odot]) \lesssim 12.2$ at $z_0 = 0$ have $\log_{10}\left(\rho_{-2}/\rho_{\rm crit}\,(z_0)\right) > 4.25$, as shown by Figs. 3.6 and 3.7. This is consistent with the results found by Shi et al. (2015) for the concentration-mass and formation time-mass relations: while the formation times show small systematic differences between GR and F6, the biggest discrepancy between the two is in the form of the concentration-mass relation at low halo masses.

The concentrations recovered in the F6 model are *higher* for lower mass haloes than in GR, as demonstrated by Fig. 3.2; this change is in the opposite sense to that seen on changing CDM for WDM. In both WDM and F6, however, low mass haloes systematically form later than their GR counterparts. In F6 gravity, although there is a systematic delay in formation histories for low-mass haloes, it is not captured by the formation time defined as in Eq (3.10).

It follows that, while in WDM the formation time-concentration relation is the

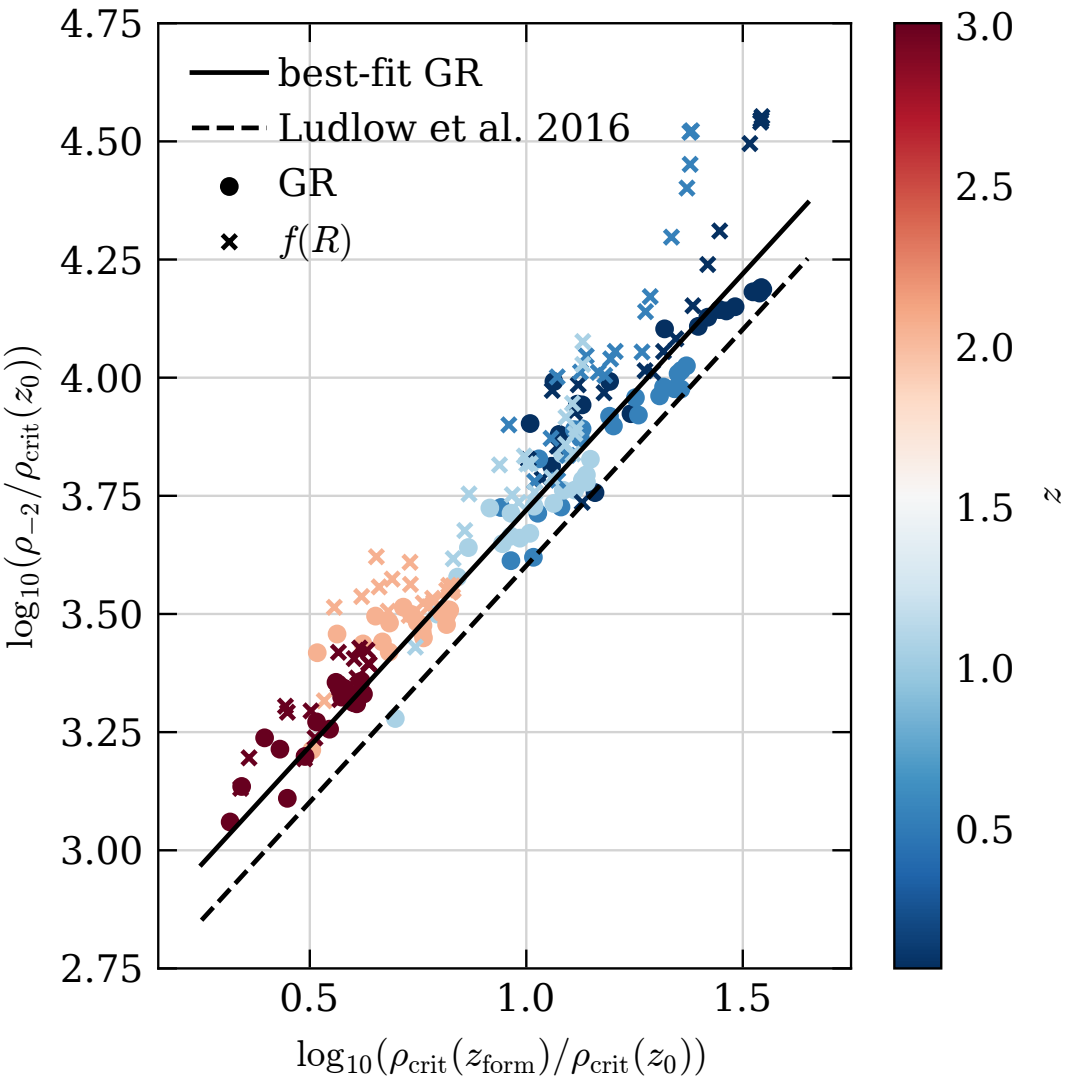

Figure 3.6: Mean enclosed density $\langle\rho_{-2}\rangle$ within the characteristic radius, r_{-2}, versus the critical density at the formation redshift, $\rho_{\rm crit}(z_{-2})$, at which a fraction $F = M_{-2}/M_0$ of the root halo mass M_0 was first contained in progenitors more massive than $f \times M_0$. Each point corresponds to median value in a logarithmically-spaced mass bin at the identification redshift z_0. All densities are normalised by $\rho_{\rm crit}\,(z_0)$, the critical density at z_0. Point types indicate the results from different gravities, as labelled. Colours indicate the identification redshift, as shown by the colour bar. Also plotted are two lines: a dashed black one which shows the Ludlow et al. (2016) scaling relation $\langle\rho_{-2}\rangle = 400 \times \rho_{\rm crit}(z_{-2})$, and a solid black one for the best-fitting GR relation $\langle\rho_{-2}\rangle = 525 \times \rho_{\rm crit}(z_{-2})$.

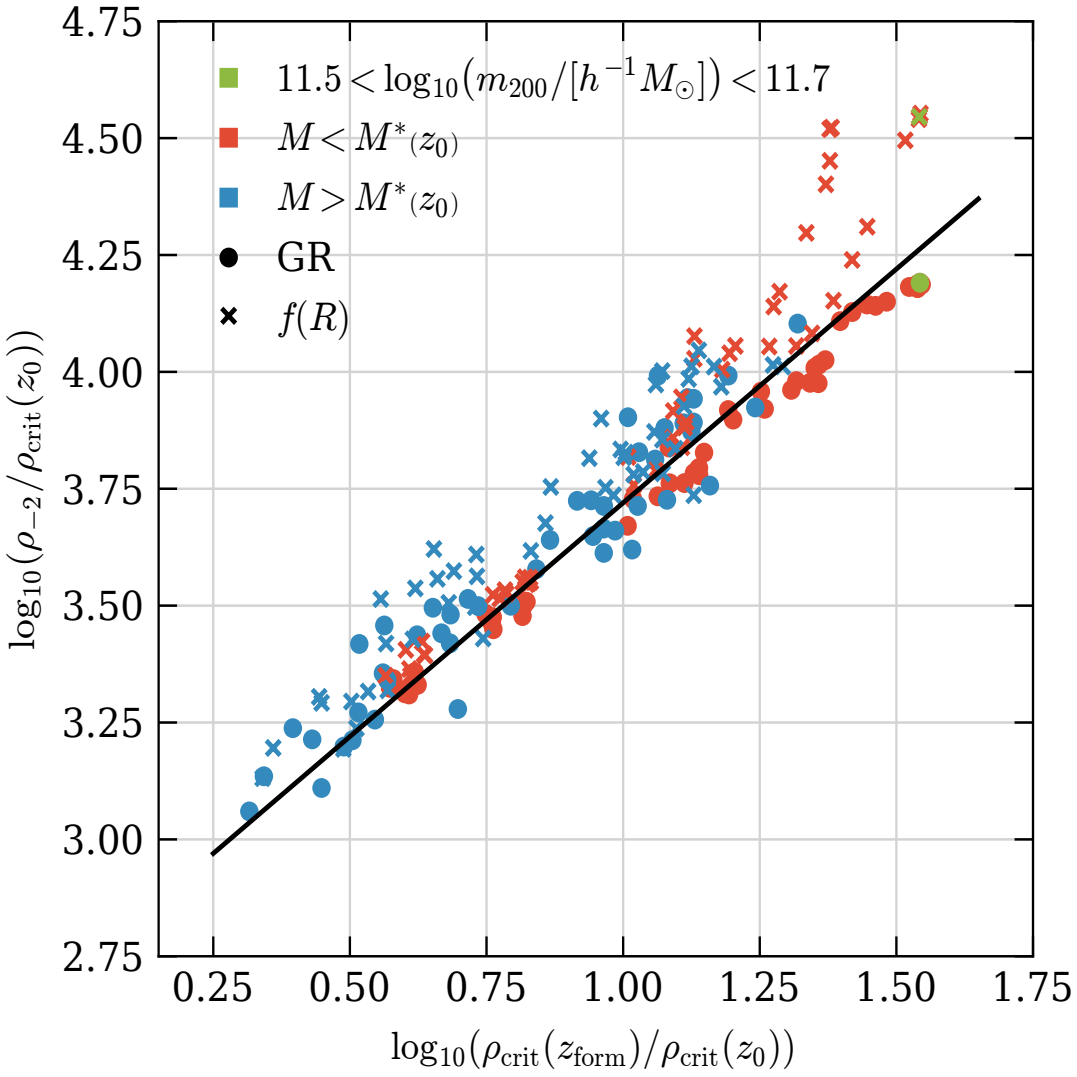

Figure 3.7: Same as Fig. 3.6, but colour–coded to indicate different halo mass ranges. The halo population has been split into two samples: one above and ones below the characteristic mass, $M^*(z_0)$, defined as $\delta_{\rm sc}(z_0)/\sigma(M^*(z_0)) = 1$ (Mo et al., 2010, Eq. 7.48). The mass bin containing haloes from Figs. 3.2 and 3.5 at z_0 is highlighted in green.

same as it is in CDM (when $z_{\rm form}$ is appropriately defined), this is not the case in $f(R)$ gravity. Even a model with an effective screening mechanism, such as F6, affects the low mass haloes identified at late times; these objects have slightly delayed formation times and notably higher concentrations, which leads to the differences between F6 and GR shown in Figs. 3.6 and 3.7.

Finally, we note that the $\langle\rho_{-2}\rangle - \rho_{\rm crit}(z_{-2})$ relation found in the GR simulation is very similar to the one reported by Ludlow et al. (2016), but with a higher intercept, ≈ 525, as shown by the solid line in Fig. 3.6. The origin of this value, which is the only free parameter of their model, is not known. It is analogous to the free parameter of the Ludlow et al. (2014a) and Correa et al. (2015) models, who also report different values. The intercept may be determined by a number of physical process and a detailed investigation of what determines its value, while worthwhile, is beyond the scope of this work.

3.3.6 Sensitivity to variation of model parameters

The parameters used to construct the CMHs (and hence to estimate z_{-2}) and to define halo characteristic densities can be varied to assess their impact the form of the $\langle\rho_{-2}\rangle - \rho_{\rm crit}(z_{-2})$ relation, and to potentially improve our understanding of the origin of the difference between F6 and GR. A few such variations have been performed: first, we modify the radius defining halo characteristic densities (using $0.3 \times r_{-2}$ and $2.0 \times r_{-2}$), and second, the mass threshold f of progenitors included in the CMH (which is varied from 0.01 to 0.1).

The results, presented in Figs. 3.8 and 3.9, confirm our intuition: increasing the progenitor mass used to construct the CMHs (by increasing f) brings the formation time closer to the identification time, z_0 (the difference is more pronounced at lower redshifts, due to the normalisation used), while increasing the radius within characteristic densities are defined decreases the mean enclosed density *and* brings the formation time closer to the identification redshift. While the parameters can be tweaked to decrease the scatter and remove the time dependence of the relation (see, e.g., Figures B1 and B2 of Ludlow et al. (2016)) the $f(R)$ haloes still exhibit a strong upwards trend in their concentrations–as well as a larger scatter than their

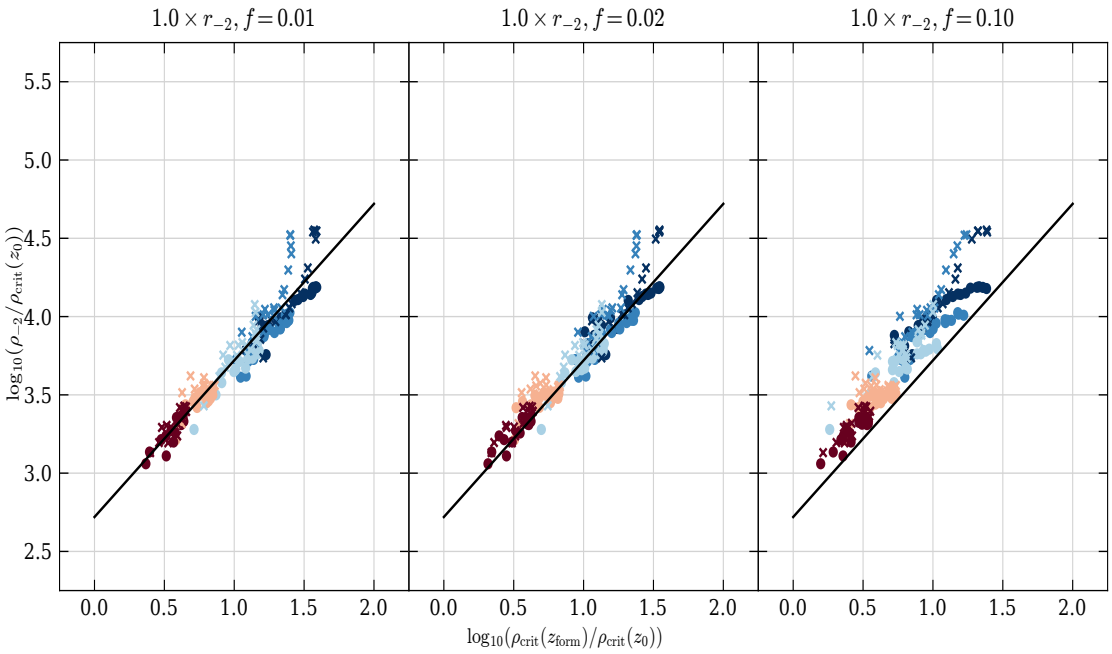

Figure 3.8: Like Fig. 3.6, but with different panels showing different collapsed mass history parameter f, as labelled above each. The solid black line shows the best-fitting GR relation, $\langle \rho_{-2} \rangle = 525 \times \rho_{\rm crit}$, and is included for comparison.

GR counterparts–for all parameter combinations. This is driven by the changes to both the $c(M, z)$ relation, and also to changes in the mass–formation time relations, which cannot be accounted for by varying the parameters mentioned above. In $f(R)$ gravity, however, the halo growth and structure are also determined by the local environment. It is therefore important to attempt to account for local effects using an environmental proxy.

3.3.7 Separation of haloes by screening

As discussed in Section 2.1, $f(R)$ gravity only affects haloes which are outside screened regions, while the screened ones grow in a manner that is largely indistinguishable from GR. It is clear from Fig. 3.6 that low mass haloes are typically the ones displaying the most prominent differences between the two simulations, implicating the fifth force as the root cause. However, it is natural that each mass bin contains both screened and unscreened objects.

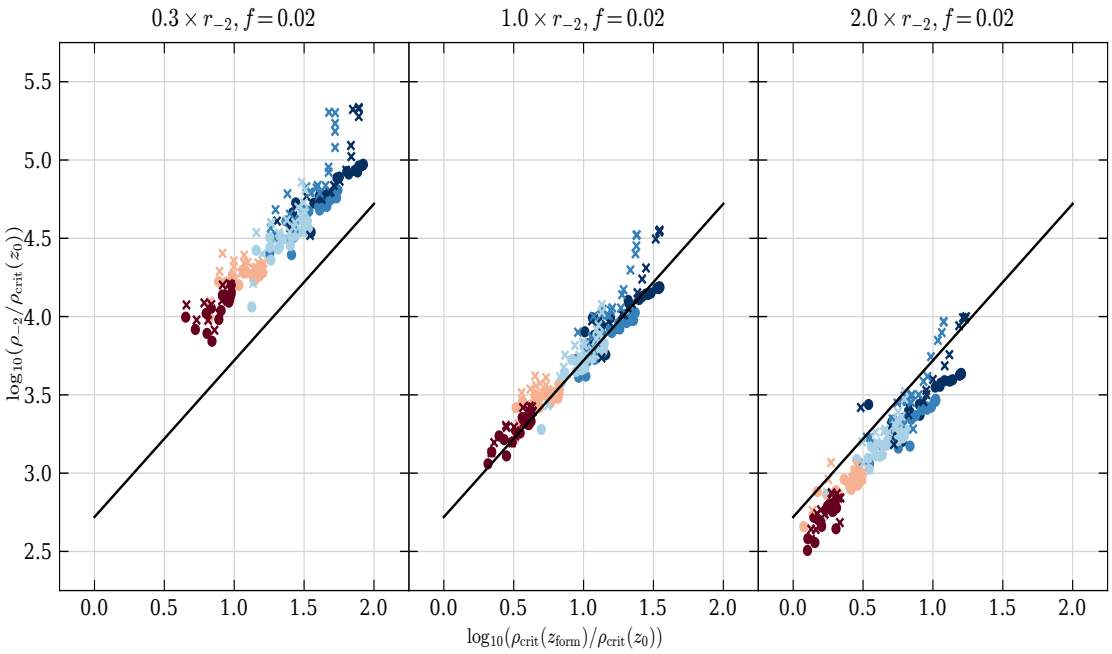

Figure 3.9: Like Fig. 3.6, but with different panels showing mean density at different fractions of a characteristic radius r_{-2}. The solid black line shows, for comparison, the best-fitting GR relation, $\langle \rho_{-2} \rangle = 525 \times \rho_{\rm crit}$.

The effectiveness of the screening mechanism (not including self-screening) is directly related to the environment in which the halo is found. Following Zhao et al. (2011); Haas et al. (2012), we use a conditional nearest neighbour distance, $D_{N,f}$, as an environmental proxy. We use the proxy in an attempt to separate haloes inside each mass bin into two populations, quantifying how strong the environmental screening effect should be.

$D_{N,f}$ for a halo of mass $\bar{M}_{200}$ is defined as the distance d (normalised to $\bar{r}_{200}$) to its $N^{\rm th}$ nearest neighbouring whose mass, M_{200}, is equal to or larger than $f \times \bar{M}_{200}$. If $D_{N,f}$ cannot be calculated (for instance, for the largest halo in a snapshot) it is assumed to be equal to ∞.

Other environment proxies, such as "experienced gravity" Φ_* (Li et al., 2011) and local spherical or shell overdensity (Shi et al., 2017) have also been proposed as methods of assessing environmental impact on formation histories. Here we use $D_{N,f}$ with $N = 1$, $f = 1.0$ since it correlates strongly with other proxies, which predict similar local enhancements to the gravitational potential (Shi et al., 2017).

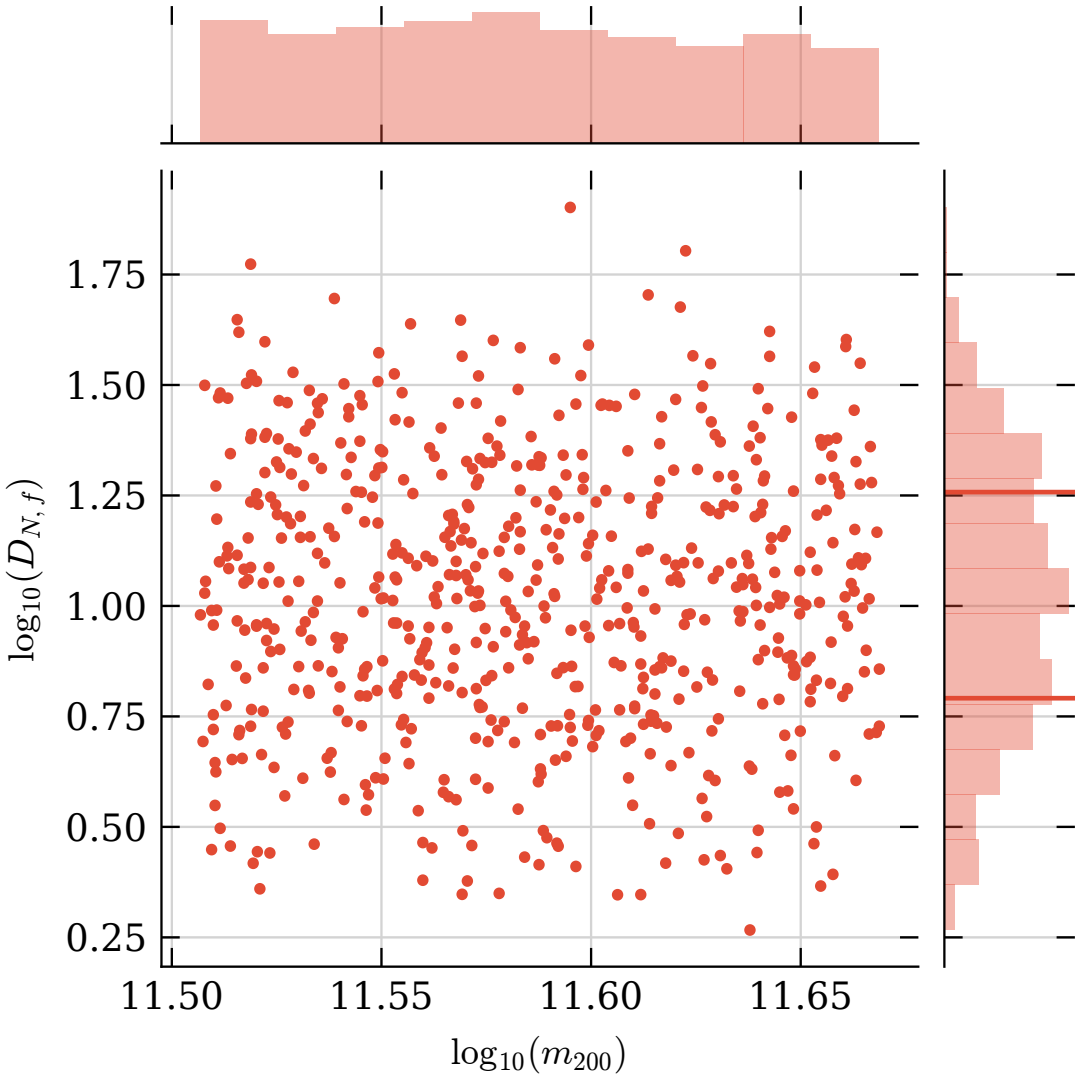

Figure 3.10: Environmental proxy $D_{N,f}$ ($N = 1$, $f = 1$) versus halo mass, M_{200}, for haloes in an example mass bin, $11.70 < \log_{10}(M_{200}/[h^{-1}M_{\odot}]) < 11.83$, at redshift $z = 0$. Distributions of $\log_{10}(M_{200}/[h^{-1}M_{\odot}])$ and $\log_{10}(D_{N,f})$ are shown at the top- and right-hand panels, respectively. The two red lines on the $\log_{10}(D_{N,f})$ histogram on the right indicate the 25^{th} and 75^{th} percentiles.

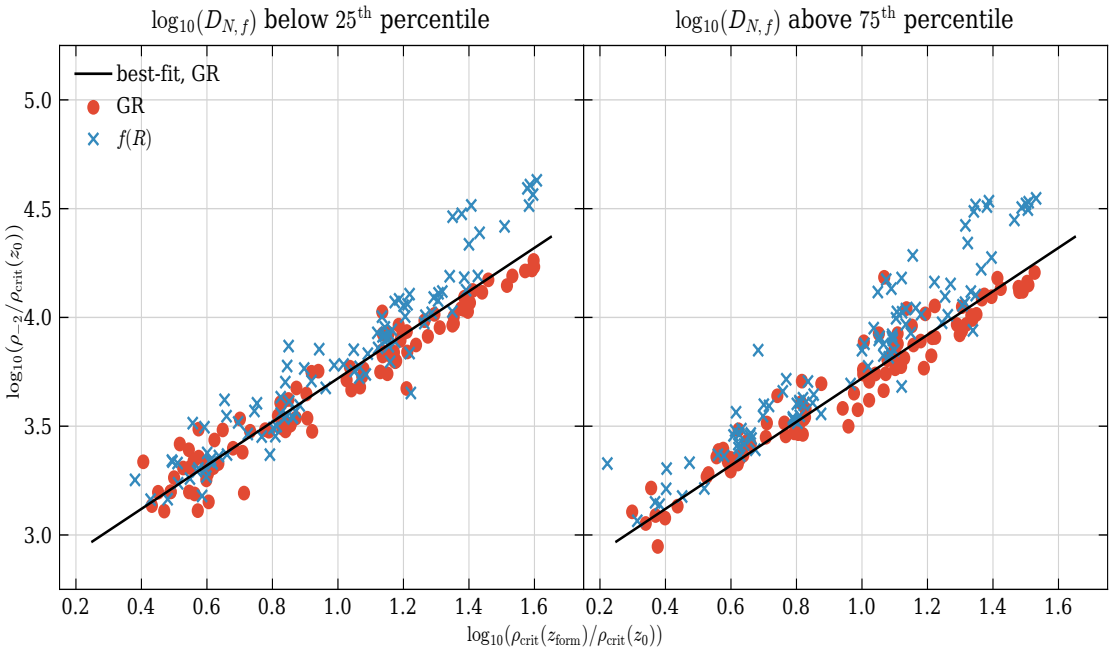

Figure 3.11: Like Fig. 3.6, but split into two populations by the environmental proxy $D_{N,f}$. The left panel shows the relation for bins including haloes below the 25th percentile; the right panel shows the same relation for bins including haloes above the 75th percentile. Colours and symbols distinguish between gravity models: red circles represent GR and blue crosses F6. Both panels include the best fitting GR relation $\langle\rho_{-2}\rangle = 525 \times \rho_{\rm crit}$ (solid black line) for reference (note that the fit is performed over the full population, regardless of the environmental proxy).

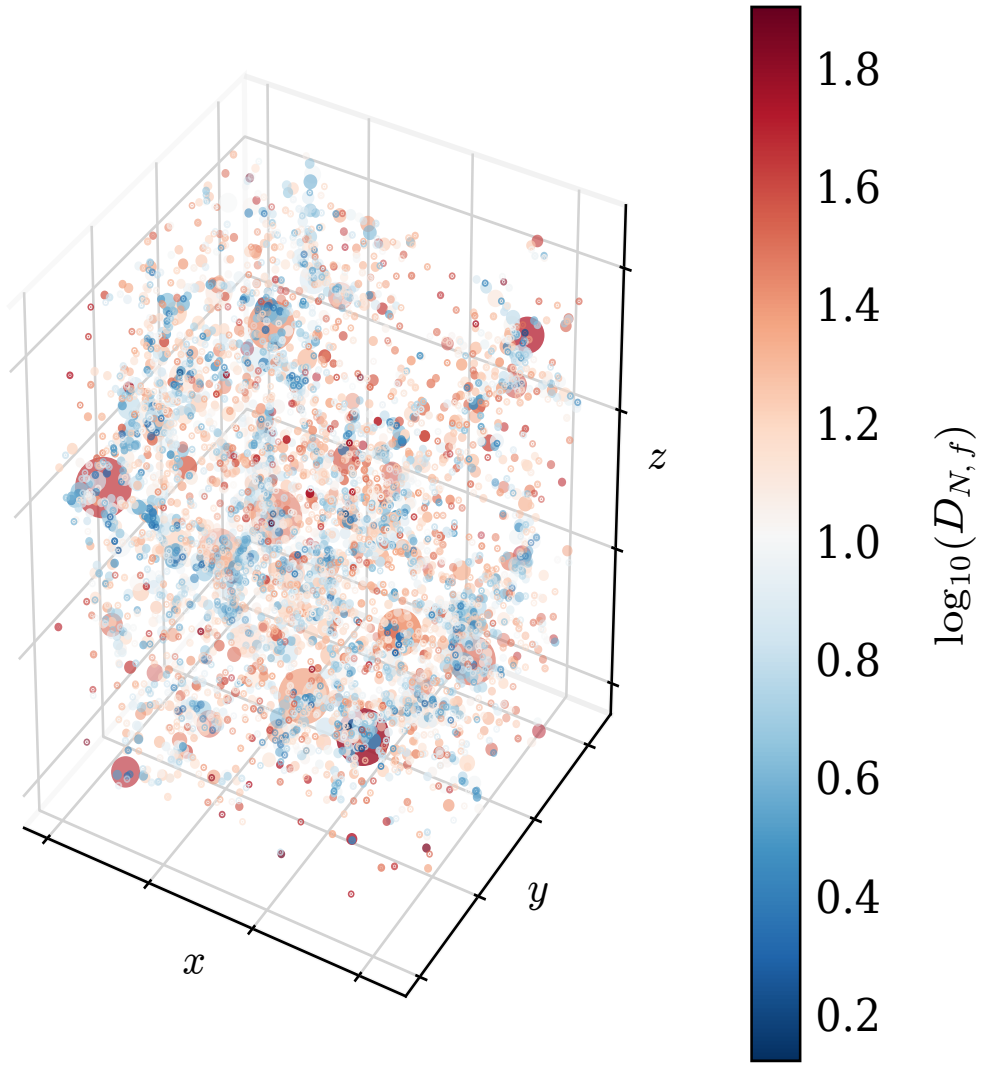

Figure 3.12: 3D visualisation of halo catalogue obtained from the $f(R)$ simulation at redshift $z_0 = 0$, consisting of over 4000 haloes. The sizes of the points reflect $log_{10}\,(M_{200})$, and the colours indicate the environmental proxy $\log_{10}\,(D_{N,f})$ used to approximate the screening and the strength of the fifth force. Note how the larger values of $D_{N,f}$ (indicating lower environmental impact on formation & evolution) are reserved for objects which are either more massive, or isolated & small, while smaller values of $D_{N,f}$ are obtained for objects which are in the vicinity of the similar-mass neighbours.

The $D_{N,f}$ values have been calculated for each halo at each redshift. Here we consider the distribution of $D_{N,f}$ in bins of halo mass focusing on the extremes of the distribution which we expect will show the biggest contrast in the efficiency of screening. The halo population at each redshift is split into two sub-groups: those below the 25^{th} and above the 75^{th} percentiles. The most massive object, with $D_{1,1} = \infty$, is excluded. The $\langle\rho_{-2}\rangle - \rho_{\text{crit}}(z_{-2})$ relations were then recalculated for each mass bin for the two sub-groups separately, and are presented in Fig. 3.11.

It is to be expected that the haloes with the lowest values of $D_{N,f}$, which are the ones that are closest to objects of comparable masses and hence in the highest density environments, will follow a concentration-formation relation closest to that displayed by GR haloes, since they are screened from the enhanced gravity. Haloes with high-$D_{N,f}$ may display a different power-law, as seen in Fig. 3.6. However, as clearly demonstrated in Fig. 3.11, while selecting haloes by their $D_{N,f}$ value has little to no effect on the GR relation, it also has little impact on the F6 haloes. This means that the difference cannot be easily accounted for by a local environmental proxy alone.

3.4 Conclusions

We have compared two high resolution dark matter only simulations, one using GR and the other F6 gravity. We constructed collapsed mass histories of haloes using their merger trees obtained from HBT+ (Han et al., 2018). We then binned the haloes by mass and calculated median enclosed mass profiles $M(r)$ and CMHs, to obtain median concentrations, c, and formation times, z_{-2}, which we used to construct the $\langle\rho_{-2}\rangle - \rho_{\text{crit}}(z_{-2})$ relation. This relation is linear in GR–and hence may be used to predict concentrations when CMHs are known–but not in F6. The differences are primarily due to a relative *enhancement* of concentration for low-mass objects in F6 which have slightly *delayed* formation times times relative to GR.

We have made several attempts to recover a linear relation from the results of the F6 simulation. For example, we varied the free parameters of the model (i.e. the fraction f of the final halo mass that a progenitor must exceed to be included in the

CMH, and fraction of the characteristic radius r_{-2} used to define the characteristic densities) to find a region in the parameter space which produces the most promising relation. While there are values of parameters which improve upon the conventional choice for GR ($f = 0.02,\ 1.0 \times r_s$), there are trade-offs with regards to scatter and gradient of the line. Furthermore, to account for the mixing of the screened and unscreened haloes in each mass bin, we split the halo catalogue into two sub-populations using an environmental proxy $D_{N,f}$, which also had little effect.

Since neither approach alone has been successful, we propose that either (1) information about haloes' sizes *as well as* environment is required, or (2) a better proxy, capable of separating haloes not only by present environment, but also by their growth histories, is required, or (3) the density-density relation in $f(R)$ is not separable into the power law and a correction.

Our overall conclusion is that the form of the concentration–formation time relation is particular to the gravitational force in the adopted cosmological model and its origin remains unknown. The key difficulty seems to lie in the question of why haloes with very similar formation redshifts can nevertheless have very different concentrations. One possibility is that the definition of formation time (z_{-2}) or assembly history (CMH)–which function well for GR models for $c(M, z)$–require amendments for $f(R)$.

Since the relation is sensitive to model parameter variation, but not to environment–based splitting, it would be interesting to further test the relation for a dependence on self-screening. This could be tested by splitting halo populations using a self-screening proxy, as well as running the analysis on other cosmologies, such as F5, F4 and enhanced (4/3 the conventional strength) gravity simulations. We believe that looking into the changes in the concentration – formation relation in different gravity regimes is a promising avenue of research into the nature and origin of the correlation between halo concentrations and formation times.

Chapter 4

Modified Semi-Analytic Galaxy Formation Model

We run the semi-analytic model (SAM) GALFORM on dark matter halo merger trees calculated from four high resolution N-body simulations, one of ΛCDM and three variants of the Hu-Sawicki form of the $f(R)$ gravity, to investigate the effects of modified gravity on galaxy properties. We explore the use of the halo-occupation distribution (HOD) for Hα emitters, as an possible discriminant between the modified gravity theories. We generate HODs for all four cosmologies at a redshift range from 0.3 to 2.2, which matches the EUCLID & WFIRST observation ranges. We observe only small differences in the predicted HODs.

4.1 Introduction

The DE theories will be constrained in the near future with observational tests by huge surveys, such as EUCLID (Laureijs et al., 2011; Amendola et al., 2013) and WFIRST (Green et al., 2012), which will probe the expansion history of the Universe in unprecendented detail. However, because galaxies are biased tracers of the large scale structure in the dark matter (Kaiser, 1984; Mo & White, 1996; Kauffmann et al., 1997), these tests depend on our understanding of galaxy formation and evolution. The new surveys will allow clustering measurements which, for the first time, will be limited by systematic errors rather than sample variance. Accurate theoretical

predictions are therefore essential to meet the scientific goals of these surveys.

In the MG theory studied here, $f(R)$, the screening mechanism depends on the mass of objects as well as their environments. We investigate three versions of the Hu & Sawicki (2007, HS) model, varying in the effectiveness of the screening mechanism. In all three cases (F6, F5 and F4 hereafter), gravity is enhanced by a factor of 1/3 compared to GR, the so-called fifth force, but the models differ with regards to the effectiveness of the screening. Nonetheless, in all three models screening is most effective at high redshift, and for the most massive haloes. Based on N-body simulations, it is expected that low mass haloes at low redshift will exhibit the biggest differences (Shi et al., 2015, 2017; Mitchell et al., 2019).

Previous theoretical and numerical studies of $f(R)$ modified gravity suggest that the concentrations and formation times of dark matter haloes (Shi et al., 2015, 2017; Oleśkiewicz et al., 2019), virial scaling relation (Fontanot et al., 2013), as well as brightness of stars (Davis et al., 2012) are expected to exhibit a small but systematic deviation from their GR counterparts.

While previous studies looked for changes in the predicted galaxy properties in statistics that cover a wide range of halo masses (e.g. the galaxy luminosity function or the global star formation rate density), we focus instead on a quantity that isolates galaxies in a narrow range of halo mass that are expected to be sensitive to the modifications to gravity. In many MG theories, including $f(R)$, the existence of the screening mechanism means that the halo population is mixed, containing both screened haloes, which feel a gravity strength equivalent to classical GR, and unscreened ones, with experience gravity that is stronger by a factor of 1/3. It is therefore imperative to look for observables which could help us separate these two classes of haloes.

One such statistic is the Halo Occupation Distribution (HOD). The HOD quantifies the bias between galaxies and dark matter, and describes how the number of galaxies per halo changes with halo mass, galaxy property, and redshift (Zheng et al., 2005; Smith et al., 2017; Contreras et al., 2019; Merson et al., 2019). The HOD for the Hα emitters likely to be seen in upcoming surveys peaks around a halo mass of $m \approx 10^{12}\, h^{-1}\, \mathrm{M}_{\odot}$ (Merson et al., 2019). The fact that this feature in

the HOD coincides with the prominent change in the concentration-mass relation in $f(R)$ (Shi et al., 2015; Mitchell et al., 2019) suggests that HODs might provide a robust observable probe of modified gravity as a viable ΛCDM alternative.

The star formation rate (SFR) is driving emission in many lines, among them OII and Hα – conversely, these lines can be used to trace SFR in the galaxy population. Gonzalez-Perez et al. (2018) predicted a peak in the HOD of OII emitters for GR, at a mass at which the screening is expected to become less effective. We expect to see a similar peak in the HODs of the Hα emitters, since they are similarly affected by the SFR (Jansen et al., 2001), and we hope that this effect will be easy to observe in modified gravity. We compare the results that the modified gravity has to what we can achieve by modifying SFR properties directly in a GALFORM model, which should result in a similar peak in an HOD. We also examine the Tully-Fisher relation and K-band luminosity function, which can be used as a probe into halo masses, virial velocities and stellar masses.

We expand on previous work – we feed the merger trees constructed from a set of high resolution N-body simulations into a semi-analytic galaxy formation model, which has been modified in two critical areas: the concentration-mass-redshift relation and virial velocity relation. However, we do not change the calculations of the hydrostatic equilibrium, as well as that of the disc and bulge sizes and stellar populations.

This Chapter is structured as follows. The theoretical background is given in Section 4.2: the N-body simulation, halo catalogue and merger tree builder are described in Section 4.2.1; the original GALFORM model and the modifications we included for an $f(R)$ version are described in Section 4.2.2; Section 4.2.3 outlines how we construct HODs. Our results, presented in Section 4.3, are always compared to the control run of a standard GALFORM model on GR merger trees, and can be divided in separate stages: first, in Section 4.3.1 we run the original GALFORM model on modified gravity merger trees; then, in Section 4.3.2, we run a modified GALFORM model, with augmented $c(m, z)$ and virial relations, on $f(R)$ merger trees; finally, in Sections 4.3.3 and 4.3.4, we include two reference runs on GR merger trees: in Section 4.3.3 we use unchanged $c(m, z)$ relation, but update it continu-

Table 4.1: Parameters of the N-body simulations used.

$\Omega_{\rm m}$	matter density	0.3089
Ω_Λ	dark energy density	0.6911
Ω_b	baryon density	0.0460
σ_8	power spectrum amplitude	0.8159
h	H_0 $[100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}]$	0.6774
L	box side $[h^{-1}\,\mathrm{Mpc}]$	62
m_p	particle mass $[h^{-1}\mathrm{M}_\odot]$	1.523×10^8
N_p	number of particles	512^3

ously (instead of just at formation times), and in Section 4.3.4 we increase the $v_{\rm hot}$ parameters to emulate increased gravitational potential of the $f(R)$ gravity.

In Section 4.4 we discuss our findings regarding luminosity function in K-band and Tully-Fisher relation (Section 4.4.1), and halo occupancy distribution (Section 4.4.2). Finally, we present our conclusions in Section 4.5.

4.2 Theory

4.2.1 N-body Simulation & Merger Trees

Since the equations of MG are non-linear, $f(R)$ simulations are more computationally expensive than comparable ΛCDM ones. However, significant progress has been made recently in designing numerical techniques specifically for this class of theories, such as `MG-GADGET`, `ECOSMOG` and `AREPO` (Li et al., 2012; Puchwein et al., 2013; Bose et al., 2015; Arnold et al., 2019b). In this Chapter we use N-body runs from the SHYBONE (Simulating HYdrodynamics BeyONd Einstein) simulation suite (Arnold et al., 2019a), consisting of 4 high-resolution boxes: one for ΛCDM and three for variations of $f(R)$ (labelled F6, F5 and F4). All runs use Planck cosmology (see Table 4.1, (Planck Collaboration et al., 2016b)), which was selected to match the calibration of the $f(R)$ correction from Mitchell et al. (2019).

Dark matter haloes were identified using the friends-of-friends (FoF) algorithm

(Davis et al., 1985), and subhaloes using the SUBFIND algorithm (Springel et al., 2001). The minimum size of subhaloes retained using SUBFIND is 20 particles. The merger trees were then constructed from the halo catalogue using the DHalo algorithm (Jiang et al., 2014), by identifying the unique descendant of each subhalo at the subsequent snapshot. The descendant of a subhalo is defined as the subhalo at the next snapshot which contains the greatest number of particles from the given subhalo.

4.2.2 GALFORM

GALFORM is a semi-analytic model (SAM) of galaxy formation and evolution (Cole et al., 2000; Bower et al., 2006; Lacey et al., 2016; Baugh et al., 2019). It is capable of reproducing many astrophysical observables, such as the luminosity function, stellar mass function and the Tully-Fisher relation, starting from dark matter halo merger trees. GALFORM populates the merger trees with galaxies by numerically solving equations that describe the bulk flow of gas, gas heating & cooling, star formation processes, as well as feedback processes such as supernovae (SNe) driven winds and heating by active galactic nuclei (AGN).

GALFORM is built up of multiple modules which implement different models of astrophysical processes. These modules can require parameters to be specified which regulate the physical process (for a detailed description of the GALFORM parameters see Lacey et al. (2016)), and all these parameters taken together define the "GALFORM model". Here we use the recalibration of the Lacey et al. (2016) model for the P-Millennium N-body simulation introduced by Baugh et al. (2019). This choice was made because both the P-Millennium and SHYBONE N-body simulations use the same Planck cosmology (Planck Collaboration et al., 2016b).

The halo merger history alone is expected to be similar between the gravity theories tested (Shi et al., 2015), and so should not have a significant impact on the galaxy properties. Halo properties such as concentration and circular velocities, on the other hand, are expected to display significant changes between gravity models (Shi et al., 2015; Mitchell et al., 2019). Therefore, we have implemented modifications to the dark matter halo models in GALFORM to model the effects of the modified

gravity: correcting the concentration-mass-redshift and virial relations.

We elect to modify the relations and distributions used for haloes in GALFORM rather than adopting the values measured directly from the N-body as these quantities are subject to noise (Bett et al., 2007). This is a better match for our use case of modelling $f(R)$ effects on baryonic processes than relying on the data measured from particle data by the halo finder, since the use of analytical models guarantees that even haloes down to the resolution limit of 20 particles have consistent values for their concentrations and velocities.

To isolate the causes of any changes in the model predictions, we have used this step-by-step approach and kept our modifications minimal and self-consistent. To understand why we have chosen to change these particular components of GALFORM (but not any others), one must remember that concentration-mass-redshift relation controls any calculation which depends on the sizes of the haloes, and similarly the velocity dispersion controls many physical processes; the virial relation has a direct impact on supernovae feedback (as described in Section 4.2.2), which in turn has a direct effect on the stellar mass function for lower mass galaxies (Cole et al., 2000; Bower et al., 2006). Moreover, other parts of the semi-analytic calculation (such as the AGN feedback) are typically relevant in the regime in which modified gravity effects are screened, and we can therefore safely assume that for these models, the equations used in the GR version of GALFORM are still a good approximation.

Concentration-redshift-mass relation

The concentration-mass-redshift relation of the dark matter haloes in GALFORM was amended to model the change in halo concentrations reported from N-body $f(R)$ simulations by Shi et al. (2015). In particular, we have implemented the Mitchell

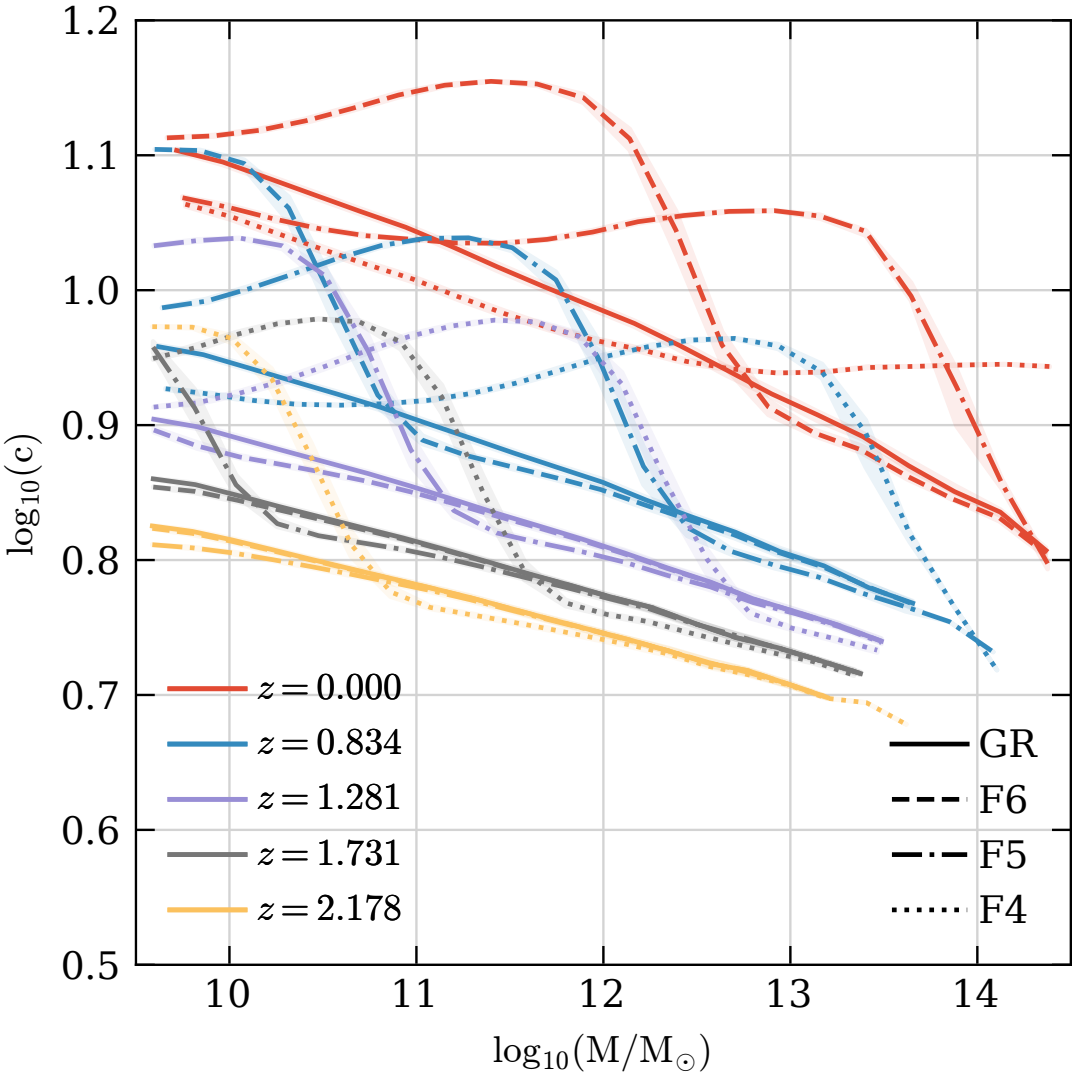

Figure 4.1: The concentration-mass-redshift relation for the dark matter haloes in the four gravity models, GR, F6, F5 and F4 (distinguished with line styles as described in the legend) at five different redshifts (colour-coded as described in the legend). Haloes are binned into 20 logarithmically spaced mass (M_{200}) bins, and the median values have been calculated in each bin. The concentrations values were computed by GALFORM using an NFW prescription, and corrected using the model of Mitchell et al. (2019), as shown in Eq (4.1).

et al. (2019) model that describes these changes:

$$\log_{10}(c/c_{\rm GR}) = \frac{1}{2}\cdot\left(\frac{\lambda}{\omega_{\rm s}}\phi(x')\left[1+{\rm erf}\left(\frac{\alpha {\rm x}'}{\sqrt{2}}\right)\right]+\gamma\right)\cdot \\ \cdot\left(1-\tanh\left(\omega_{\rm t}\left[x+\xi_{\rm t}\right]\right)\right) \tag{4.1}$$

$$x' = (x-\xi_{\rm s})/\omega_{\rm s} \tag{4.2}$$

$$p_2 = 1.503\log_{10}\left(\frac{|\bar{f}_R|}{1+z}\right)+21.64 \tag{4.3}$$

$$x = \log_{10}(M_{500}/10^{p_2}), \tag{4.4}$$

where $c_{\rm GR}$ is the base concentration value, ϕ is a normal distribution, erf is the error function, p_2 is a gravity-dependent measure of the amount of screening of a halo, $\bar{f}_R$ is the mean background value of Eq (2.3), and ($\lambda = 0.458$, $\xi_{\rm s} = -0.324$, $\omega_{\rm s} = 1.49$, $\alpha = -6.17$, $\gamma = -0.038$, $\omega_{\rm t} = 0.82$, $\xi_{\rm t}0.01$) are free parameters[1].

GALFORM has no information about the local environment of the haloes, other than the implicit differences in clustering, driven by halo assembly bias. For this reason, the concentrations it calculates are essentially median values for haloes of a given mass. Likewise, only the median concentration corrections are applied, without splitting haloes into screened and unscreened sub-catalogues. For each halo we calculate M_{500}, which is defined as the mass enclosed inside a radius within which the mean density is 500 times critical density of the universe; then, we calculate a dimensionless x from Eq (4.4), and concentration correction from Eq (4.1).

In order to implement the new concentration mass relation, it was necessary to alter the way in which GALFORM updates halo concentrations. In its present form, GALFORM does not technically compute the $c = (m, z)$ value, but rather calculates $c = c(m, z_f)$ instead – halo concentrations and velocities are only updated at formation times of the haloes. This has two effects on the $c - m$ relation output by GALFORM: firstly, all values are systematically lower than they would be if concentrations were measured from the N-body simulation; and secondly, the relation exhibits a large scatter, caused by the fact that while halo concentrations are

[1]for further details on fitting of this relation and the associated uncertainties see Mitchell et al. (2019)

set deterministically from their masses, each mass bins contains haloes with vastly different formation times.

Therefore, we have made one significant change to the algorithm used in GALFORM to update halo concentrations. Previous models only update the concentration, circular velocity and spin of the haloes at their formation times, and then propagate these values "up" the merger tree. In our updated model, halo concentrations are recalculated at each snapshot, and therefore do not propagate up the merger tree. This is close to the cooling model developed by Hou et al. (2017), and ensures that the $c(m, z)$ relation is correct at each redshift.

Virial relation for dark matter haloes

Another modification required to the GALFORM model is to change the relation between halo mass and velocity dispersion in modified gravity. In GR the circular velocity of a halo is proportional to the square root of halo mass, $v_{200} \propto \sqrt{M_{200}}$. Since concentration-mass relation for the GR case follows a power law, as shown by Gao et al. (2008), and the concentration parameter is known to be modified in $f(R)$ as given by Eq (4.1), v_{200} can be scaled as $\sqrt{c_{200}}$.

In practice, the values calculated from this method closely resemble the approach of Fontanot et al. (2013). In their model the halo catalogue is split into screened and unscreened sub-populations, depending on their velocity dispersion: $\tilde{\sigma}_{200} \approx \sqrt{(1/2)\, c^2 |\bar{f}_R\,(z)|}$; the velocities of haloes which are considered unscreened are then increased by $\sqrt{4/3} - 1 \approx 0.1547$. This value comes from the fact that in $f(R)$ the gravity enhancement is never larger that $4/3$ – the only difference between the different versions of the $f(R)$ model is the effectiveness of the screening mechanism.

The model of Fontanot et al. (2013) is equivalent to multiplying all velocities (and concentrations, if implemented using the relation from Section 4.2.2) by a predefined number if they fall into the unscreened category. While this agrees well with the increased concentrations shown in Shi et al. (2015, Fig. 4), it falls short of correctly reproducing the concentration-mass-redshift dynamics across multiple redshifts and multiple theories. As shown by Mitchell et al. (2019), dark matter halo concentrations are not simply increased above a certain threshold mass value –

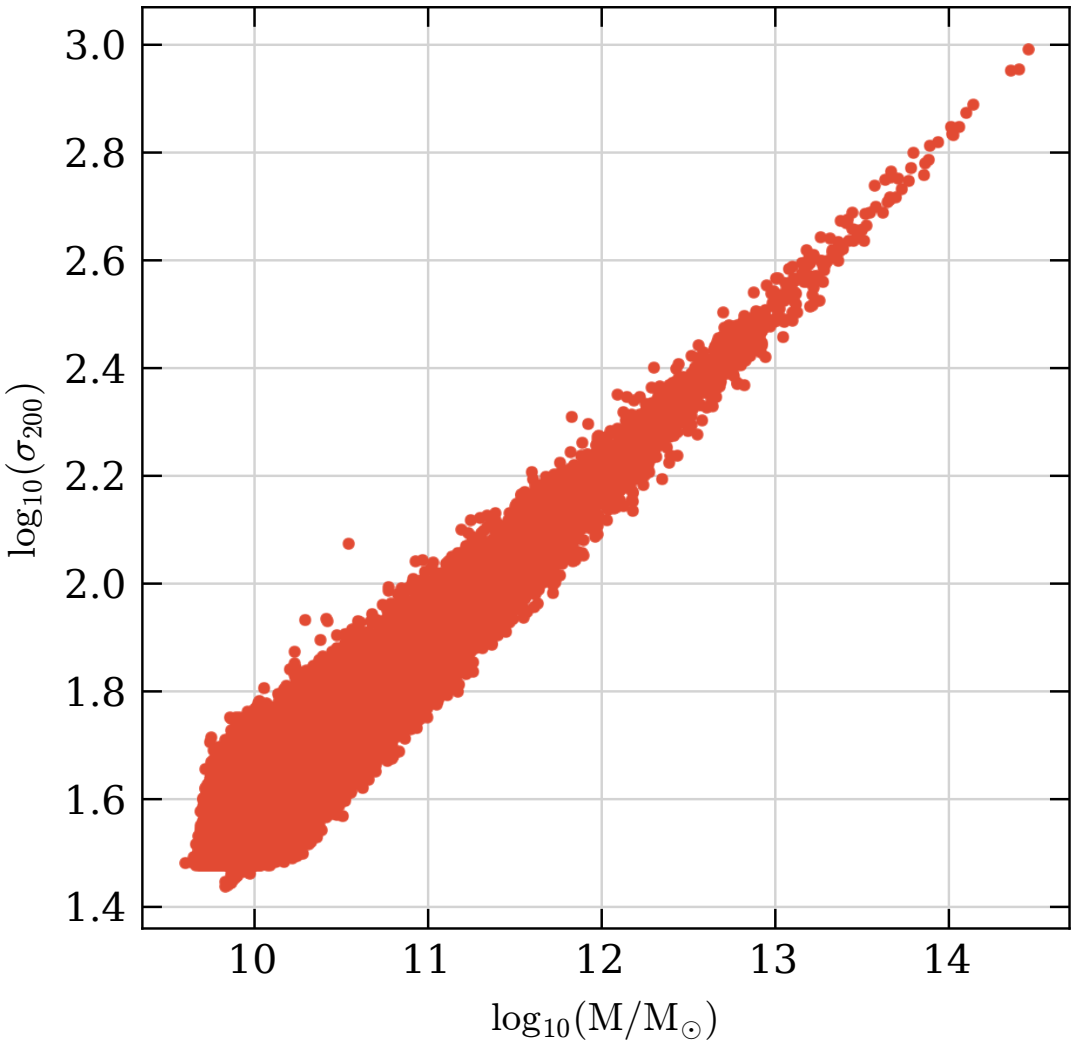

Figure 4.2: The virial scaling relation between M_{200} and σ_{200} for dark matter haloes in ΛCDM.

instead, they have a complex non-linear response which depends on both mass and redshift.

Therefore, our velocity dispersion transition is modelled just like the concentrations, with velocities multiplied by a square root of the correction factor.

Supernovae feedback parameter variation

in GALFORM supernovae (SNe) feedback is modelled as a process which ejects cold gas from the galaxy to a reservoir of mass $M_{\rm res}$, at the rate of

$$\dot{m}_{\rm disc} = \beta_{\rm disc} \times {\rm SFR}_{\rm disc}, \tag{4.5}$$

where SFR is the total star formation rate, and

$$\beta_{\rm disc} \sim \left(\frac{v_{\rm c}}{v_{\rm hot,disc}} \right)^{-\alpha_{\rm hot}}, \tag{4.6}$$

with parameters $\alpha_{\rm hot}$ specifying the dependence of β on circular velocity ($v_{\rm c}$), and $v_{\rm hot}$ normalising the circular velocity. Note that the subscript "disc" refers to the disc component, and that there is a separate gas reservoir ("burst"), with the identical feedback equations, controlling burst mode of the star formation process. These reservoirs are responsible for quenching star formation in quiescent and starburst modes respectively.

From Eq (4.6) it is evident that changes in the circular velocity driven by changes in the gravity model can be compensated for by modifying the $v_{\rm hot}$ parameters. Therefore, two additional runs of GALFORM have been performed: one with increased value of $v_{\rm hot}$ and the other with a decreased value. As discussed in this section, these parameters directly scale the impact which changes in halo velocity dispersion have on baryonic processes, and are likely to be the largest first-order effects caused by the virial and $c(m, z)$ relations. In essence, this is an attempt to understand whether it is possible to emulate the $f(R)$ fifth force by manipulating the model parameters directly.

With this question in mind, two additional runs were performed, with $v_{\rm hot}$ increased and decreased by 15%. The modified gravity interpretation of these runs is a cosmology in which the screening mechanism for the SNe feedback is turned off

for all haloes at all snapshots, whereas in the actual model these effects are both mass and redshift dependent.

4.2.3 HOD

As explained in Section 4.2.2, GALFORM uses the merger history of dark matter haloes to track the evolution of baryons, with the end-product being a prediction for the number of galaxies inside a dark matter halo along with properties for these galaxies. This allows GALFORM to predict how galaxies trace the underlying dark matter distribution. The halo occupation distribution (HOD) quantifies this in terms of the probability that a halo of mass M contains N galaxies which satisfy some constraint on a galactic property, P, which could be the broad-band magnitude of the galaxy in some filter or the luminosity of an emission line. For instance, the mean number of galaxies brighter than luminosity L measured in the Hα emission line in a halo of mass M is $\langle N\left(> L_{\mathrm{H}\alpha}\,|M\right)\rangle$ (Zheng et al., 2005; Smith et al., 2017).

While HODs can be parametrised and tuned using the observed abundance and clustering of galaxies (Zehavi et al., 2011), here we use the galaxy catalogue computed by the GALFORM model, which makes a direct *prediction* of the form of the HOD. This important as any differences introduced by changing gravity are hard to anticipate and model using a parametric form, and indeed, may not be described by the same parametrization that worked in the case of GR. We use GALFORM to calculate a range of luminosity-dependent HODs across luminosities and redshifts.

The HODs were computed for the redshift range $0.3 \lesssim z \lesssim 2.2$. This includes the ranges of the EUCLID ($1 \lesssim z \lesssim 2$) and WFIRST ($0.9 \lesssim z \lesssim 1.8$) H$\alpha$ surveys, as well as even lower redshifts since that is where the most prominent differences between $f(R)$ and ΛCDM are expected.

The Hα line luminosity is computed from the number of Lyman continuum photons, which is calculated by GALFORM for each galaxy given its star formation history and the metallicity with which stars are being produced, and a model for HII regions (Stasinska, 1990). We have applied attenuation due to dust extinction as calculated by GALFORM for the continuum stellar emission at the wavelength of the Hα line, 6563 (see Gonzalez-Perez et al. (2014) and Lacey et al. (2016) for a

description of the calculation of the dust attenuation).

At each snapshot the galaxy catalogue is split into 20 cumulative luminosity-limited samples, with the $L_{\mathrm{H}\alpha}$ limits spaced evenly in log-space from $10^{38}\,h^{-2}\,\mathrm{erg\,s^{-1}}$ to $10^{42}\,h^{-2}\,\mathrm{erg\,s^{-1}}$. The HOD is constructed for each luminosity-limited sample by computing the mean number of galaxies per halo, binned by halo mass M_{200} into 20 evenly spaced bins in log-mass from $10^{11}\,h^{-1}\,\mathrm{M}_{\odot}$ to $10^{14}\,h^{-1}\,\mathrm{M}_{\odot}$. M_{200} is defined as the mass within an over-density with an average density corresponding to 200 times the mean density of the universe.

4.3 Results

As discussed in Section 4.2.2, the modifications to GALFORM are implemented in steps to isolate the drivers of any changes in the model predictions. In this section we take one control run (using ΛCDM cosmology) calibrated for implementation in the P-Millennium N-body simulation, and compare it to the three $f(R)$ simulations at each stage: first (Section 4.3.1) by only using the merger histories from the $f(R)$ simulations; then (Section 4.3.2) by applying the $c(m, z)$ and virial relation corrections on the modified merger trees. Moreover, we present two reference runs: in Section 4.3.3, of GALFORM ran on GR merger trees, but with continuous $c(m, z)$ calculation, in order to check the effects of the continuous updates to concentration; and in Section 4.3.4, of GALFORM on GR merger trees with v_{hot} parameters increased by 15%, to emulate a global increase in a gravitational potential.

For each of these runs, we look at the model predictions for three observables: the K-band luminosity function (hereafter LF), the Tully-Fisher (hereafter TF) relation i.e. the relation between circular velocity and luminosity for disc-dominated galaxies, and the HOD (Section 4.2.3).

For a modified gravity cosmology to pass the local astrophysical constraints, it must give a good fit to the observed K-band galaxy LF at $z = 0$. We have picked this band as it most closely traces the galaxy stellar mass function, with a weak dependence on the age and metallicity of stellar populations. We compare the GALFORM calculation with the observed LF directly, instead of relying on the

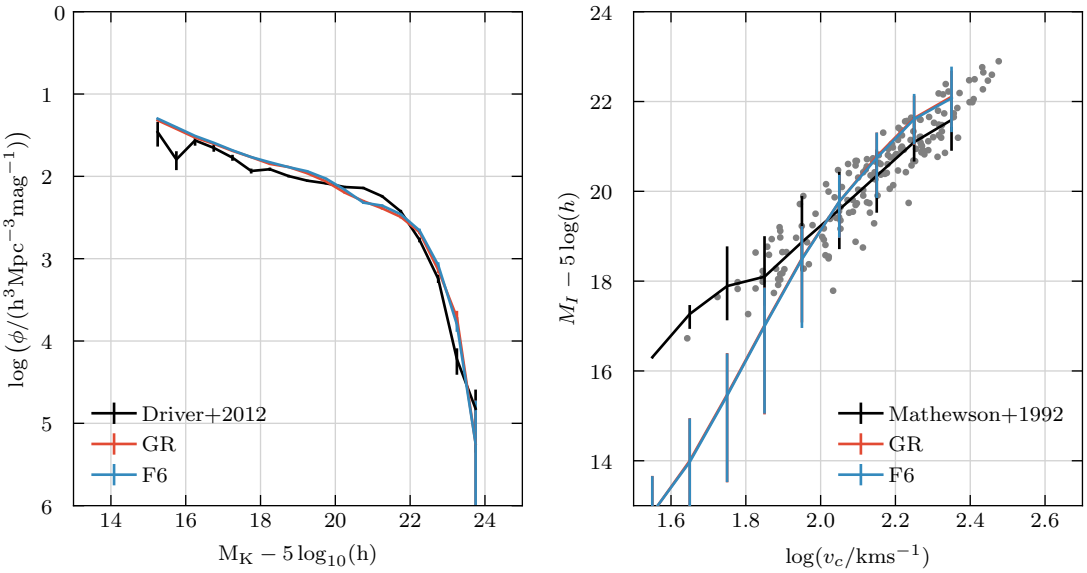

Figure 4.3: Same as Fig. 4.15, but the red line represents unmodified GALFORM model ran on F6 merger trees.

SED fitting (see e.g. Mitchell et al. (2013)). While there is ongoing research on the stellar dynamics and evolution in the modified gravity theories (Davis et al., 2012; Sakstein, 2015), the effects of the modified gravity on the IMF and stellar evolution are unknown. We have therefore retained the IMF and stellar population synthesis models used in GR and have applied these in the modified gravity implementations of GALFORM .

The TF relation at $z = 0$ is plotted in terms of the I-band magnitudes and circular velocities of spiral galaxies measured at the half mass radius of the model galaxy. The I-band magnitudes were calculated in the rest frame using the face-on values, and with attenuation by dust included. The circular velocities were calculated at the half-mass radius of the disc by integrating the mass included in the NFW profile (Cole et al., 2000). GALFORM galaxies are selected to have a B-band bulge-to-total ratio $(B/T)_B < 0.2$ (including dust attenuation), gas fractions $M_{\rm cold}/M_{\rm star} > 0.1$, to replicate the observational sample selection.

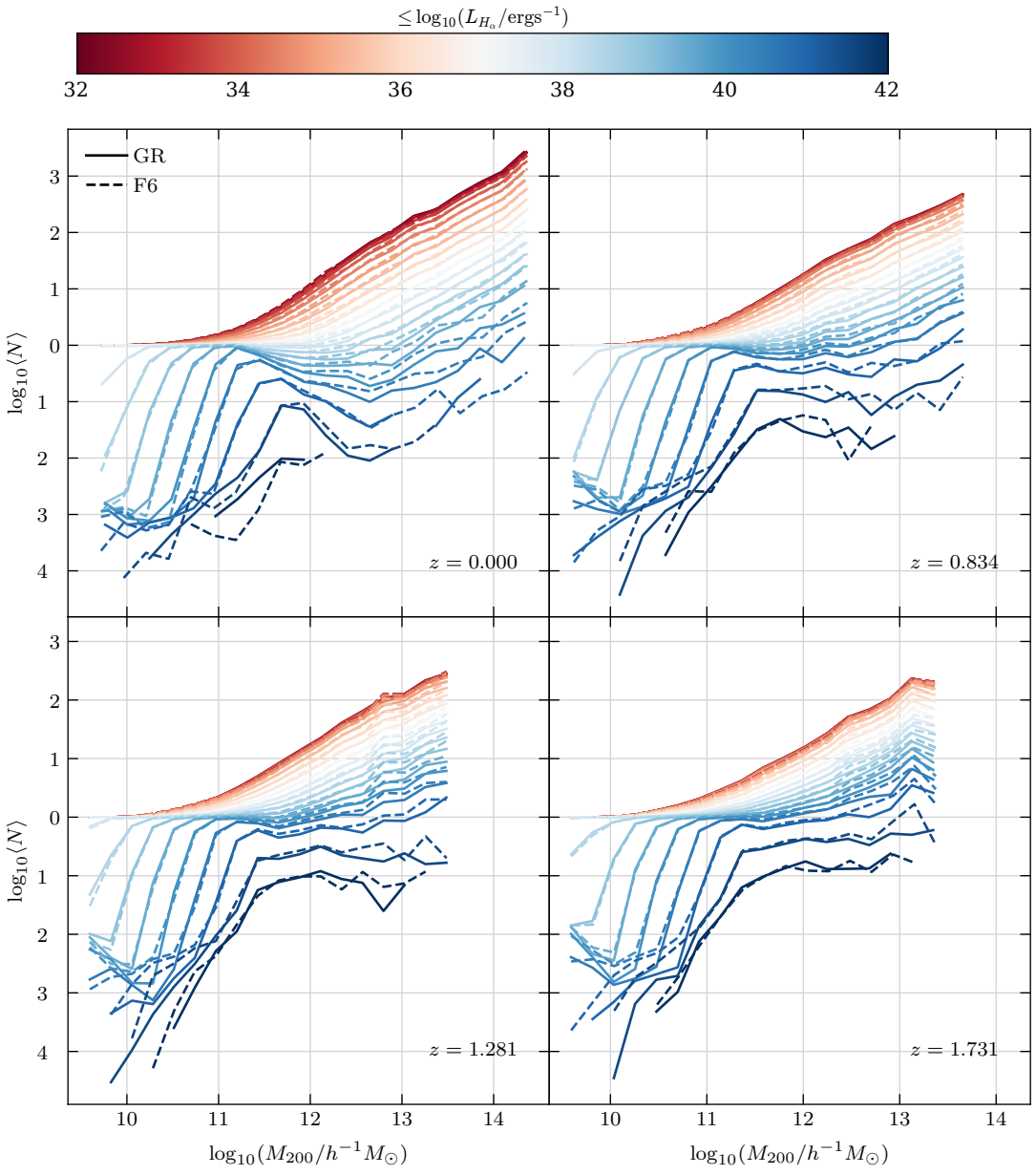

Figure 4.4: Same as Fig. 4.16, but the dashed lines correspond to HODs calculated from unchanged GALFORM model running on F6 merger trees.

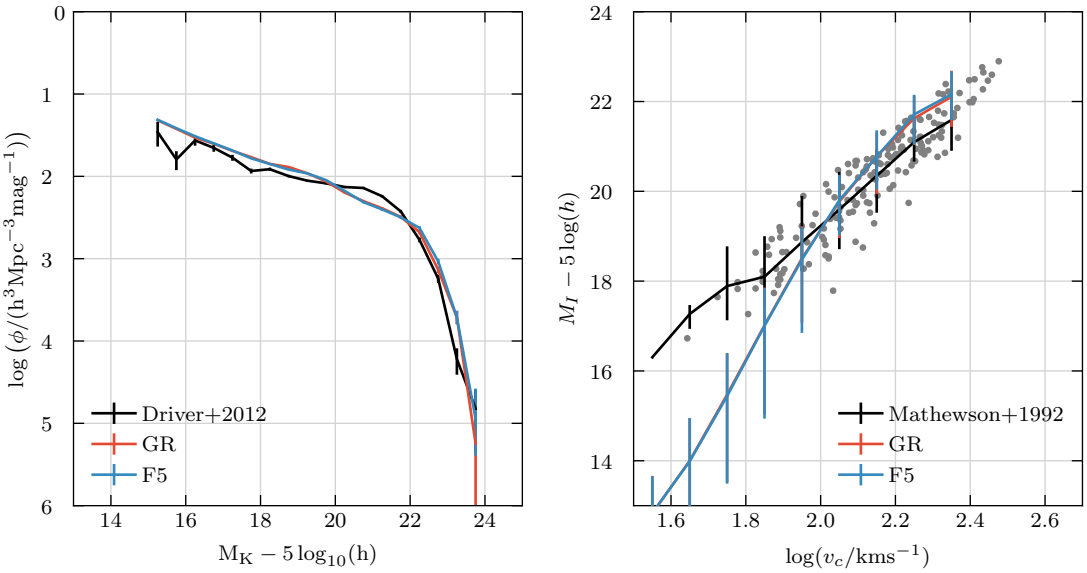

Figure 4.5: Same as Fig. 4.15, but the red line represents unmodified GALFORM model ran on F5 merger trees.

4.3.1 GALFORM on $f(R)$ Merger Trees

In the first stage, the only change to the GALFORM model is to use the dark matter halo merger trees from each simulation rather than the trees from the GR simulation. In this part, we do not observe significant changes in the GALFORM model predictions, since the differences between the ΛCDM and the $f(R)$ merger trees are small (Shi et al., 2015).

4.3.2 Concentration-mass-redshift & virial relation corrections

In the second stage, the GALFORM model used the halo merger trees from the modified gravity simulations, and corrections were applied to the halo concentrations and virial velocities according to Eq (4.1), as described in Section 4.2.2. Since the correction depends on the present redshift, ant not formation redshift (Section 4.2.2), we further modify GALFORM, to recalculate halo concentration at every timestep, instead of only at formation times (as is done in standard GALFORM).

Both corrections were applied simultaneously since concentration and velocity

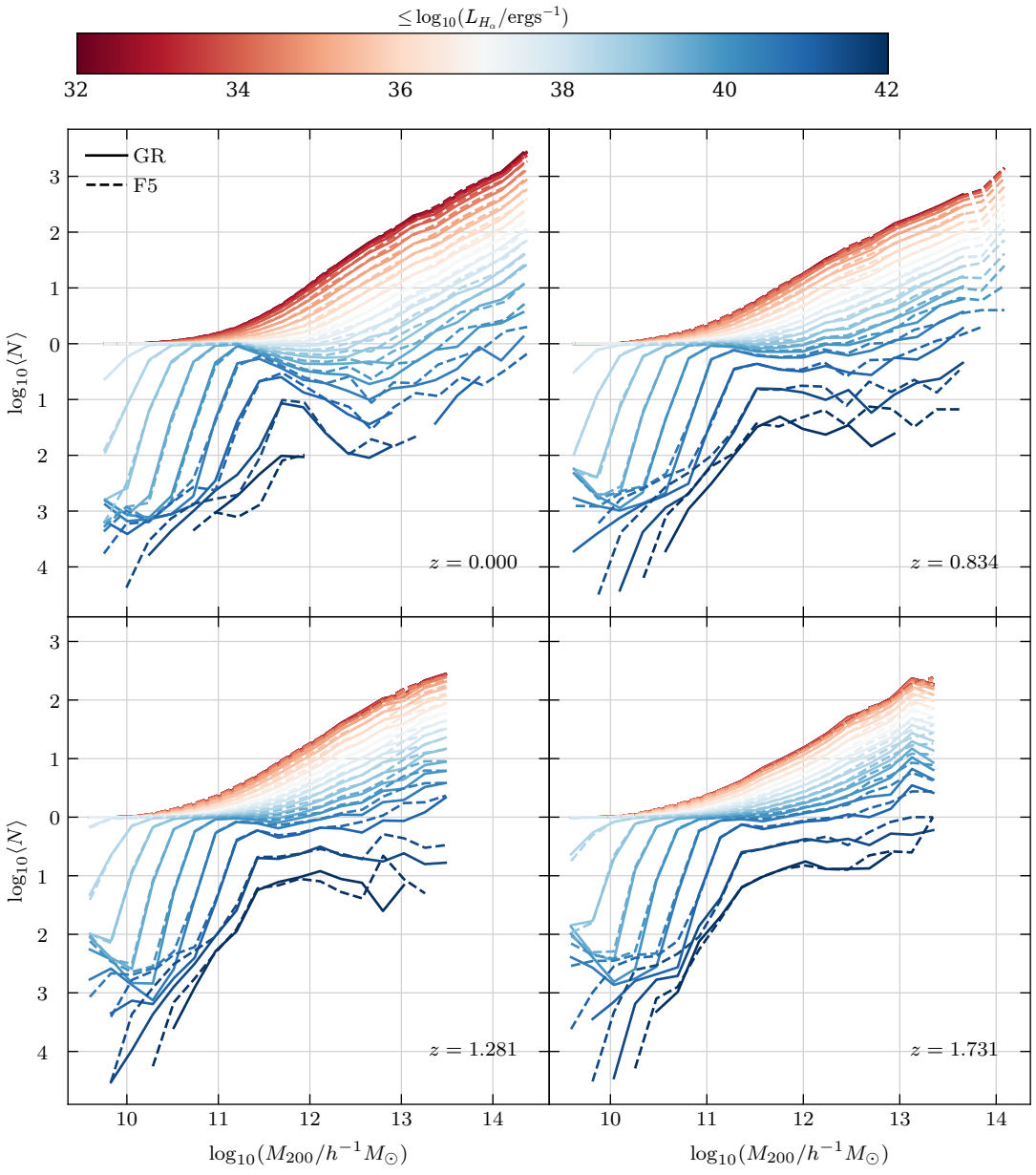

Figure 4.6: Same as Fig. 4.16, but the dashed lines correspond to HODs calculated from unchanged GALFORM model running on F5 merger trees.

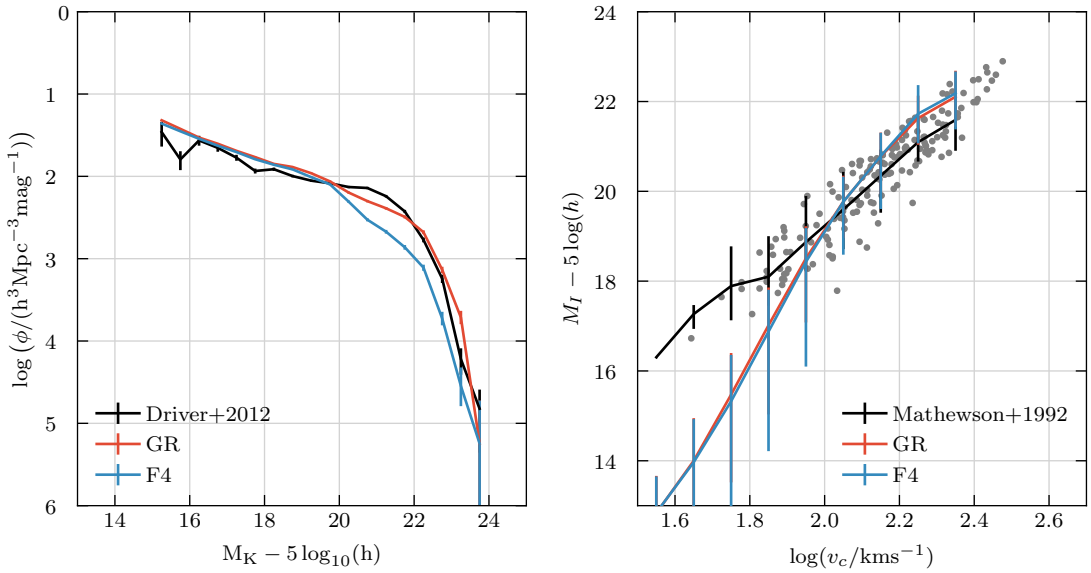

Figure 4.7: Same as Fig. 4.15, but the red line represents unmodified GALFORM model ran on F4 merger trees.

are inter-dependent, and modifying one without another would be inconsistent.

4.3.3 Continuous $c(m, z)$ Calculation

In this reference run, we modify GALFORM to recalculate halo concentration at every timestep, instead of only at formation times (as is done in standard GALFORM). Our aim is to verify to what extent are the effects observed in the previous subsection caused by continuous concentration recalculation, versus by the modified gravity correction.

4.3.4 $v_{\rm hot}$ Parameter Variation

In this reference run, the GALFORM model uses the modified gravity merger trees without changing the halo concentrations and virial relations. In this case we make a change to the GALFORM parameters controlling SNe feedback, increasing $v_{\rm hot,disc}$ and $v_{\rm hot,burst}$ parameters by 15%. As described in Section 4.2.2, since GALFORM calculates SNe feedback from Eq (4.6), we can approximate the effect that modified gravity has on the circular velocity by increasing the parameter by which

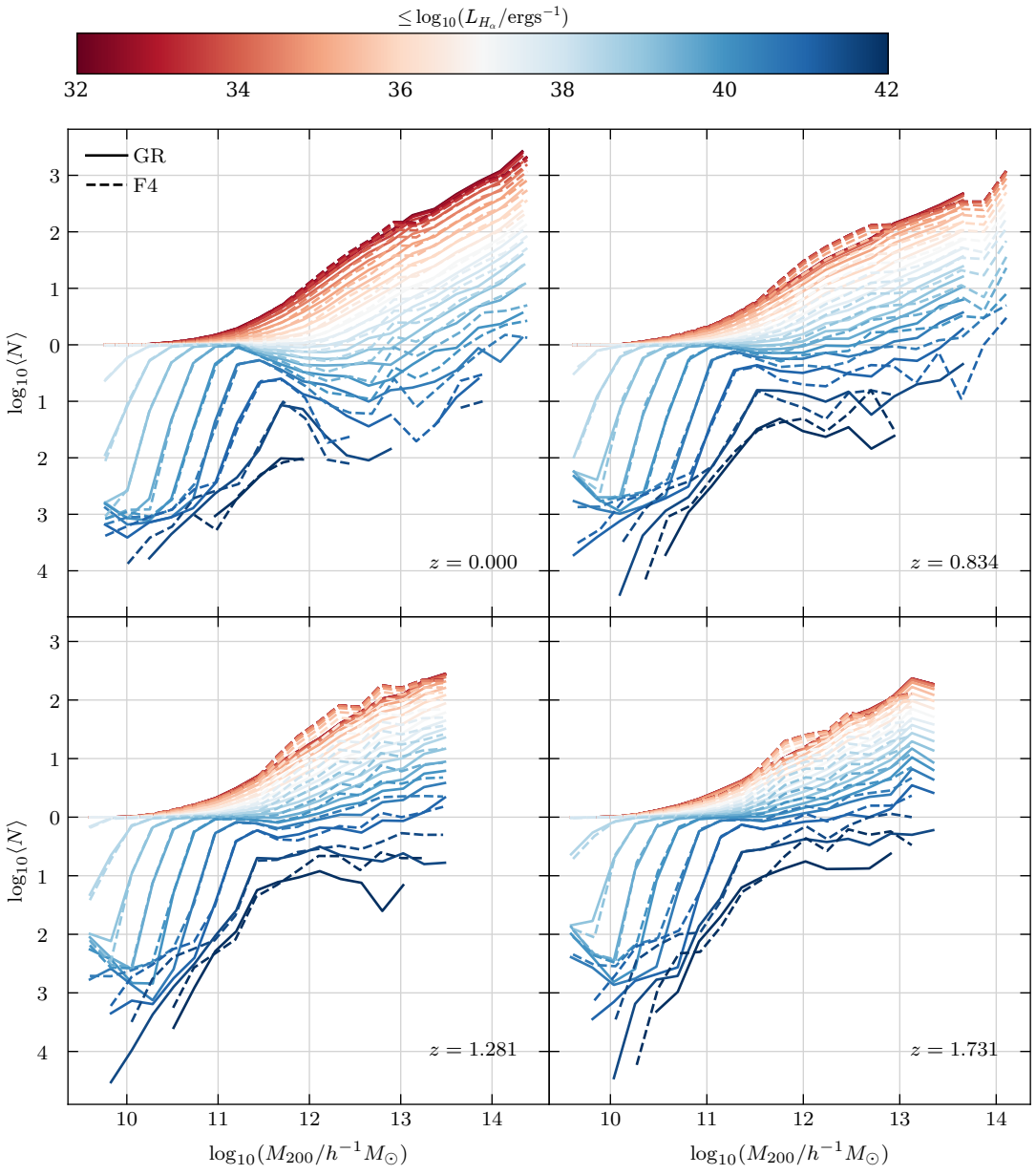

Figure 4.8: Same as Fig. 4.16, but the dashed lines correspond to HODs calculated from unchanged GALFORM model running on F4 merger trees.

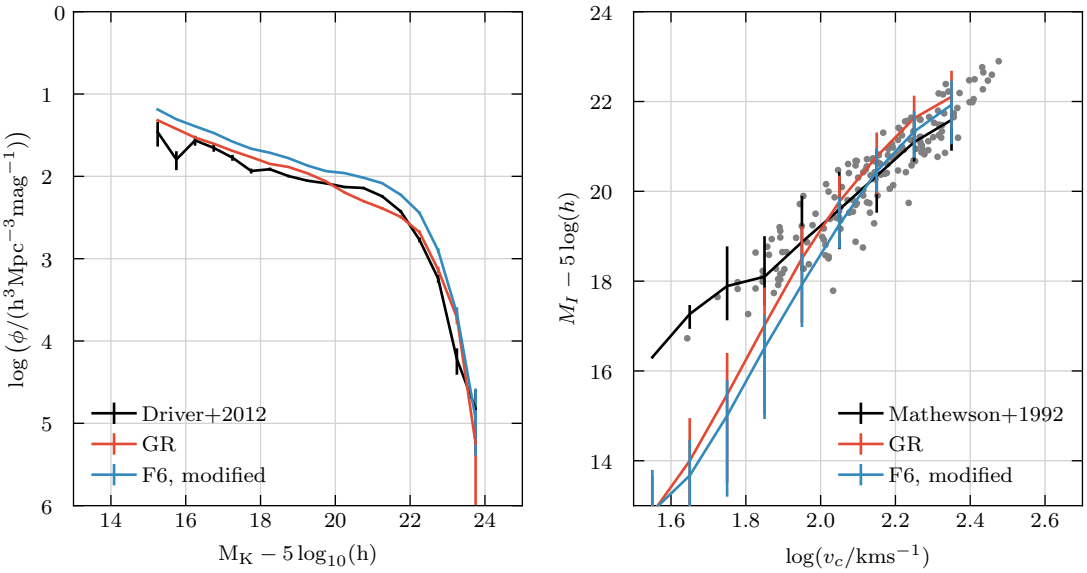

Figure 4.9: Same as Fig. 4.15, but the red line represents GALFORM model with modified $c(m, z)$ and virial relations (see Section 4.2.2), ran on F6 merger trees.

the velocity is normalised. This is equivalent to modelling SNe feedback as being unscreened for all $f(R)$ models.

4.4 Discussion

4.4.1 The K-band luminosity function and Tully-Fisher relation

As expected, the first stage (only changing the underlying merger histories of the dark matter haloes) does not have a perceptible impact on the LF or the TF relation.

In the second stage, when changing concentrations and virial velocities in addition to the merger histories, the TF relation in $f(R)$ exhibits a predictable deviation from ΛCDM – the circular velocities are systematically higher. The change is within the 10%−90% percentile range, which makes both GR and F6 models good fits to the fainter end of the TF relation. The K-band LF is also changed by the modifications to the concentration and virial relations, even in F6 which has the most effective screening mechanism and is therefore the most similar cosmology to ΛCDM . The

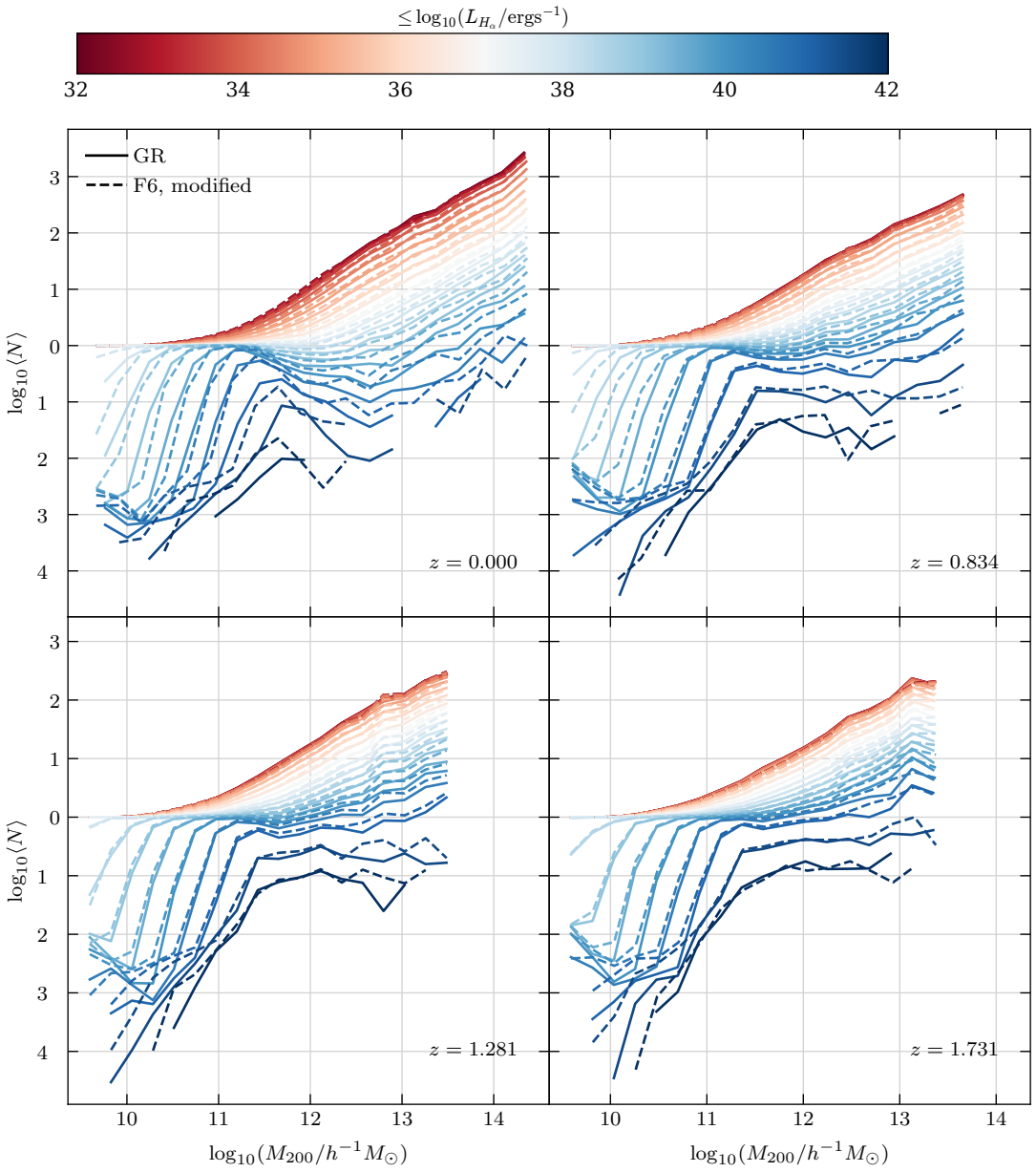

Figure 4.10: Same as Fig. 4.16, but the dashed lines correspond to HODs calculated from GALFORM model with modified $c(m, z)$ and virial relations (see Section 4.2.2), ran on F6 merger trees.

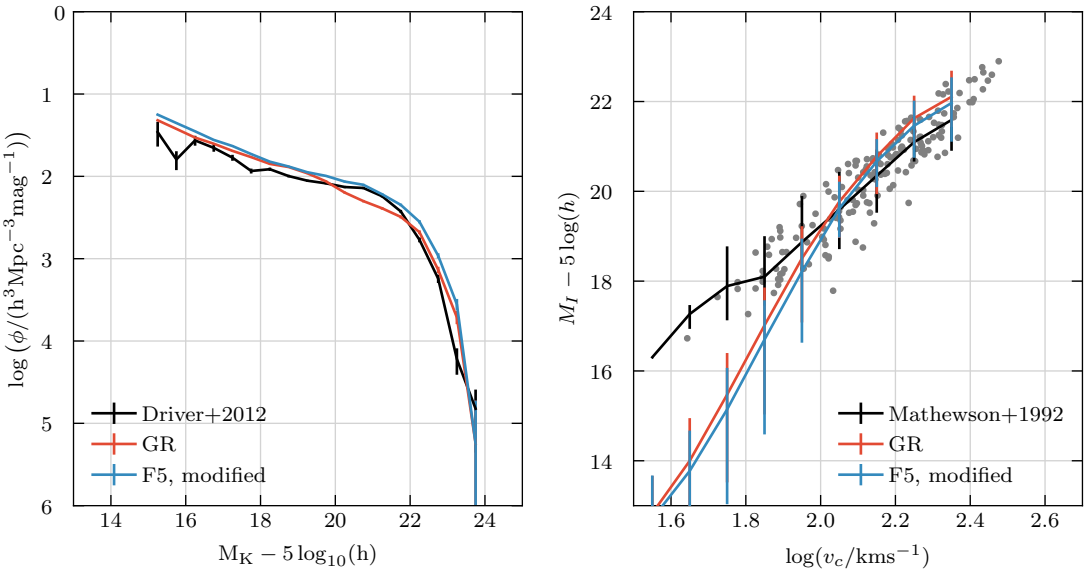

Figure 4.11: Same as Fig. 4.15, but the red line represents GALFORM model with modified $c(m, z)$ and virial relations (see Section 4.2.2), ran on F5 merger trees.

K-band LF is systematically higher in $f(R)$ than in the GR counterpart, with the most profound increase at L^*. Nonetheless, both observables exhibit similar systematic effects, which indicates that the impact of $f(R)$ on some observables can be accounted for using only a simple change to the parameter values of a model. Interestingly, LF and TF in $f(R)$ run after applying modified gravity $c(m, z)$ correction are closer to the GR version for F4 gravity than they are for F5 and F6. This is not surprising, as it can be easily explained by comparing the concentrations from Fig. 4.1 – the curves for F4 and GR and actually closer than, for instance, the curves for F6 and GR.

In order to isolate the modified gravity impact on the observables, we also include the reference run on GR merger trees, with unmodified concentration being updated at every timestep. By comparing Fig. 4.9 and Fig. 4.15, we can conclude that the modified gravity has an impact on both LF and TF relation which is distinct from merely changing the manner in which concentrations are calculated. However, for a more robust comparison, we recommend further study comparing two runs of GALFORM calibrated for continuous concentration updates: one without the $c(m, z)$ correction, and one with.

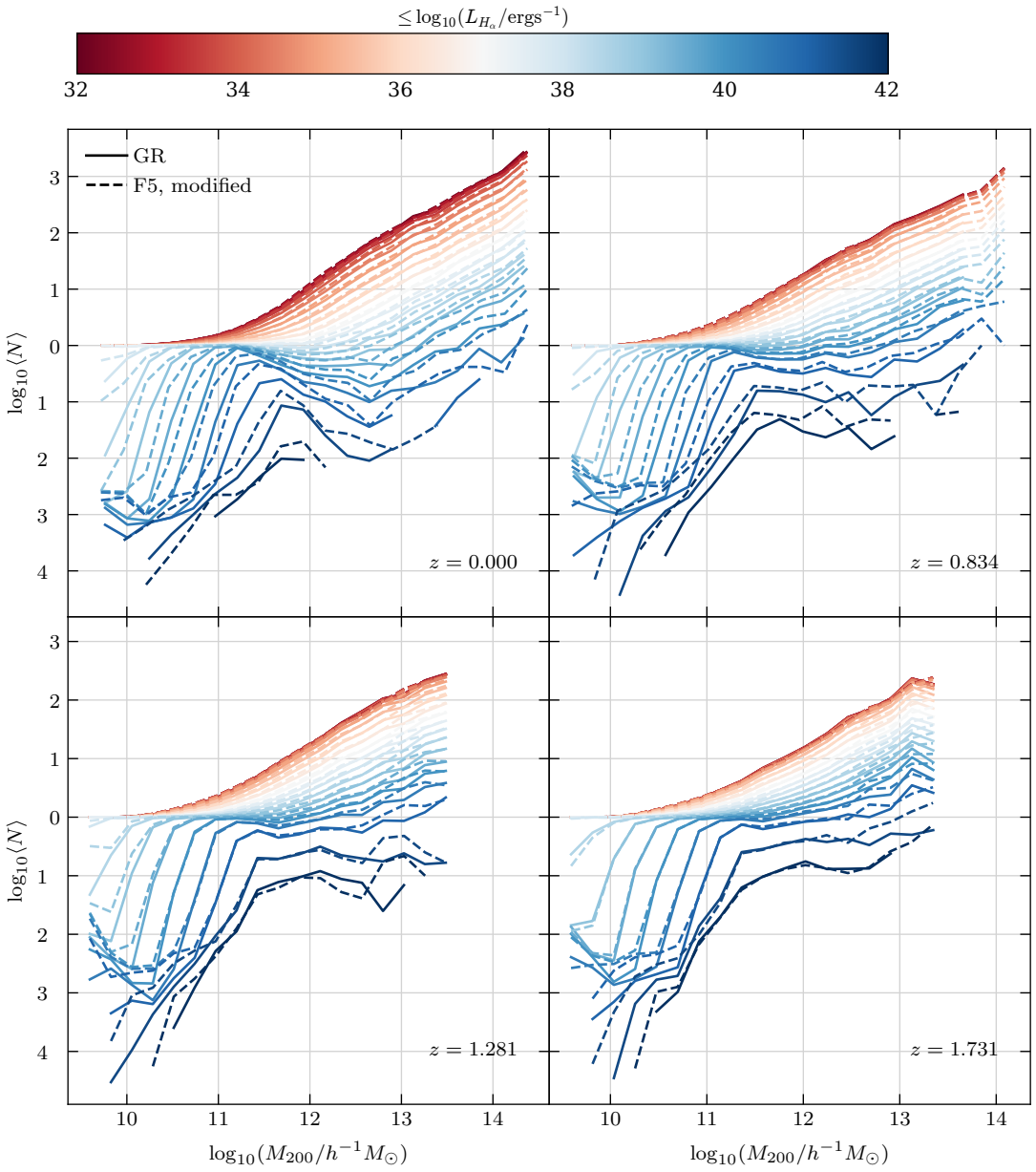

Figure 4.12: Same as Fig. 4.16, but the dashed lines correspond to HODs calculated from GALFORM model with modified $c(m, z)$ and virial relations (see Section 4.2.2), ran on F5 merger trees.

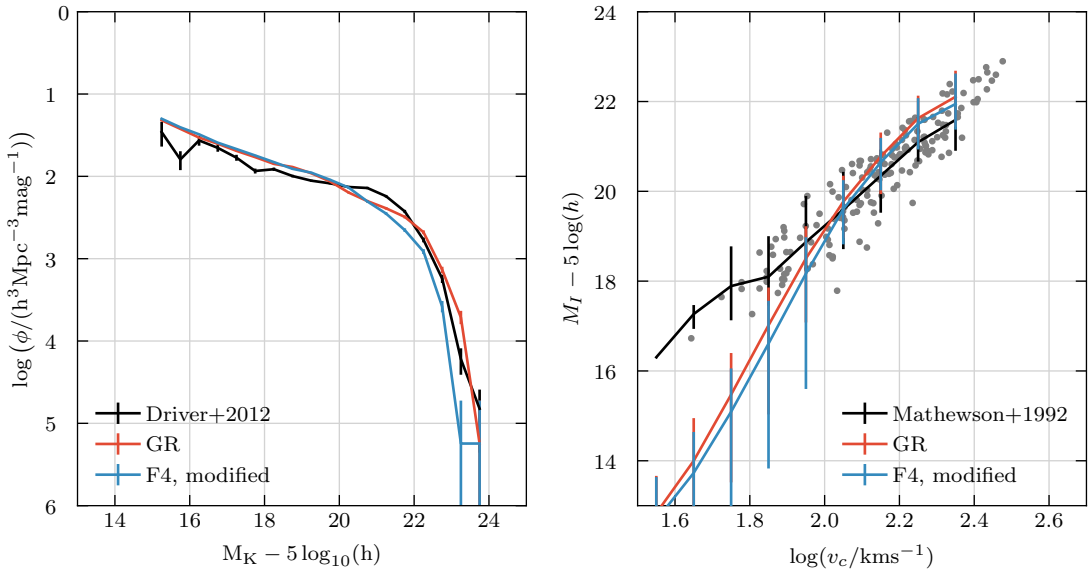

Figure 4.13: Same as Fig. 4.15, but the red line represents GALFORM model with modified $c(m, z)$ and virial relations (see Section 4.2.2), ran on F4 merger trees.

Finally, when we increase the $v_{\mathrm{hot,disc}}$ and $v_{\mathrm{hot,burst}}$ parameters by 15%, the TF relation is shifted higher on the plot, which is to be expected. Interestingly, the K-band LF is affected in the opposite way it was in the second stage – the LF is mildly suppressed at the faint end, and strongly suppressed at the bright end.

4.4.2 HODs in $f(R)$ cosmologies

Similar to the results from the previous section, there is little difference between the control run and the first stage.

In the second stage the HODs exhibit differences at low redshift for low mass haloes, and at the highest Hα luminosity cut. This coincides with the objects for which the $f(R)$ screening switches off. The HODs peaks for $f(R)$ typically coincide with the GR peaks, but the number of galaxies is systematically lower. As expected, at higher redshifts the HODs of GR and F6 have no perceptible differences, which serves as a control check that the model predictions are indeed driven by the modified gravity phenomenology.

This can be explained by either the fact that fewer haloes successfully formed galaxies, or that the galaxies in $f(R)$ have lower Hα luminosities as their SNe

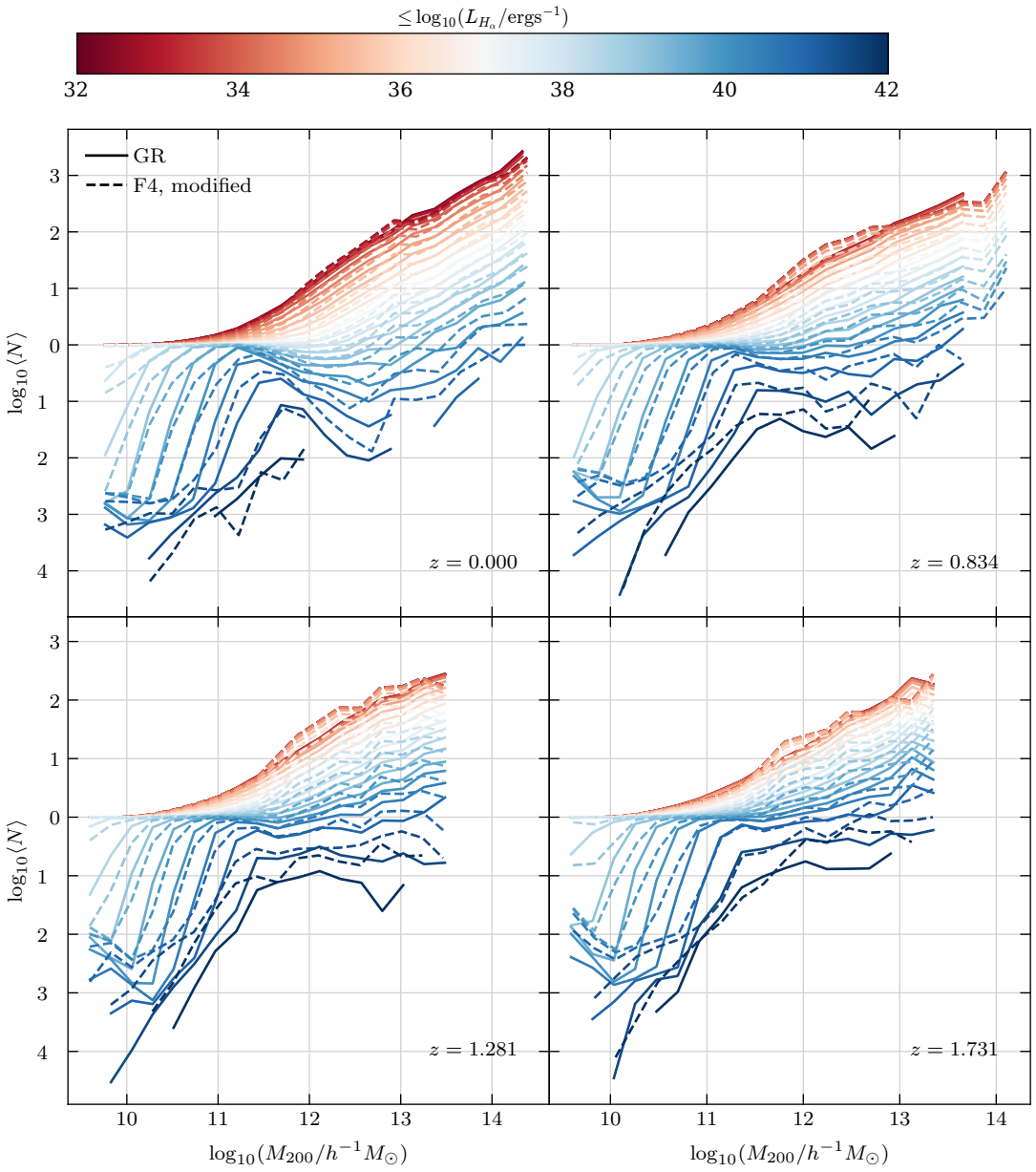

Figure 4.14: Same as Fig. 4.16, but the dashed lines correspond to HODs calculated from GALFORM model with modified $c(m, z)$ and virial relations (see Section 4.2.2), ran on F4 merger trees.

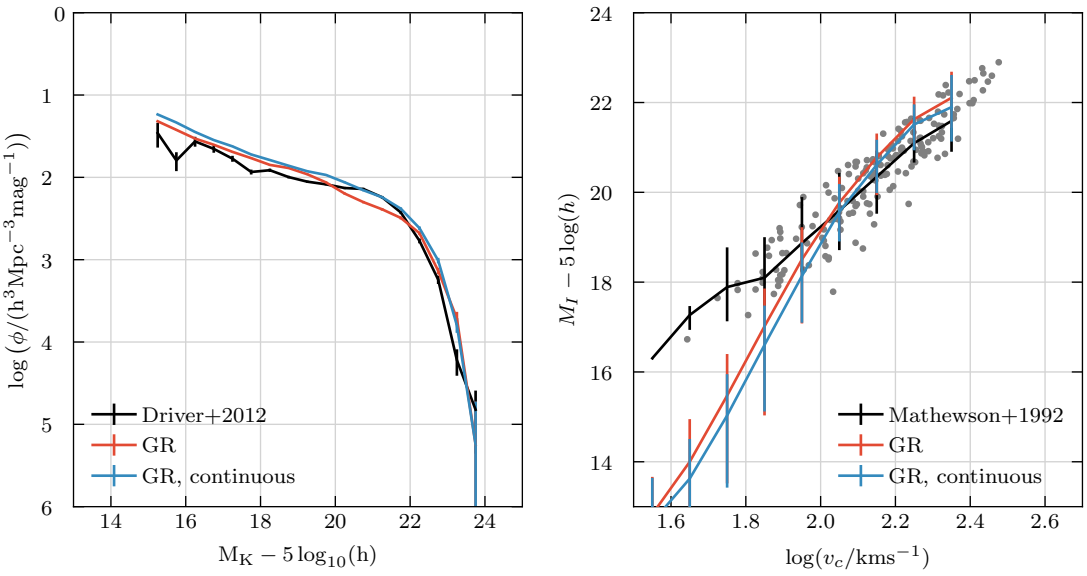

Figure 4.15: Predictions of the rest-frame K-band LF with dust extinction and TF relation, both at $z = 0$. The results from the simulation are compared to the observational data from Driver et al. (2012) (for the K-band LF), and to the data from Mathewson et al. (1992) (for the TF relation). The blue line represents the reference GR run, and the red one the GR run with GALFORM modified such that concentrations are updated continuously (see Section 4.2.2).

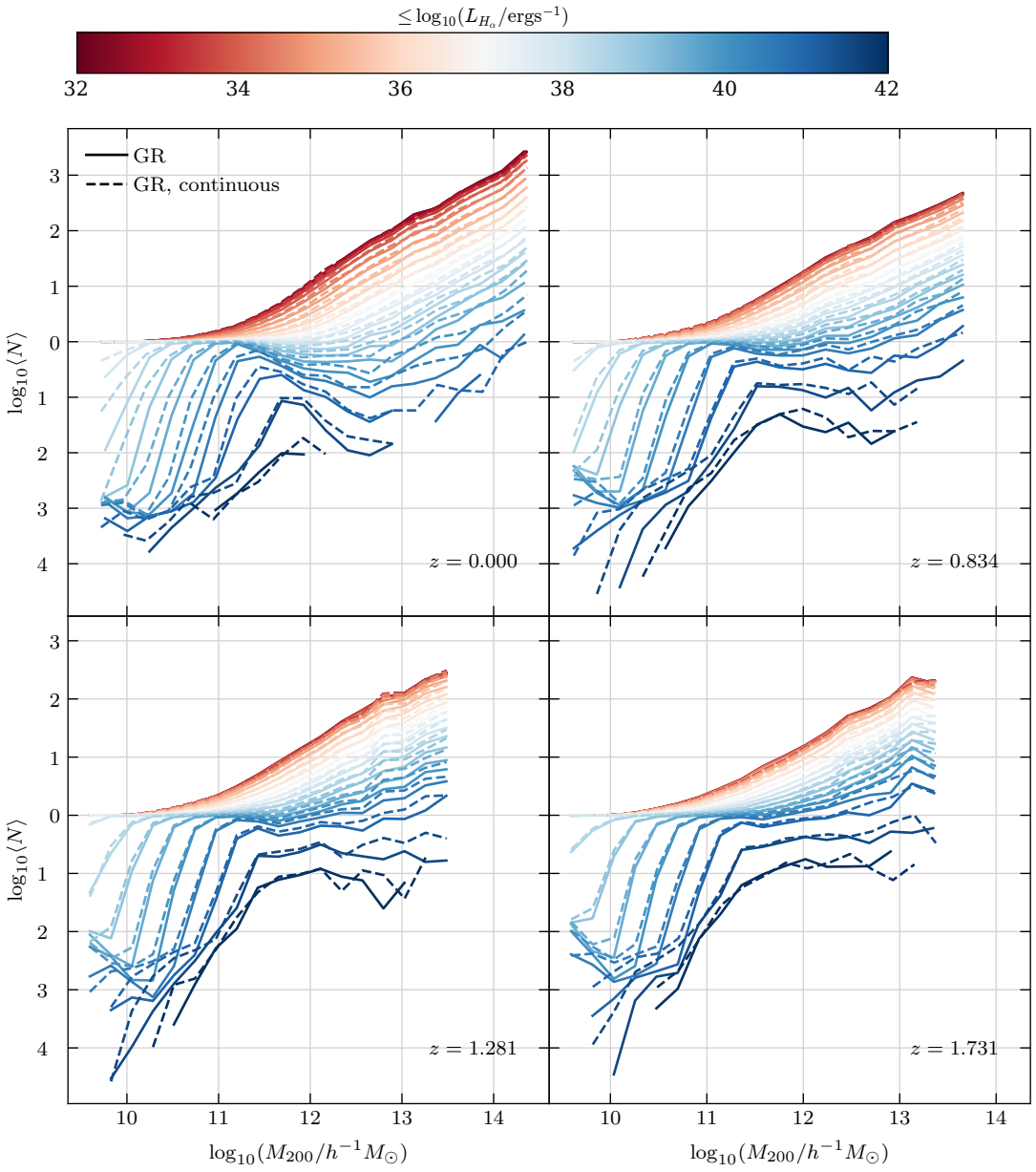

Figure 4.16: Predictions of the HODs for GR at four redshifts as labelled in each panel. The colours of the lines represent the Hα luminosity limits, as displayed by the colourbar on top; the styles of the lines correspond to standard GALFORM (solid line) and GALFORM with continuously updated $c(m, z)$ relation (dashed line; also see Section 4.2.2), as labelled in the legend in the top-left panel.

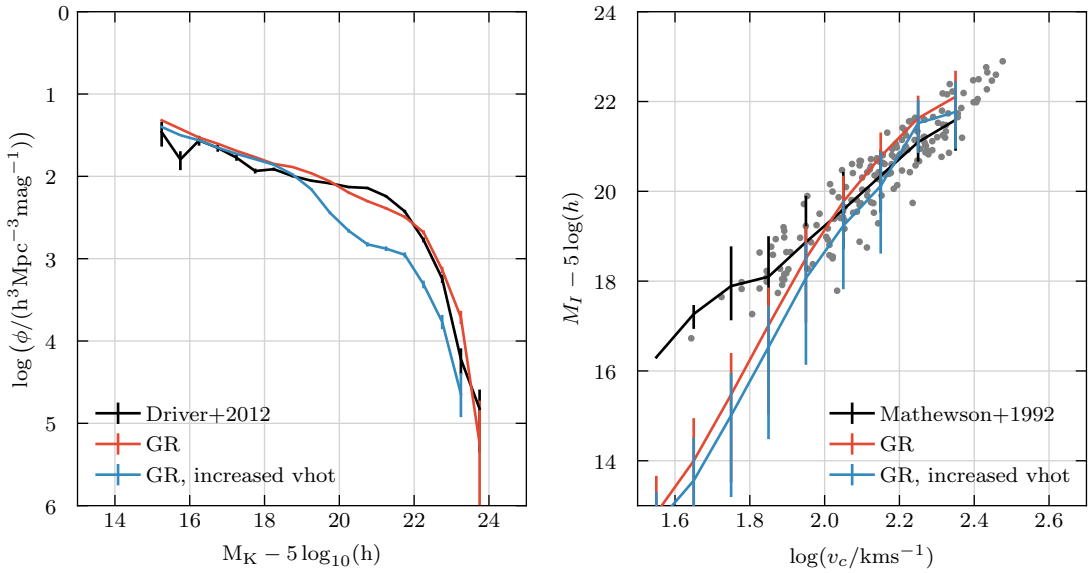

Figure 4.17: Same as Fig. 4.15, but the red line represents GALFORM model with $v$hot, burst and $v$hot, disc parameters increased by 15%.

feedback is more effective. If the latter conclusion is correct, it provides an interesting counterpoint to Davis et al. (2012), which argues that stars and galaxies in a $f(R)$ universe should be brighter than those in a ΛCDM one. A more complete model, with stellar population synthesis models with take into account modifications to gravity, would be necessary to assess the relative importance of these two opposing effects.

When the circular velocities are changed for all haloes at all redshifts, as they were in the third stage, the HODs are all impacted similarly – the number of galaxies per halo mass bins, for a given Hα luminosity cut, are lower. Therefore, it proves that the impact of $f(R)$ gravity on some observables *cannot* be modelled using only a simple changes to the model parameters.

4.5 Conclusions

We have implemented the GALFORM SAM on merger trees obtained from N-body simulations of modified gravity cosmologies and compared the predicted K-band luminosity function, Tully-Fisher relation, and Halo Occupancy Distributions. To

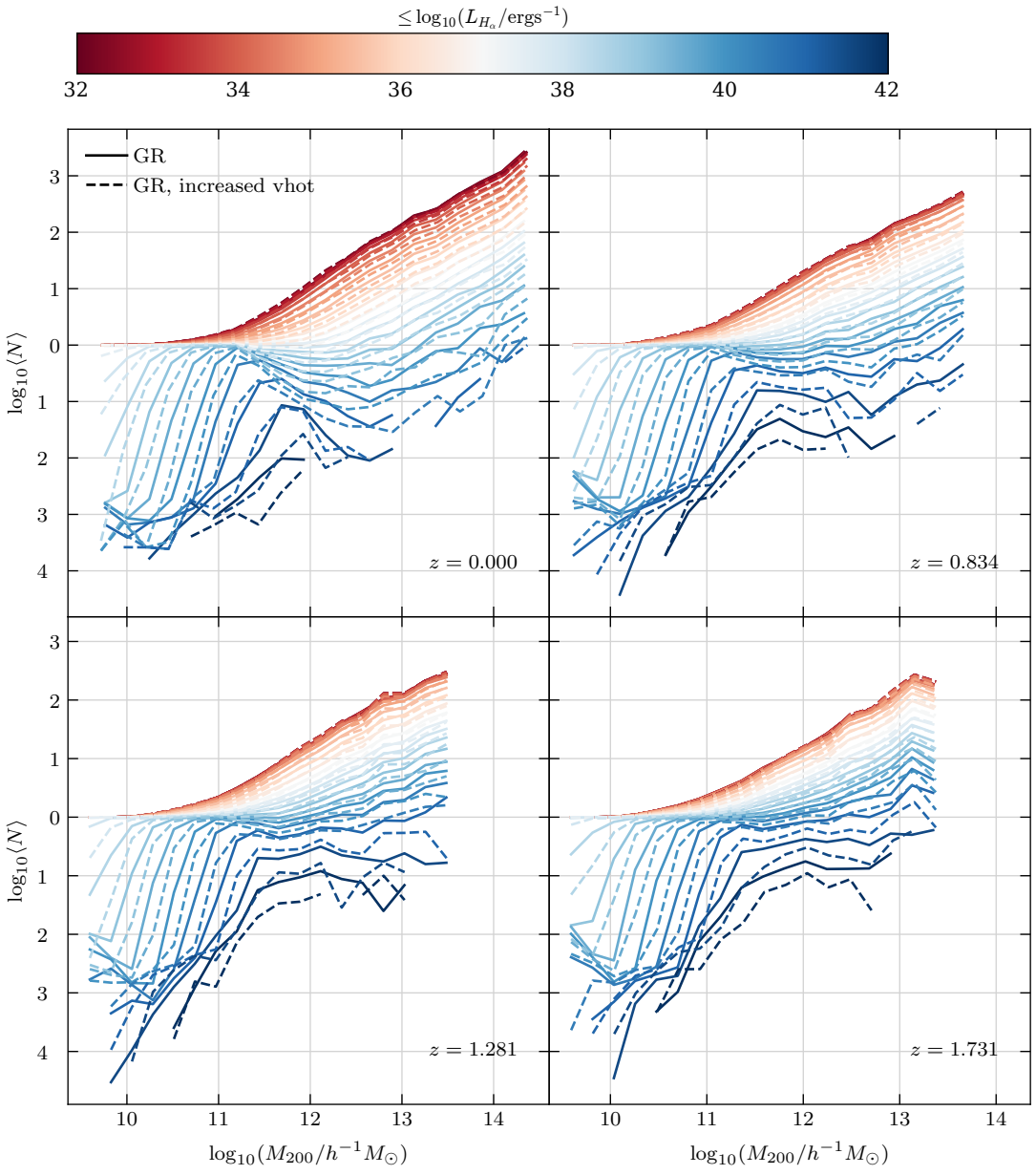

Figure 4.18: Same as Fig. 4.16, but the dashed lines correspond to HODs calculated from GALFORM model with $v_{\rm hot}$ parameters increased by 15%.

model the effects of $f(R)$ gravity we have modified the SAM in a number of ways – by running an unchanged version on different merger trees, applying an empirical correction to the $c(m, z)$ and virial relations, and increasing the $v_{\rm hot}$ parameters by 15%. Each of these stages was run separately, to carefully distinguish the effect of the modification on the results, and evaluate its relative importance.

Semi-analytic modelling of the effects of the modified gravity on the galaxy formation and evolution is still a nascent area of research, and our results are preliminary. The difficulty lies in the combinations of complexity of the galaxy formation SAMs, which contain many non-linear sub-models and implicit Newtonian assumptions, as well as the multitude of modified gravity theories and the sheer strangeness they bring to the otherwise well understood gravitational equations.

Changing the merger trees of the underlying dark matter haloes on its own has very little impact on the GALFORM outputs, as most of the phenomenology is driven by the baryonic processes. However, after also modifying the $c(m, z)$ and virial relations in GALFORM, we find small but systematic changes in K-band LF, TF relation and HODs. While $f(R)$ effects are non-linear, their impact on LF and TF relations are systematic and relatively easy to explain. Similarly, characteristic "peaks" of the $L_{\rm H\alpha}$-limited abundances in the HODs shift predictably, moving with decreased screening efficiency, decreasing redshift, decreasing mass and increasing luminosity. Each of the observables individually are straight-forward to model analytically, but the difficulty lies in designing a physically-motivated model which will produce those outputs from first principles.

Specifically, finding an observable which could serve as a modified gravity "smoking gun" on a galactic scale (and timescale), and predicting it with a SAM would be a big step forwards in the modified gravity research.

A common criticism of modified gravity theories is that they are finely tuned to evade observational evidence rather than provide falsifiable predictions or fit their models. While it is important that gravity theories are bound by observational constraints, there is another role of modified gravity modelling. Currently, due to unprecedented success of GR, SAMs are tightly coupled to the Newtonian gravity equations. This leads to a situation in which it becomes difficult to test gravity the-

ories with predictions from SAMs, and in which all SAM predictions are dependent on one dominant theory. It will therefore benefit SAMs if gravity assumptions they make could be made explicit, and modular, in the same way in which dark matter halo density profiles are.

Finally, we would like to point at two more difficulties in carrying out modified gravity research on SAMs – selecting model outputs and measuring relative importance of SAM components and modified gravity corrections. Even observables whose response to modified gravity models is qualitatively well-understood, such as HODs, are difficult to discretise to a form in which we can directly compare different models. On the other hand, there is a well-established technique for measuring relative importance of model parameters on model outputs, called Sensitivity Analysis, which is discussed in more detail in Chapter 5.

Chapter 5

Sensitivity Analysis of Galaxy Formation Models

We perform a multi-parameter exploration of the GALFORM semi-analytic galaxy formation model, to compute how sensitive the present-day K-band luminosity function is to independently varying different model parameters using variance-based sensitivity analysis (SA). We demonstrate the usefulness of the SA approach by varying just two model parameters, one which controls supernova feedback and the other the heating of gas by AGN, which matches our physical intuition regarding how these parameters affect the predictions for different parts of the galaxy luminosity function. Subsequently, we use SA to compute Sobol' sensitivity indices varying seven model parameters simultaneously, connecting the variance in the model output to the variance in the input parameters. We discover that the SA correctly identifies the least and most important parameters, and that it is able to capture the combined responses of varying multiple parameters at the same time. Our study marks a much needed step away from a traditional, "one-at-a-time" parameter variation, and improves the transparency of multi-parameter models. It is also the first application of a variance-based SA to a model that aims to predict the evolution and properties of the whole galaxy population.

5.1 Introduction

Galaxy formation is a complex process which we are only now just starting to understand through a combination of observations, numerical simulations and analytical modelling. Two main theoretical techniques are used to model the formation and evolution of galaxies: semi-analytical modelling (SAM) and hydrodynamic simulations (for a review see Somerville & Davé (2015)). SAMs use physically motivated, simplified mathematical relations to describe the evolution of baryons in growing dark matter haloes (Baugh, 2006; Benson, 2010). Hydrodynamic simulations, on the other hand, tend to make fewer assumptions and approximations than SAMs and solve the fluid equations governing the dynamics of baryons. Nevertheless, in hydrodynamic simulations many processes, such as star formation, remain "sub-grid" due to the finite numerical resolution of the simulation and our inability to write down the precise equations describing some processes (Crain et al., 2015; Ludlow et al., 2019). In the absence of a complete mathematical description, physical processes are described in both SAMs and hydrodynamic simulations by approximate equations that contain parameters. Values have to be chosen for these parameters to specify a model. Here, we present a new application of an established statistical method to assess the impact of changes in model parameters on the output of a model.

The past few years have seen tremendous breakthroughs in the hydrodynamic simulation of galaxy formation for significant galaxy populations in cosmological volumes (Vogelsberger et al., 2014; Schaye et al., 2014; Pillepich et al., 2018). Nevertheless, SAMs remain an attractive and valuable complement to hydrodynamical simulations due to their flexibility and speed. These properties of SAMs mean that they can be used to build intuition about physical processes, by running thorough investigations of the impact of varying model parameters (e.g. see the comprehensive exploration of perturbations around the fiducial model presented by Lacey et al. (2016)). Also, SAMs remain the method of choice to populate large volume N-body simulations using a physical galaxy formation model: the fiducial simulation volumes used in SAMs are around 100 times bigger than those used in the current state-of-the-art hydrodynamical simulations. The predictions of SAMs have reached

an impressive level of maturity through careful comparisons between the predictions of different groups and techniques (e.g. Contreras et al. (2013); Knebe et al. (2015); Guo et al. (2016); Mitchell et al. (2018)).

Nevertheless, some scepticism remains regarding SAMs, much of which can be traced to the way in which the model parameters are set. Traditionally models have been calibrated by developing physical intuition about how the model responds to changes in selected parameter values, such as those which control the mass loading of winds driven by supernovae, and then varying one parameter at a time to hone in on a best-fitting model. Often the quality of the model reproduction of the calibration data is judged by eye and compromises are made in order to match multiple datasets; these steps are hard to quantify and therefore difficult to reproduce. The "best-fitting" model is reported as a single choice of parameter set that defines the model. The primary motivation for producing a single model is the desire to build mock catalogues for galaxy surveys (Baugh, 2013). However, users often want to know the uncertainty on the model predictions and how the predictions respond to changes in the input parameters.

The range of processes modelled by SAMs lends them the flexibility to predict varied observation but at the cost of having to specify a number of parameters which complicate model optimisation or calibration. A number of techniques have been devised to reduce the complexity or dimensionality of the parameter space and to perform efficient searchs of the parameter space: principal component analysis (Benson & Bower, 2010, hereafter PCA), Bayesian emulators (Bower et al., 2010; Gómez et al., 2012), particle swarm optimiser (Ruiz et al., 2015, hereafter PSO), Markov Chain Monte Carlo (Henriques et al., 2009; Lu et al., 2011, 2012; **?**; Mutch et al., 2013; Martindale et al., 2017), and Latin-hypercube sampling (Bower et al., 2010; Rodrigues et al., 2017).

Here we apply sensitivity analysis to quantify the dependence of the model output on the variation in the values of the model input parameters. The analysis of Gómez et al. (2014) using the ChemTreeN SAM of Tumlinson (2009) is similar in scope to our work. They use an analysis of variance technique for variance decomposition instead of sensitivity indices, and Gaussian processes for model fitting.

Here we use the GALFORM SAM effectively as a black-box model, and evaluate the sensitivity of the model outputs to the variation of the input parameters. A SAM is an ideal candidate for sensitivity analysis, as the interactions between parameters are complex enough to develop a black-box-like behaviour ("becomes easier to experiment with than to understand", Golovin et al. (2017)); however, many parameters have a natural physical interpretation, and hence it will be straightforward to develop intuition about how sensitive the model outputs should be to changing the inputs. Many parameters also have either physically motivated bounds, or at least a plausible range of possible values.

A criticism often aimed at SAMs is that they contain too many free parameters. This is usually rebuffed with the insistence that the parameters are physical, not statistical. Model fitting alone is therefore insufficient for interpreting how well a SAM is performing. A different research question, one this study tries to address, is how sensitive the model is to the parameter variation – in other words, how well do we understand the impact of the physical processes and their interactions on the model predictions?

Sensitivity analysis (SA) (Fisher, 1918; Sobol', 1993, 2001; Saltelli et al., 2010) is an area of statistical modelling which analyses how the variance of the output of a model is affected by variance in the model inputs. It is closely related to uncertainty analysis and model optimisation, and can be used to test the robustness of the model predictions to uncertainty in the input parameters, quantify dependence of the outputs of a model on different parameters, identify model non-linearities, and guide subsequent model optimisation. This addresses a common criticism of black-box models, namely that after adding sufficiently many free parameters they can be fine tuned to match any observations, and provide a single set of predictions. While model optimisation can be used to compute confidence intervals, SA is uniquely positioned to quantify model responses and the relative importance of the inputs. This addresses the complaint about SAMs listed above, that providing a spread of model predictions is preferable to fitting to the observations. Using SA, we will be able to not only tell how much model predictions vary for individual outputs, but also quantify how much of this variance can be attributed to individual model inputs

(or their combinations).

There are several SA techniques, not all of which are suitable for analysing non-linear models with a high-dimensional parameter space. With a few exceptions[1], SA is done in 3 stages:

1. sampling of the parameter space
2. model evaluation in the parameter space
3. computation of sensitivity indices

Here, we use a variance-based SA which adopts the improvement of introduced by Saltelli et al. (2019) over the Sobol' indices. Variance-based methods aim to decompose the variance of the model output into the contributions from individual parameter variances, as well as the combined variances of the interactions of multiple combinations of parameters changing at once. In order to avoid a computational penalty for evaluating all possible parameter combinations, input parameters are treated as probability distributions, and the sensitivity of the model output is estimated approximately. Moreover, a number of numerical optimisations have been introduced into the sampling and index calculation techniques, to improve the convergence of the indices and average over the values which are too difficult to compute efficiently.

This work diverges from previous studies in two important ways: firstly, we narrow the scope of this investigation to computing only sensitivity indices, and we do not attempt to provide the best-fitting values for a galaxy formation model. We believe that SA is not the best tool for this task, as it investigates model responses at the extreme values of input parameters, and often for unusual combinations of inputs, where the model no longer reproduces the observable values. Secondly, we do not limit ourselves to measuring responses of the model to individual parameters and their linear combinations. Instead, we use sensitivity indices to capture both individual and combined impacts of parameters. Lastly, this study focuses

[1]Some methods, such as Gaussian processes, use parameter exploration to simultaneously measure model sensitivity and maximise goodness-of-fit for model output(s).

exclusively on one observable, the K-band luminosity function, calculated using the GALFORM SAM, and probes how this specific model reacts to changes in the input parameters. Our scope is narrower, but also deeper than any previous study in this area.

The layout of this Chapter is as follows. In Section 5.2 we set out the theoretical background, introducing the GALFORM model and, for completeness, giving the equations for the processes that we vary (Section 5.2.1). We then discuss variance based sensitivity analysis (Section 5.2.2), the concept of low-discrepancy sampling (Section 5.2.3), the exploration of parameter space using Saltelli sampling (Section 5.2.4), define the sensitivity indices (Section 5.2.5) and illustrate these ideas with a toy model (Section 5.2.6). Our results using GALFORM are presented in Section 5.3 and our conclusions are given in Section 5.4.

5.2 Theoretical background

Here we set out the theoretical ideas used in this Chapter. Section 5.2.1 gives a brief overview of the GALFORM semi-analytical model, introducing the processes that are varied in the sensitivity analysis. Section 5.2.2 introduces variance based sensitivity analysis, Section 5.2.3 discusses the sampling of a model parameter space and Section 5.2.4 covers Saltelli sampling. Section 5.2.5 defines the sensitivity indices and Section 5.2.6 illustrates their use with a toy model. Section 5.2.7 discusses the use of GALFORM output in the sensitivity analysis.

5.2.1 GALFORM

As explained in Section 2.2, GALFORM is a SAM which aims to predict the properties of galaxies starting from dark matter halo merger histories that are typically extracted from an N-body simulation (Cole et al., 2000; Baugh, 2006; Bower et al., 2006; Lacey et al., 2016). GALFORM models the processes which shape the galaxy population using a set of physically motivated, non-linear differential equations which track the exchange of mass, energy and angular momentum between the different components of a galaxy.

Table 5.1: Planck Collaboration et al. (2014) cosmology used in the P-Millennium simulation; the last two rows give the simulation box length and the number of particles used.

parameter	value
Ω_Λ	0.693
Ω_M	0.307
$\Omega_{\rm baryon}$	0.04825
h	0.6777
σ_8	0.8288
n	0.967
$L[h^{-1}{\rm Mpc}]$	542.16
$N_{\rm P}$	5040^3

All the processes are modelled by equations that contain parameters, and a GALFORM model corresponds to a set of parameters whose values have been chosen so that the model reproduces a particular set of observations. Some of these parameters govern different choices for processes in the model, such as the radial density profile assumed for the hot gas within a halo or the stellar initial mass function (IMF) which describes the number of stars of different masses produced in episodes of star formation. For example, the Gonzalez-Perez et al. (2014) model assumes a universal, solar neighbourhood IMF whereas the Lacey et al. (2016) model invokes a top-heavy IMF in bursts of star formation and a solar neighbourhood IMF in quiescent star formation. Even though these two models are implemented in the same N-body simulation, the choices made regarding the IMF and the slightly different emphasis on which observations the model should reproduce most closely means that there are several differences in the values of the parameters which define these galaxy formation models.

Here we use the recalibration of the Gonzalez-Perez et al. (2014) model introduced by Baugh et al. (2019) for the Planck Millennium N-body simulation, which we refer to as GP14.PMILL. The Planck Millennium N-body simulation (hereafter the PMILL simulation) adopts the Planck cosmology (Planck Collaboration et al.

Table 5.2: The GALFORM parameter ranges used in this work. The parameter ranges have been taken from previous analyses (Bower et al., 2010; Rodrigues et al., 2017).

process	parameter	min	max
star formation	$\nu_{\rm SF}$ [Gyr^{-1}]	0.2	1.2
supernova feedback	$\gamma_{\rm SN}$	1.0	4.0
	$\alpha_{\rm ret}$	0.2	1.2
	$V_{\rm hot,disc}$ [km/ sec]	100	550
	$V_{\rm hot,burst}$ [km/ sec]	100	550
AGN feedback	$\alpha_{\rm cool}$	0.2	1.2
disc instabilities	$f_{\rm stab}$	0.61	1.1

(2014); see Table 5.1) and has superior mass resolution and halo merger histories that are better sampled in time compared with earlier N-body simulations into which GALFORM was implemented (see Table 5.1). Below we review the processes that we vary in the sensitivity analysis. A more complete description of GALFORM can be found in Section 2.2 and Lacey et al. (2016).

Parameter selection

We consider the relative importance of the processes described in Section 5.2.1 by performing a SA on the parameters that describe these phenomena. The parameters and the ranges over which they are varied are listed in Table 5.2. In some instances, the parameter range is reasonably well defined, such as $f_{\rm stab}$, as discussed in Section 2.2.7. In other cases, the choice of range of parameter values is less well defined. For example, using simple conservation arguments, $\gamma_{\rm SN}$ could take on values of 1 and 2 in the momentum and energy conserving phases of the wind evolution (Ostriker & McKee, 1988; Lagos et al., 2013). Numerical simulations of winds have suggested different values of $\gamma_{\rm SN}$. The other parameters defining the GALFORM model beyond those listed in Table 5.2are held fixed.

Model output

After the formation and evolution of galaxies is calculated over the merger history of the dark matter haloes in the PMILL simulation, galaxy luminosities can be obtained from the predicted star formation rate and metallicity of the stars produced using a stellar population synthesis model. Dust extinction is calculated in post-processing, based on the size and gas metallicity of each galaxy (Gonzalez-Perez et al., 2014; Lacey et al., 2016). The model output that we focus on here is the K-band luminosity function at $z = 0$.

5.2.2 Variance-based sensitivity analysis

The SA method we use here closely follows those used by Sobol' (2001) and Saltelli et al. (2019), which are designed to decompose variance in the model output into the variances of the input parameters and their interactions using as few model evaluations as possible.

Many SA approaches suffer from a number of shortcomings which make them unsuitable for analysing non-linear models. By non-linear models we mean here ones that are characterised by interactions between the inputs[2] and which therefore cannot be analysed effectively using regression or one-at-a-time (OAT) parameter variation techniques (Morris, 1991).

Unlike other methods, variance-based SA allows a full exploration of the input space, and therefore accounts for the interactions between parameters and non-linear responses of the model. It follows that variance-based methods are able to evaluate the total effect indices (see below) and rank the parameters in order of their influence on the output (Chan et al., 1997; Sobol', 2001; Saltelli et al., 2010).

Finally, we note that all SA methods assume that the model inputs are independent, which might not hold in general for complex models. For instance, correlations between inputs, or unphysical combinations of their values, cannot be recognised by

[2]Interactions between inputs occur, for example, when varying two or more input parameters produces a significantly different response from the model than would be expected from summing the change produced by varying the parameters independently.

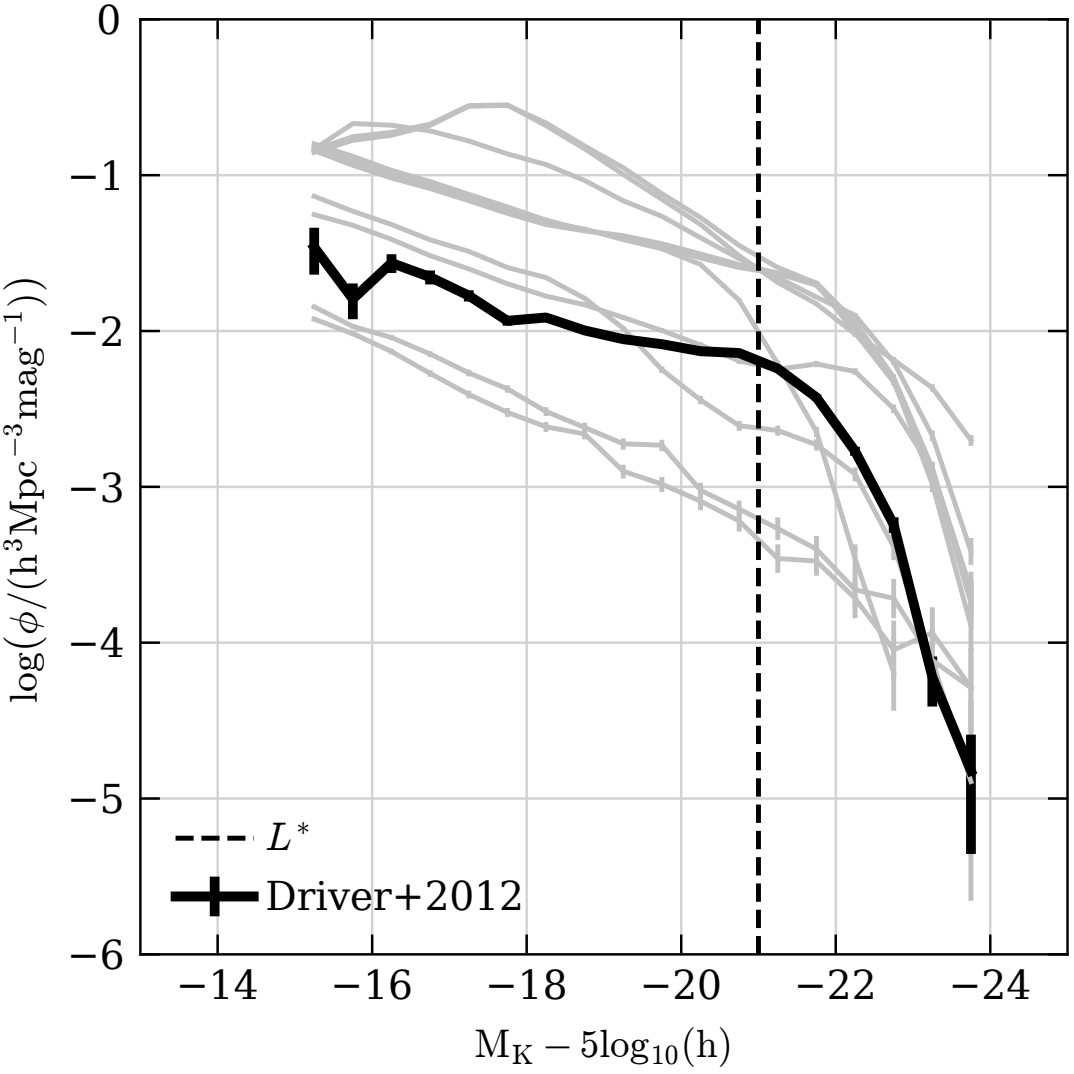

Figure 5.1: The K-band LF at $z = 0$ in the AB magnitude system. Gray lines represent 10 GALFORM model realisations randomly chosen from the 1600-model run series. The black line represents the observational data from Driver et al. (2012). The black vertical line is drawn at $L = L^*$, and separates the bright and the faint ends of the LF.

SA techniques. Similarly, variance-based SA currently assumes that the model output is a scalar. This means that the model outputs are independent of one another; for example in the case of the luminosity function, the model prediction in a luminosity bin is considered to be independent of the results in other bins and from other outputs, e.g. other galaxy statistics. Even if the output of the model is multi-dimensional, and even if it is correlated across one or more dimensions[3], each of the outputs must be analysed in isolation from the others. Unfortunately, at the time of writing there are no well-established techniques which quantify or alleviate these two shortcomings. However, these limitations do not apply to GALFORM: the input parameters can all be varied independently and freely across the entire parameter space, and our outputs will be quantised and analysed independently.

5.2.3 Sampling parameter space

Sampling the high-dimensional parameter space of a complex model requires a trade-off between the accuracy of the sampling and computational expense. The accuracy of the sampling describes how well the space is probed – have any potentially interesting regions of the parameter space been overlooked because too few points have been sampled or because the method used has left gaps in the space?

The accuracy of a sampling scheme can be assessed formally in terms of its "discrepancy". The lowest discrepancy sampling possible is a regular grid. However, this is subject to "aliasing" or a lack of resolution due to the fixed gaps in the parameter space between the model evaluations; interesting model behaviour could be hidden in the unsampled parts of the parameter space. The convergence of the exploration of the parameter space is slow with a regular grid. The aliasing can be reduced and the convergence rate sped up by using a random sequence to sample the parameter space, which leads to a higher density of sampling in some parts of parameter space compared to a regular grid. The drawback in this case is that some

[3]We know that this is the case in galaxy formation models because if the luminosity function changes in a given bin this will lead, for example, to a change in the luminosity - circular velocity relation. Benson (2014) argued that correlations between bins in the observed luminosity function are important in setting model parameters.

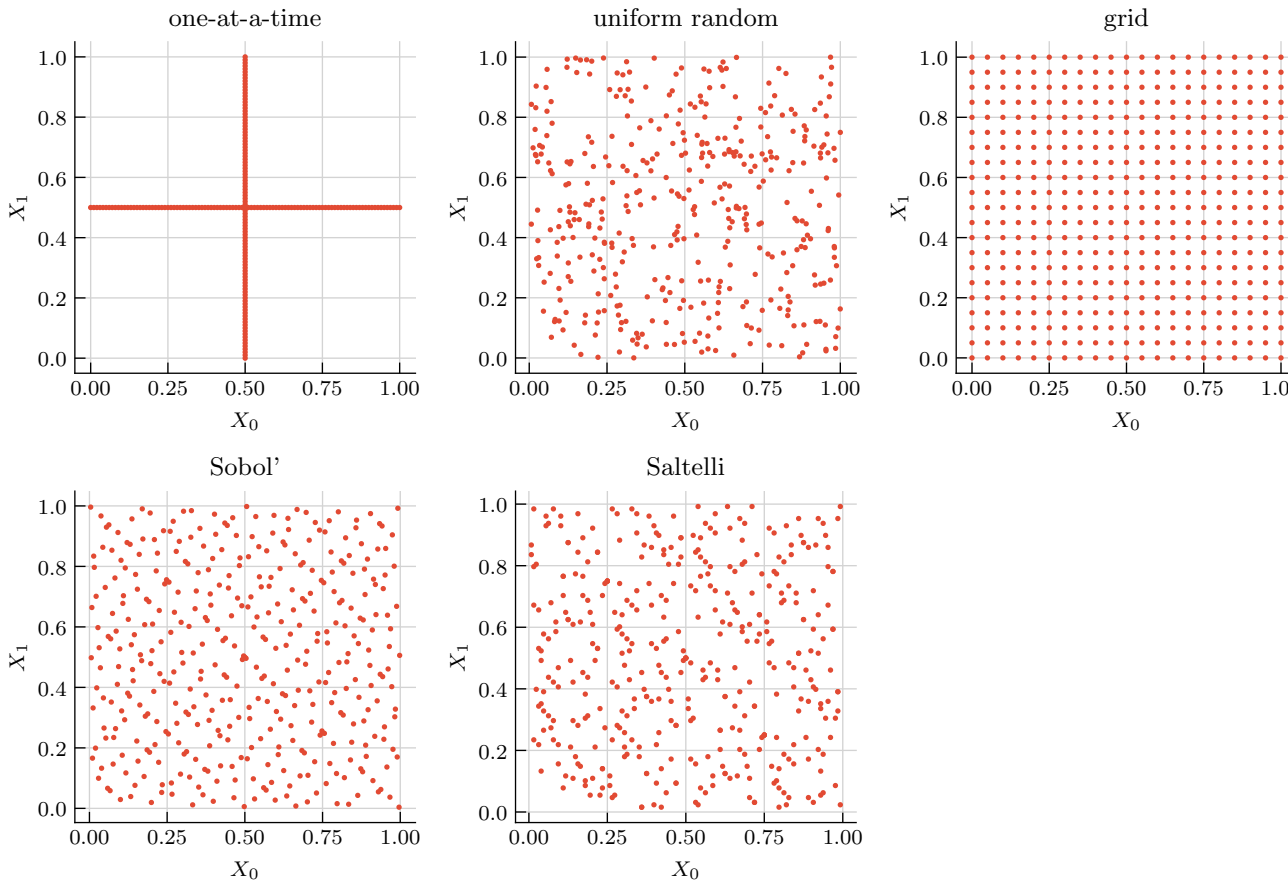

Figure 5.2: Comparison of selected parameter space sampling strategies. Each panel contains 400 points sampled between $[0, 1]$ in the X_0 and X_1 dimensions using different methods, as labelled in each panel.

regions of the parameter space will be more sparsely sampled than they were using a regular grid. A random sequence is formally described as the highest discrepancy sampling. Ideally, for a fixed number of sampling points, we want to strike a balance between avoiding the regular sampling achieved using a grid and leaving big gaps unsampled in the parameter space, as happens with random sampling.

Several quasi-random techniques have been proposed to generate sequences that approach this ideal of "low discrepancy" sampling, and which also ensure fast convergence of the uncertainties in the sensitivity indices. A quasi-random sequence is one designed to generate points in d-dimensions which appear random, but which are generated deterministically to have certain desired properties. Unlike pseudo- and truly-random sequences, successive points in a quasi-random sequence fill the gaps left by the previous points in the parameter space. The "random" part of the name is technically a misnomer, as the sequence is fully deterministic, but yields a uniform distribution when projected onto any dimension of the parameter space.

A quasi-random sequence can be designed to minimise its discrepancy. For a low-discrepancy sequence, all of its subsequences also have low discrepancy. If a given sequence is uniform, its discrepancy tends to zero as its length increases. For these reasons, quasi-random low-discrepancy sequences are used to maintain a balance between rapid convergence of numerical algorithms, a thorough coverage of the parameter space, and a high uniformity of a resulting sample along all dimensions of the parameter space (Press et al., 2007, §7.8). Quasi-random sequences are therefore an attractive replacement for pseudo-random sequences in many applications which require a high quality sampling.

Sampling based on low discrepancy sequences, such as the recurrent additive sequence (Ulam, 1960), Halton sequence (Halton, 1964), Latin hypercube (Stein, 1987) or Sobol' sequence (Sobol', 1967; Levitan et al., 1988) can be used in numerical integration and model optimisation and have been shown to outperform schemes based on truly random, or pseudo-random number generators, while achieving significantly faster convergence rates (Sobol', 1993). The advantage of these sequences over truly random and pseudeo-random sequences can be attributed to the fact that the low discrepancy property guarantees gap-less sampling over the entire parameter space.

The low discrepancy quasi-random sequence typically used in SA is the Sobol' sequence (Sobol', 1967). It can be efficiently calculated, and produces a sample which quickly converges to the correct set of sensitivity indices, as verified by checking against analytically calculated values for test models. Even though it is impossible to estimate the required number of model evaluations prior to running the SA, there exists a natural convergence criterion – the sum of the first-order indices, defined in Eq (5.4), has to add up to unity. Moreover, even if the SA did not converge after the initial run, additional evaluations can be easily added (see the example in the next subsection).

5.2.4 Saltelli sequence sampling

The Sobol' sequence was originally proposed as a method of improving the convergence of numerical integration (Sobol', 1967). Antonov & Saleev (1979) developed an efficient computational method to implement Sobol' sampling. Saltelli et al. (2010) combined multiple Sobol' sequences to further reduce the number of points required for the estimation of the sensitivity indices, improving the convergence rate.

Hereafter we refer to the Sobol' sequence as an N by d matrix, where N is the number of points of a d dimensional parameter space.

The Saltelli sequence is obtained as follows: first we generate an N by $2d$ Sobol' sequence, (as demonstrated for the case of $N = 4$, $d = 3$ in the first line of Eq (5.1)). Let the first d columns be called submatrix $\mathbf{A}$ (blue), and the last d submatrix $\mathbf{B}$ (red). The values in the matrices indicate the locations in parameter space at which the model is to be evaluated, for parameters which can take on values over the range 0 to 1. We next construct a number d of N by d matrices $\mathbf{A}_{\mathbf{B}}^{(\mathrm{i})}$, for $i \in \{1, 2, ..., d\}$, such that for each $\mathbf{A}_{\mathbf{B}}^{(\mathrm{i})}$ the i^{th} column is taken from matrix $\mathbf{B}$, while the remaining columns come from matrix $\mathbf{A}$. The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{A}_{\mathbf{B}}^{(\mathrm{i})}$ specify all the points of the parameter space at which the model is to be evaluated (one point per row), giving a total of $N \times (2 + d)$ evaluations which are required to calculate the first order sensitivity indices.

$$\text{Sobol}(4,3) = \begin{bmatrix} 0.500 & 0.500 & 0.500 & 0.500 & 0.500 & 0.500 \\ 0.250 & 0.750 & 0.250 & 0.750 & 0.250 & 0.750 \\ 0.750 & 0.250 & 0.750 & 0.250 & 0.750 & 0.250 \\ 0.125 & 0.625 & 0.875 & 0.875 & 0.625 & 0.125 \end{bmatrix} \tag{5.1}$$

$$\mathbf{A}_{\mathbf{B}}^{(1)} = \begin{bmatrix} 0.500 & 0.500 & 0.500 \\ 0.750 & 0.750 & 0.250 \\ 0.250 & 0.250 & 0.750 \\ 0.875 & 0.625 & 0.875 \end{bmatrix}$$

$$\mathbf{A}_{\mathbf{B}}^{(2)} = \begin{bmatrix} 0.500 & 0.500 & 0.500 \\ 0.250 & 0.250 & 0.250 \\ 0.750 & 0.750 & 0.750 \\ 0.125 & 0.625 & 0.875 \end{bmatrix}$$

$$\mathbf{A}_{\mathbf{B}}^{(3)} = \begin{bmatrix} 0.500 & 0.500 & 0.500 \\ 0.250 & 0.750 & 0.750 \\ 0.750 & 0.250 & 0.250 \\ 0.125 & 0.625 & 0.125 \end{bmatrix}$$

A visual impression of the different sampling approaches is given by Fig. 5.2 which shows five commonly used types of sampling: OAT, uniform pseudo-random number generator, uniform grid sampling, a two-dimensional Sobol' sequence and Saltelli sampling. The OAT approach is often used with far fewer evaluations than shown here, which makes it computationally cheaper than the other approaches. The drawback of this method is clear from the vast areas of the parameter space that are left unexplored. This problem is only exacerbated on increasing the dimensionality of the parameter space. The pseudo-random number generation suffers from poor convergence, as randomness often results in over and under sampling of many regions. The Sobol' and Saltelli sequences uniformly sample the parameter space and achieve the low discrepancy target at a reasonable computational cost.

5.2.5 Sensitivity indices

Given a scalar model Y with independent inputs, we can define the first order effect of the variance in the input X_i as:

$$E_i = E_{\mathbf{X}_{\sim \mathrm{i}}}\left(Y|X_i\right) = \int Y(X_i)\mathrm{pdf}(X_i)\prod_{i\neq j}^{d}\mathrm{d}X_i \tag{5.2}$$

$$V_i = \mathrm{Var}_{X_i}\left(E_i\right) = \int \left(E_i - E(Y)\right)^2 \mathrm{pdf}(X_i)\mathrm{d}X_i, \tag{5.3}$$

where X_i is i^{th} model input, V_i is the variance integrated in X_i space over dimension i, and E_i is the mean Y value, integrated over the d-dimensional X space in all dimensions except i. Variance is used in the conventional sense, as the expected value of the squared deviation from the mean; sensitivity of the model is understood as variance of the mean of each model parameter, normalised by the total variance. Since V_i can only take values between 0 and $\mathrm{Var}(Y)$, the total variance in the model output, we define first-order sensitivity index S_i as

$$S_i = \frac{V_i}{\mathrm{Var}\left(Y\right)}, \tag{5.4}$$

which measures the effect that varying the input X_i has on the output, averaged over variations of all other inputs ($\mathrm{Var}(Y)$ is defined in Eq (5.6)). S_i is the isolated, normalised variance of model output, averaged over all input parameters except parameter i. If $S_i = 1$, all variance in Y comes from the variance in X_i, whereas if $S_i = 0$, none of it does, and Y is independent of X_i.

In order to measure the interactions between model parameters, we can define higher order indices. For second order interactions, the combined variance is

$$V_{ij} = \mathrm{Var}_{X_{ij}}\left(E_{\mathbf{X}_{\sim ij}}\left(Y|X_i, X_j\right)\right) - V_i - V_j, \tag{5.5}$$

from which $S_{i,j}$ can be calculated analogously to S_i.

It should now be obvious from the definition of the model variance why the OAT methods are inappropriate for complex models – they do not consider the full contribution to the model variance given by Eq (5.2) (which averages over *all* values of the other inputs, instead of being measured only at a designated slice, as shown in the relevant panel of Fig. 5.2), nor does OAT treat the combined variance of two (Eq (5.5)) or more variables correctly.

For a deterministic model, the only source of variance in the output is the variances of the inputs. Therefore, from variance decomposition it follows that

$$\sum_{i=1}^{d} V_i + \sum_{i<j}^{d} V_{ij} + ... + V_{12...d} = \mathrm{Var}(Y), \tag{5.6}$$

which we normalise to obtain the sensitivity indices of all orders

$$\sum_{i=1}^{d} S_i + \sum_{i<j}^{d} S_{ij} + ... + S_{12...d} = 1. \tag{5.7}$$

A direct consequence is that, in order to analytically decompose the total variance of the model, one needs to compute variances of $2^d - 1$ variables, which can be computationally expensive for complex models. However, if we assume that the indices decrease as their order increases (which is correct for the model of interest here), we might be less interested in the precise values of higher-order contributions, and focus instead on the total higher-order response of a given variable. In this case it is convenient to combine the higher-order terms into a total-order index

$$S_{Ti} = \frac{E_{\mathbf{X}_{\sim \mathrm{i}}}\left(\mathrm{Var}_{X_i}\left(Y|\,\mathbf{X}_{\sim \mathrm{i}}\right)\right)}{\mathrm{Var}(Y)} = 1 - \frac{\mathrm{Var}_{\mathbf{X}_{\sim \mathrm{i}}}\left(E_{X_i}\left(Y|\,\mathbf{X}_{\sim \mathrm{i}}\right)\right)}{\mathrm{Var}(Y)}, \tag{5.8}$$

containing all terms of the decomposed output variance which include X_i. Unlike the first-order indices, the S_{Ti} do not have to add up to 1, as they include all the input interactions[4].

Higher order effects can also be calculated in simpler analyses, such as Analysis of Variance (Fisher, 1918, ANOVA), High Dimensional Model Representations (Sobol', 1993, HDMR) or derivative-based methods. However, the total indices are a unique feature of the variance-based SA, and are a major advantage of this methodology, as they allow for a direct comparison of the linear and non-linear impacts of the input parameters.

Since model evaluations produce a discrete data set, we use the numerical approximations in order to evaluate Eqs. (5.4) and (5.8). The appoximate expressions

[4]In this case, the whole is literally more than the sum of the parts.

are based on the sampling matrices ($\mathbf{A}, \mathbf{B}$, and $\mathbf{A}_{\mathbf{B}}^{(\mathrm{i})}$),

$$\mathrm{Var}_{X_i}\left(E_{\mathbf{X}_{\sim \mathrm{i}}}\left(Y|X_i\right)\right) \approx \frac{1}{N}\sum_{j=1}^{N} f\left(\mathbf{B}\right)_j \left(f\left(\mathbf{A}_{\mathbf{B}}^{(\mathrm{i})}\right)_j - f\left(\mathbf{A}\right)_j\right) \tag{5.9}$$

$$E_{\mathbf{X}_{\sim \mathrm{i}}}\left(\mathrm{Var}_{X_i}\left(Y|\mathbf{X}_{\sim \mathrm{i}}\right)\right) \approx \frac{1}{2N}\sum_{j=1}^{N} \left(f\left(\mathbf{A}\right)_j - f\left(\mathbf{A}_{\mathbf{B}}^{(\mathrm{i})}\right)_j\right)^2, \tag{5.10}$$

where $f\left(\mathbf{X}\right)$ is the model f evaluated at point X (Jansen, 1999; Sobol', 2001; Saltelli et al., 2010).

5.2.6 Illustrative sensitivity analysis of a toy model

The performance of the sensitivity analysis estimator can be demonstrated using a toy model. The Ishigami function is an example of such a model, and is commonly used to test the predictions of sensitivity analysis because it contains non-linear interacting terms. Nevertheless, the sensitivity indices can be calculated analytically and compared to the estimated values.

The Ishigami function is defined by Eq. 14 of Ishigami & Homma (1991) as

$$Y\left(X_1, X_2, X_3\right) = \sin\left(X_1\right) + a\sin^2\left(X_2\right) + b\,X_3^4\sin\left(X_1\right), \tag{5.11}$$

where the X_i are random variables uniformly distributed between $-\pi$ and π, such that $\mathrm{pdf}(X_i) = U(-\pi, \pi)$, and a, b are numerical constants, here chosen to be 7 and 0.1 respectively.

The SA was carried out by running the model on inputs generated by a 3-dimensional Sobol' sequence for 500 realisations, which resulted in $4000 = 500 \times (2 + 2 \times 3)$ values of X_i (as explained in Section 5.2.4). Next, Eq (5.11) was evaluated at each X_i point, giving a vector Y of length 4000. Finally, the vector Y was analysed using the `SALib` Python package (Herman & Usher, 2017).

The evaluations of Eq (5.11) are shown in Fig. 5.3, and the first- and total-order sensitivity indices of the three input parameters are shown in Fig. 5.4. It is interesting to draw some qualitative observations from Fig. 5.3:

- varying X_1 and X_2 in isolation results in large changes in Y; this is reflected by large values for S_1 and S_2.

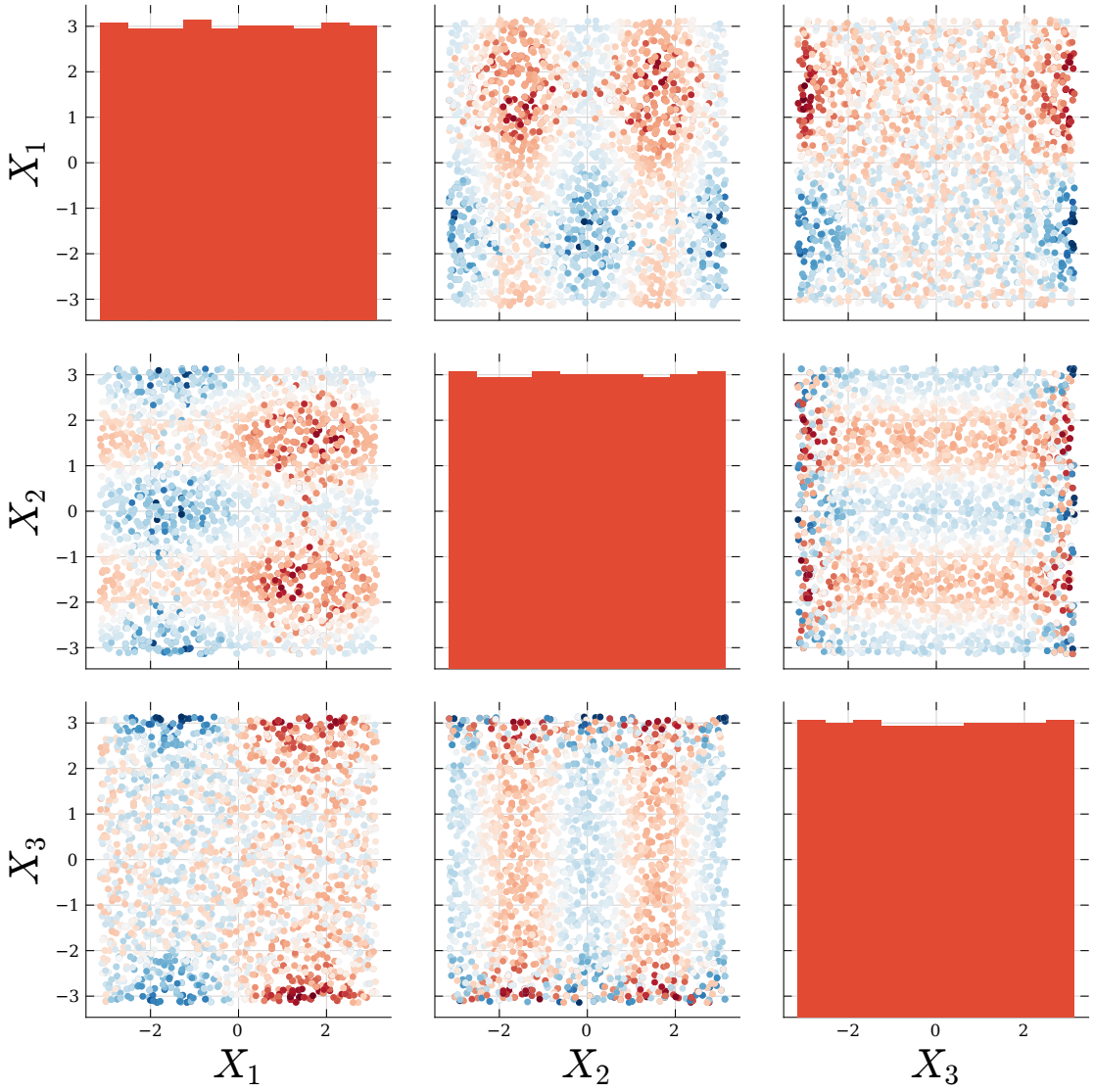

Figure 5.3: The 4000 evaluations of the Ishigami function (Eq (5.11). On-diagonal histograms show distribution of the X_i parameters. Off-diagonal scatter plots show pairs of parameters and are colour-coded to show the value of output Y.

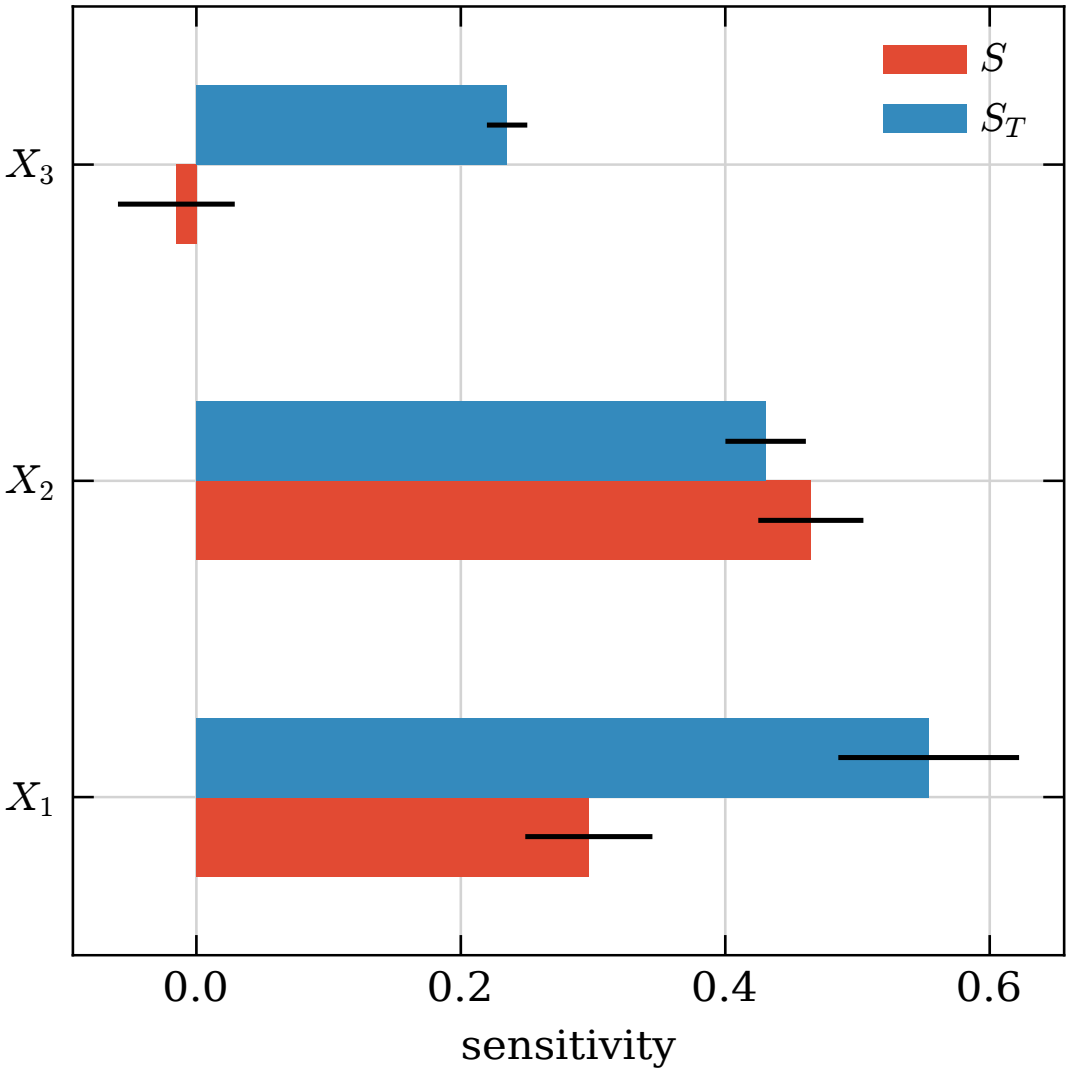

Figure 5.4: First-order (red) and total (blue) sensitivity indices of the three input parameters of the Ishigami function, with 1σ confidence bars (black).

- varying X_1 and X_2 together has a large effect on Y; this is reflected by large values for S_{T1} and S_{T2}.

- varying X_3 for mid-range values produces little effect, but varying other parameters at extreme X_3 values produces a large change in Y; correspondingly, S_3 is nearly zero, but S_{T3}, which captures the global response of Y to X_3, is larger

- S_1 is negative, despite being defined in terms of non-zero variance (Eq (5.4)); this is the result of using a numerical approximation instead of an analytical formula to estimate S_1; however, note that the confidence interval includes the origin, and so the value of S_1 is consistent with zero.

A more complete SA would involve computing second order indices, and comparing sensitivity indices for different versions of the Ishigami function, such as with different a, b parameters, or over different X_i ranges. However, this more complete analysis is beyond the scope of this section, as it is only meant for demonstration purposes. The source code used to reproduce this analysis has been made public: `https://github.com/oleskiewicz/sensitivity/releases/tag/v1.0`.

5.2.7 GALFORM output used in the sensitivity analysis

When applying SA to a model with a multi-dimensional output, it is necessary to select the most interesting outputs manually. The Sobol' index method assumes, and can only be calculated for, separate one-dimensional output vectors Y. From the formal standpoint this is problematic as the sensitivity indices contain no information about any correlations between various model outputs. However, in practice one could perform model runs which follow the Saltelli sampling and then carry out separate sensitivity analyses for any desired number of model outputs, since running the model is more time consuming than calculating the Sobol' indices.

Here we focus on the prediction of GALFORM for the K-band luminosity function (LF) at $z = 0$, calculated as described in Section 5.2.1. We have chosen to consider this statistic due to the well-understood influence of the model parameters on the form of the luminosity function (see the extensive discussion in Lacey et al. (2016)).

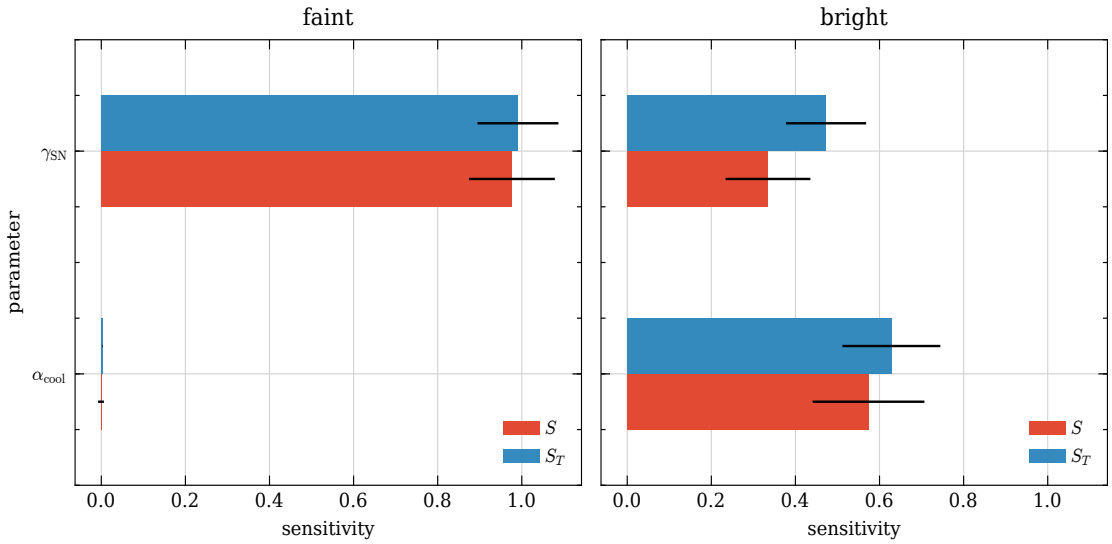

Figure 5.5: Sensitivity indices for the first series of runs, when varying two parameters in GALFORM: $\alpha_{\rm cool}$ and $\gamma_{\rm SN}$, for the K-band luminosity function measured in two coarse luminosity bins. The colours of the bars indicate different indices, first (blue) (Eq (5.4)) and total (red) (Eq (5.8)) order for a given variable. The left panel shows indices for $L < L^*$, and the left for $L > L^*$ (see Eq (5.12)). The black bars show the 1σ confidence interval for the sensitivity indices.

Varying the parameters around the values used in the fiducial model shows that the bright and faint ends of the luminosity function are regulated by different physical processes. Therefore, the sensitivity indices could be easily verified for errors, and we will be able to quantify our intuition regarding the relative importance of the different feedback modes on the abundance of galaxies at different luminosities.

We have elected to perform the analysis on the model output values normalised by the observational data (see Eq (5.13)) instead of on the model output itself. This way, the values we focused on were close to the ones typically used for model optimisation, and had a reduced dynamic range, being effectively normalised by the observational values. Analysing a SAM independently of the observational constraints, while interesting in its own merit, is outside the scope of this work.

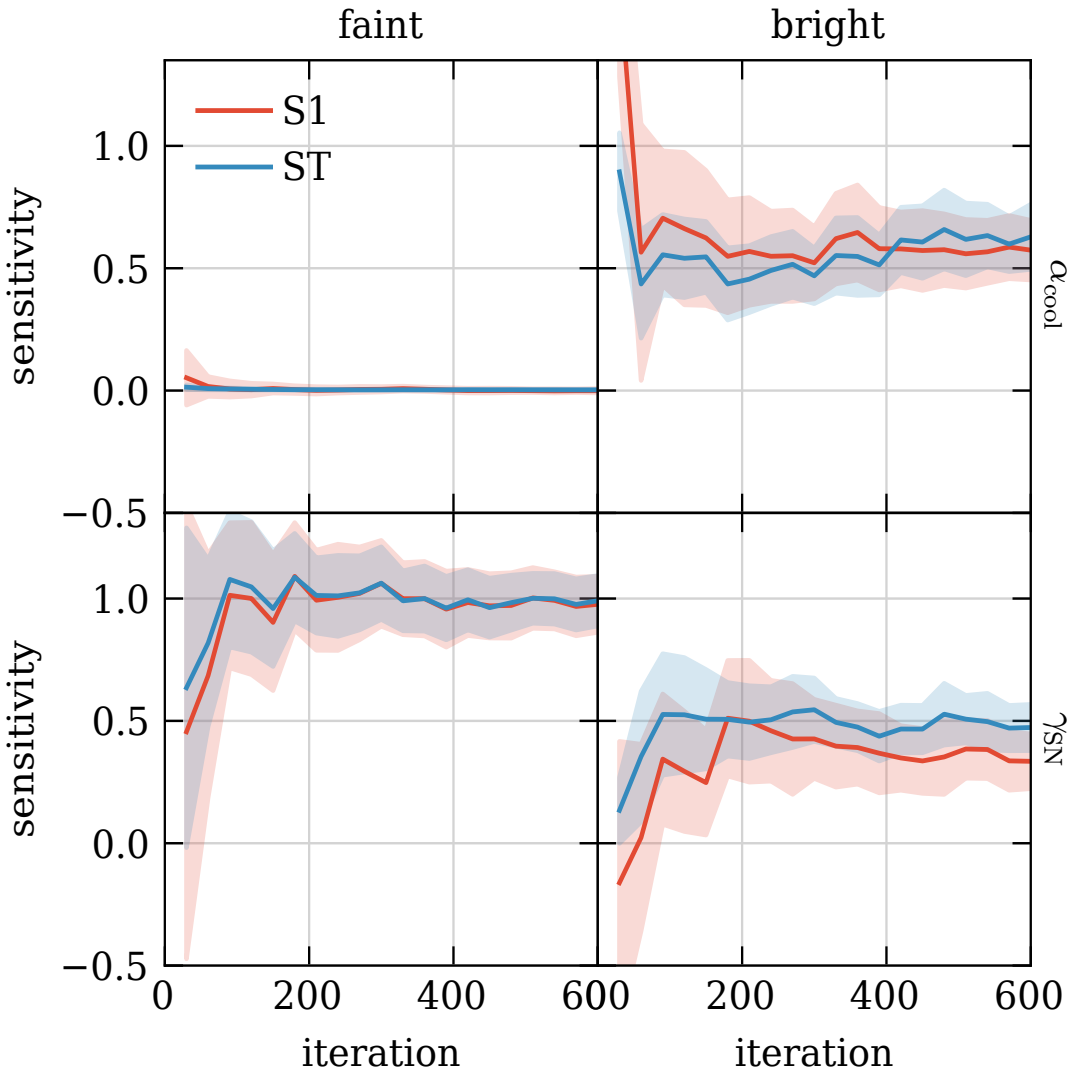

Figure 5.6: Convergence of the first- and total-order sensitivity indices for the first series of runs, when varying two parameters of GALFORM, $\alpha_{\rm cool}$ and $\gamma_{\rm SN}$, as a function of a number of samples. The sensitivity indices in this case are computed in each of two broad luminosity bins, covering, respectively, the faint and bright ends of the luminosity function. Individual subplots show the results for the faint and bright end of the K-band LF (columns, labelled on the top), and the α parameters (rows, labelled on the right). Solid lines correspond to the values of the indices, and the shaded regions to the 1σ confidence band of the values, both colour-coded by the order of the indices as labelled in the legend.

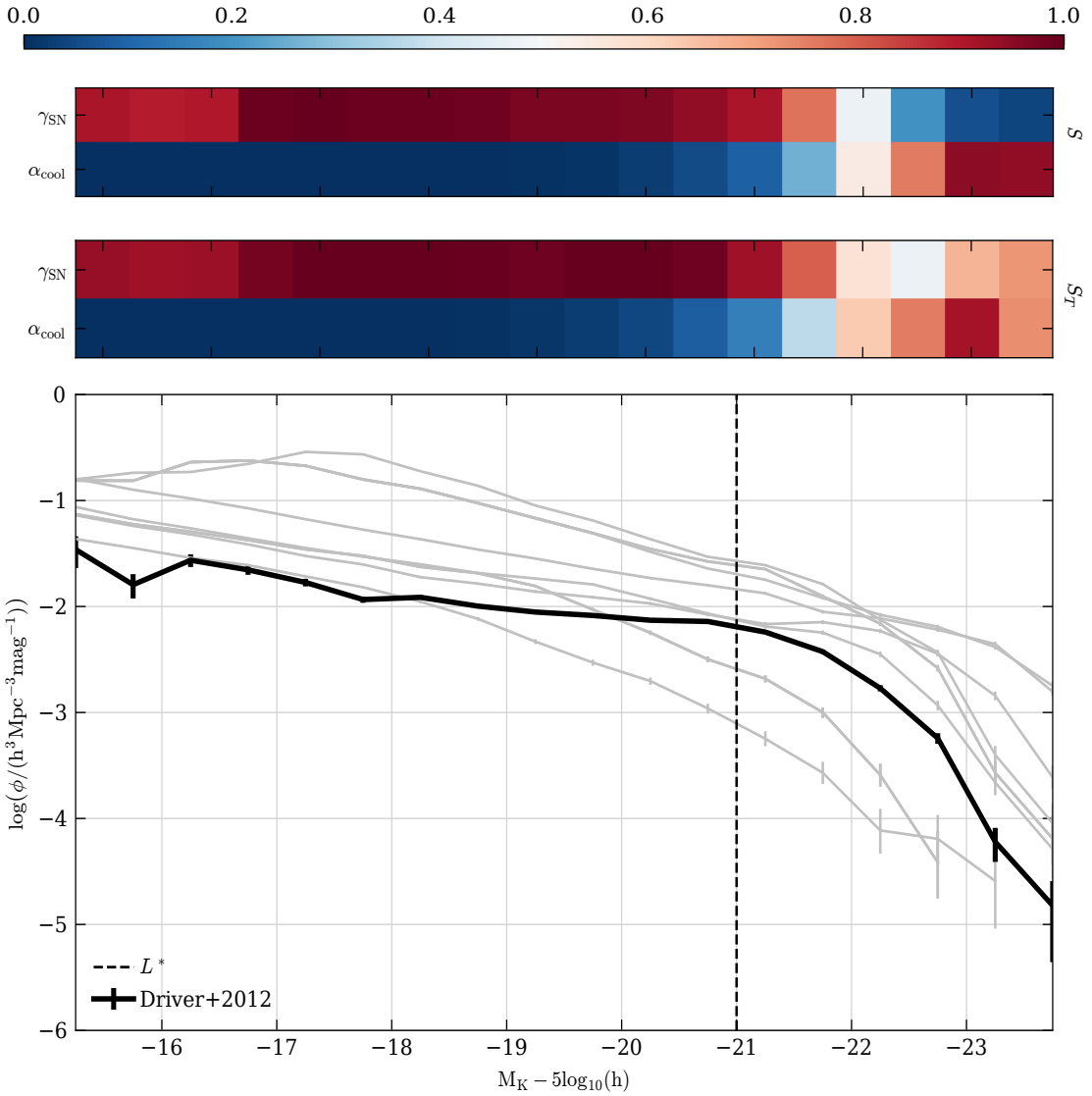

Figure 5.7: First- and total-order sensitivity indices (Eq (5.4)) for the first series of runs, when varying two parameters in GALFORM: $\alpha_{\rm cool}$ and $\gamma_{\rm SN}$. *Bottom panel:* K-band luminosity function at $z = 0$ like in Fig. 5.1; gray lines correspond to 10 randomly chosen runs; black line is the observational data from Driver et al. (2012); dashed vertical line corresponds to L^*. *Top panels:* first- and total-order (as labelled on the right) sensitivity indices of two variables (y axis) for 18 individual magnitude bins (x axis), colour-coded by value between 0 (not sensitive) and 1 (most sensitive) as labelled by the colourbar at the top.

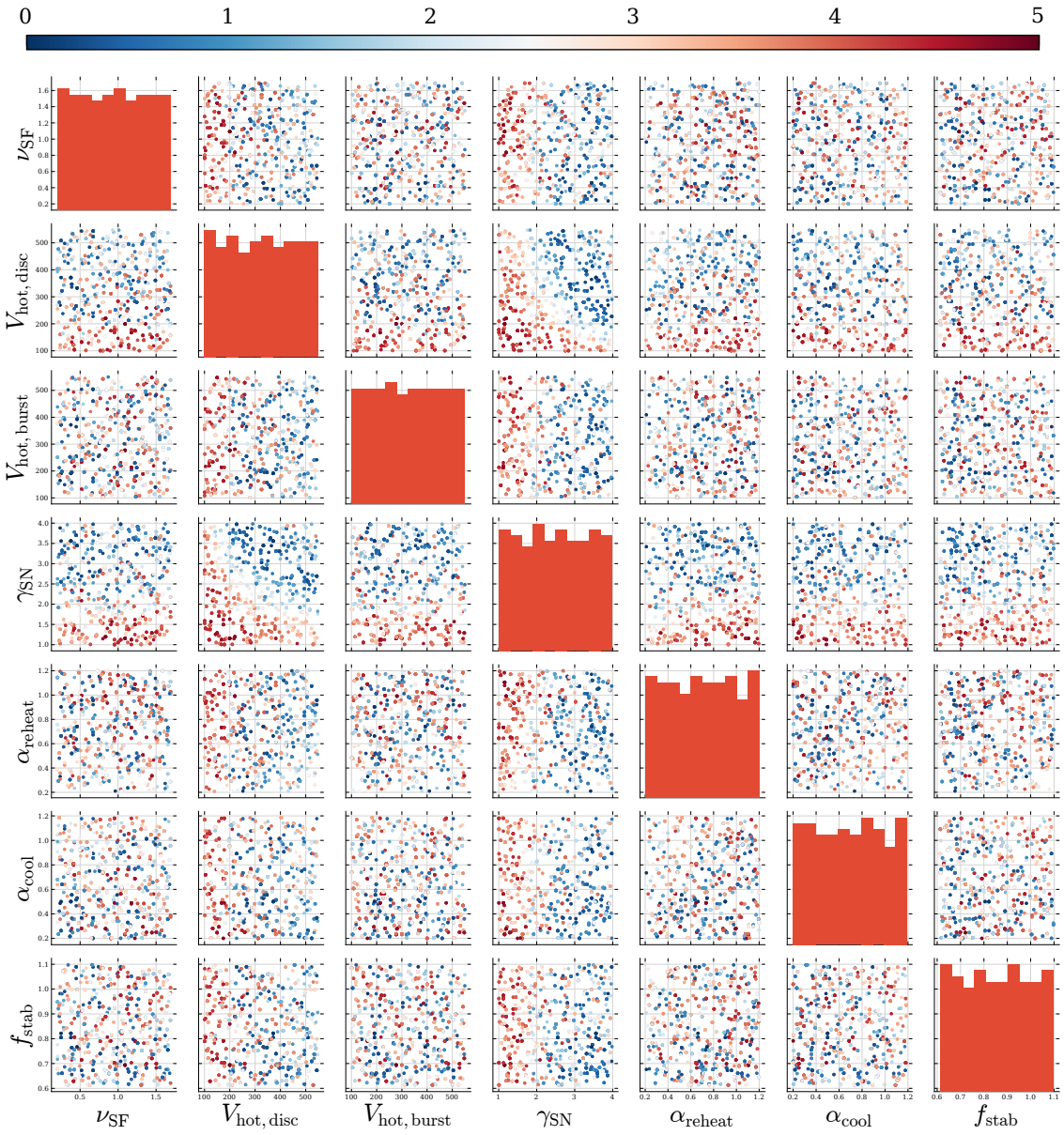

Figure 5.8: The parameter space of the second GALFORM experiment in which 7 parameters are varied over 1600 realisations of the model. On-diagonal histograms show the nearly-uniform distribution of the individual parameters, as expected for Saltelli sampling. Off-diagonal scatter plots show the parameter space for pairs of parameters, colour-coded by the goodness-of-fit χ^2 (Eq (5.14)) of the model prediction for the K-band luminosity function using 18 luminosity bins (Fig. 5.1), to the observational estimate from Driver et al. (2012); blue points correspond to runs with low values of χ^2, as labelled by the colourbar.

5.3 Results

5.3.1 Sensitivity analysis experiments

We have carried out two separate sensitivity analyses using GALFORM: (1) 600 GALFORM model runs varying two parameters ($\alpha_{\rm cool}$ and $\gamma_{\rm SN}$), and (2) 1600 model GALFORM runs varying seven parameters (see Table 5.2).

For both series of runs, a SA was carried out on the K-band LF at $z = 0$, with two different binnings of the LF used to compute the sensitivity indices, as explained below.

In the first instance we performed a simple analysis by splitting the LF into two broad luminosity bins, one covering a range of luminosities brighter than L^* and the other luminosities fainter than L^* (see Fig. 5.1). For each run, we calculated two model outputs covering the bright and faint ends of the LF, $d_{\rm faint}$ and $d_{\rm bright}$, defined by summing the normalised differences between the observed and predicted values of the luminosity function for luminosities brighter and fainter than L^*. e.g.:

$$d_{\rm faint} = \sum_{L<L^*} \frac{\log_{10}(\phi) - \log_{10}(\hat{\phi})}{\log_{10}(\phi)}, \tag{5.12}$$

with $d_{\rm bright}$ defined analogously for $L > L^*$. The observed luminosity function $\hat{\phi}$ is taken from Driver et al. (2012). Unlike a traditional measure of model fitness, we *do not* take the absolute value or square of the distance between the model prediction and the data. This is because the sign of the output (i.e. the sense of the discrepancy) is valuable information for the sensitivity indices, as it contains the direction of the model response.

This coarse analysis is quantitatively identical to measuring the LF using only two broad luminosity bins. This exercise has two goals: (1) to verify that SA produces explainable results which can be interpreted in accordance with our physical intuition about the galaxy formation model, and (2) to check the convergence of the sensitivity indices and their confidence intervals, which can be estimated as explained in Section 5.2.2.

After this coarse two-bin analysis, in the second case we calculate sensitivity

indices for each of the 18 luminosity bins, L_i, using the quantity:

$$d_{\rm i} = \frac{\log_{10}(\phi_i) - \log_{10}(\hat{\phi}_i)}{\log_{10}(\phi_i)}. \tag{5.13}$$

This serves as a fine-grained analysis, which can quantify the relative impact of different parameters on the individual segments of the LF, as well as uncovering interactions between model parameters.

5.3.2 Feedback processes and the luminosity function

The first series of runs, which analysed the effects of changing two of the parameters which specify different feedback processes in GALFORM, $\alpha_{\rm cool}$ and $\gamma_{\rm SN}$, was carried out to verify the usefulness of the SA and to evaluate its effectiveness, given our physical intuition, regarding the expected impact on the LF of varying these model parameters. Only two parameters were allowed to vary to speed-up the analysis and allow for an easier interpretation of results: $\alpha_{\rm cool}$ and $\gamma_{\rm SN}$ (see Table 5.2 for the range of parameter values considered). Recall that $\gamma_{\rm SN}$ controls the mass loading of SNe driven winds and $\alpha_{\rm cool}$ determines the halo mass above which AGN heating shuts down the cooling flow.

Fig. 5.5 shows the first- and total-order sensitivity indices calculated from 600 GALFORM model runs for the coarse-bin analysis of luminosity function using two bins, one fainter and one brighter than L^*, as presscribed by Eq (5.12). The results are striking, but not unexpected: it is clear that $\gamma_{\rm SN}$ is the dominant parameter out of the two in shaping the model output for galaxies fainter than L^* (and hence, that such galaxies are mainly affected by SNe feedback) and that both parameters have similar significance for galaxies brighter than L^* (albeit $\alpha_{\rm cool}$ is slightly more important), and so bright galaxies are affected by SNe feedback and AGN heating. Moreover, $\rm S_1$ and $\rm S_T$ are comparable in all cases, which means that the model response to varying these parameters is mostly linear.

Figure 5.6 shows the convergence of the indices from the Fig. 5.5 as a function of the number of samples N. The indices do not change substantial after 100 GALFORM runs.

Figure 5.7 shows the first- and total-order sensitivity indices for the fine-grained

Table 5.3: The best-fit GALFORM parameters found in this work, in relation to Driver et al. (2012).

parameter	value
$\nu_{\rm SF}$ [Gyr^{-1}]	0.46
$\gamma_{\rm SN}$	3.45
$\alpha_{\rm reheat}$	0.74
$V_{\rm hot,disc}$ [km/ sec]	332.69
$V_{\rm hot,burst}$ [km/ sec]	392.90
$\alpha_{\rm cool}$	0.58
$f_{\rm stab}$	0.77

analysis of the LF using multiple luminosity bins, using Eq (5.13) as model output. We can see that L^* is close to coinciding with the bin at which AGN heating starts to become important, which explains the results shown in Fig. 5.5. We also learn that while SNe feedback does not interact with the AGN heating at the faint end of the LF, their influence over the bright end is strongly correlated.

We did not consider the best-fit model for this two parameter case, since we perform a rudimentary estimate of the best-fitting parameter set in the next section, when varying more GALFORM model parameters at the same time.

5.3.3 Sensitivity analysis over a multi-dimensional parameter space

The design of the second experiment, in which seven GALFORM parameters are varied simultaneously (Table 5.2), is inspired by the work on parameter optimisation using Bayesian emulators by Bower et al. (2010); Rodrigues et al. (2017). For comparison, we use the same parameter ranges adopted in their studies. This exercise requires significantly more model realisations than the first one, since we sample a higher dimensional parameter space and aim to observe interactions between more parameters.

Fig. 5.8 shows the parameter space and its sampling, colour-coded by the goodness-

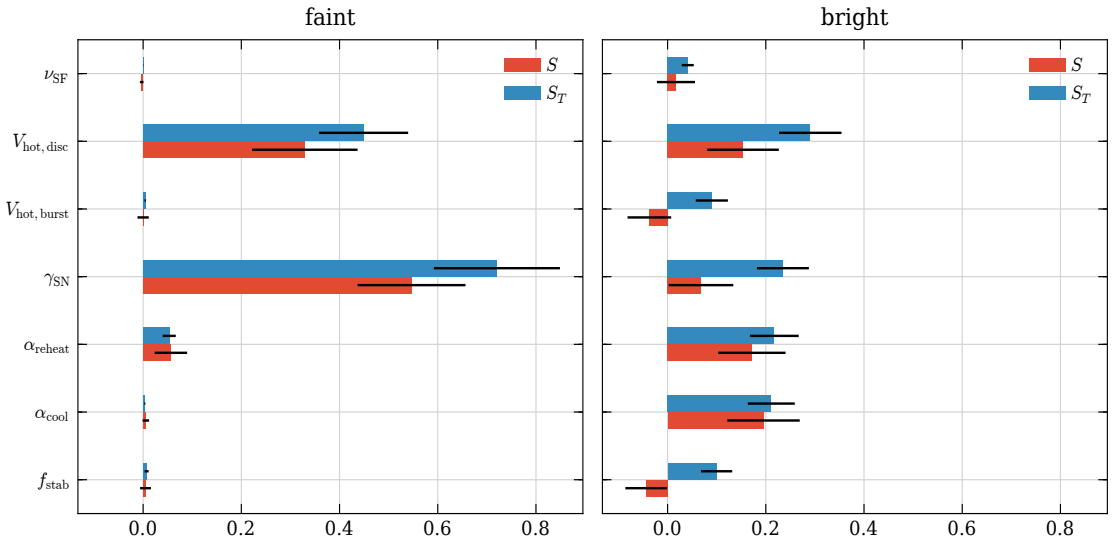

Figure 5.9: Sensitivity indices for the second series of 1600 GALFORM runs, varying 7 model parameters (Table 5.2), computed using the coarse two bin description of the luminosity function. The bar colours indicate the values of different indices, the first index (Eq (5.4)) (S1, red) and total order index (Eq (5.8)) (ST, blue) for each parameter. The left panel shows indices for galaxies in the luminosity bin fainter than L^*, and the right panel for galaxies in the bin brighter than L^* (Eq (5.12)). The black bars show the 1σ confidence intervals for the sensitivity indices.

of-fit measure

$$\chi^2 = \sum_i \frac{\left(\log_{10}(\phi_i) - \log_{10}(\hat{\phi}_i)\right)^2}{\log_{10}(\phi_i)}, \tag{5.14}$$

where the sum is carried out over all luminosity bins and low values of χ^2 are blue. While χ^2 is not a robust model output for SA, as it does not contain information about the *direction* of the model response as explained in Section 5.3, it is still a useful measure of a global model response or "quality of fit". The on-diagonal histograms indicate that the Saltelli sampling produces a nearly uniform sampling of parameter space, as expected from a low-discrepancy sequence. The off-diagonal scatter plots give a first indication of some of the first-order index results: the χ^2 of the model LFs is sensitive to variation of $\gamma_{\rm SN}$, is degenerate in the $\gamma_{\rm SN}$–$v_{\rm hot,disc}$ plane (which follows directly from Eq (2.15)), and depends only weakly on other parameters.

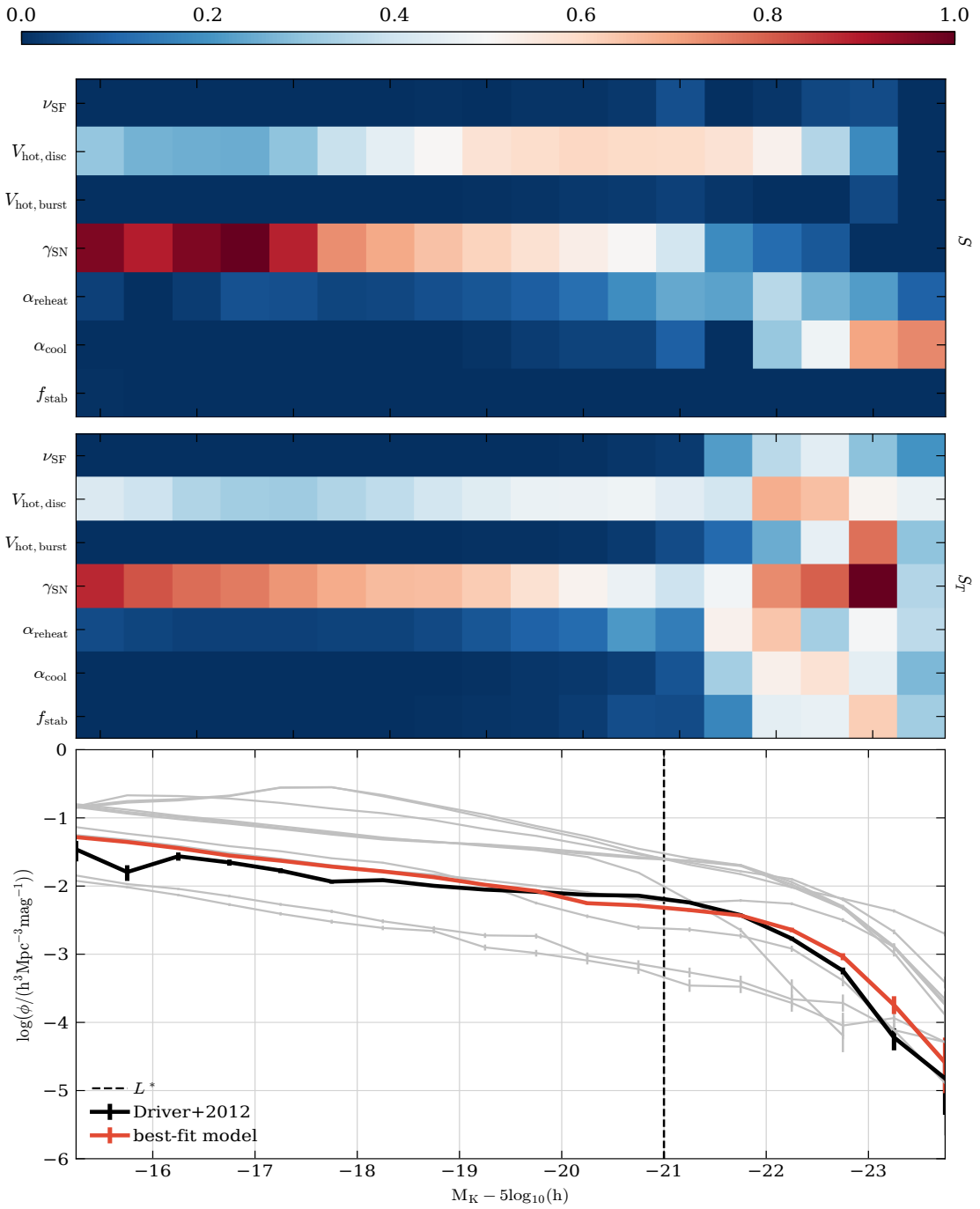

Figure 5.10: First- and total-order sensitivity indices (Eq (5.4)) for the second series of runs, when varying 7 GALFORM parameters (Table 5.2). *Bottom panel*: K-band luminosity function at $z = 0$ as in Fig. 5.1; grey lines show 10 randomly chosen GALFORM models; the black line connects observational data from Driver et al. (2012); dashed vertical line corresponds to L^*; the solid red line shows the best-fitting model. *Top panels:* first- and total-order (as labelled on the right) sensitivity indices of two variables (y axis) for individual magnitude bins (x axis), colour-coded by value between 0 (not sensitive) and 1 (most sensitive) as labelled by the colour-bar at the top.

Figure 5.9 shows the first- and total-order sensitivity indices for the coarse-binned analysis of the LF (Eq (5.12)). Since sensitivity indices are derived from the normalised variance (Eqs. (5.4) and (5.8)), the values should always be between 0 and 1, which they are (including the confidence interval). Similarly to Fig. 5.5, in Fig. 5.9 we see two different types of behaviour of the GALFORM model: the faint end is dominated by SNe feedback, while the bright end has a mixed, non-linear response to many parameters. Interestingly, while AGN feedback (via $\alpha_{\rm cool}$) has the highest first-order sensitivity index (S1) for the bright end of the LF, the total-order indices (ST) of SNe feedback processes dominate. Of particular interest are the $f_{\rm stab}$ and $v_{\rm hot,burst}$ parameters. These parameters have nearly zero first-order response indices (which means that their impact cannot be detected by an OAT analysis), but their combined higher-order responses are significant.

It is instructive to see the origin of the values reported in Fig. 5.9, by inspecting how the sensitivity changes bin-by-bin (Eq (5.13)) in Fig. 5.10. The results are consistent with Section 5.3.2, and together provide an interpretation of the behaviour of the GALFORM model. Moreover, displaying the model output together with model sensitivity can be of use when manually tweaking the model, allowing for a fine, manual control over the precise details of the LF (or, indeed, other outputs).

Finally, we note that Fig. 5.10 also shows the LF for the best-fitting model, as determined by the smallest value of Eq (5.14). This can be considered an additional benefit of running SA – requiring so many model realisations naturally finds one which is likely to be close to a global optimum. The best-fitting parameter values are reported in Table 5.3. Note that the values diverge from those reported in Lacey et al. (2016), due to different fitting method and the fact that this study only considered the K-band LF, whereas Lacey et al. (2016) took into account multiple observations in a manual parameter tuning. Of particular interest is the value of $V_{\rm hot,disc}$, which is over 20% larger than in the previous calibration of this GALFORM model. We attribute this difference to the fact that, as discussed in Section 5.3.3 and shown in Fig. 5.9, the combined total-order sensitivity index of $V_{\rm hot,disc}$ outweighs the first-order index for both ends of the K-band LF. This suggests that the optimal value of this parameter could be missed by OAT model fitting. The differences in the

other parameter values are not as significant as they might seem – the variables with the highest sensitivity match the previously reported values pretty closely (e.g. $\gamma_{\rm SN}$ is within 7%), and the variables with low sensitivity that diverge by a significant margin by definition of the sensitivity indices do not have significant impact on the K-band LF.

5.4 Conclusions

We have used variance-based sensitivity analysis to analyse the sensitivity of the K-band luminosity function predicted using the GALFORM semi-analytical model of galaxy formation to the variation of the model parameters. We have shown that sensitivity analysis is a useful tool, which goes beyond simple model fitting and one-at-a-time parameter variation, and we have demonstrated that it can be applied to a challenging problem in computational astrophysics. Variance-based sensitivity analysis is perhaps particularly useful for the semi-analytic modelling of galaxy formation modelling, due to the computational expense of searching a multi-dimensional parameter space and the non-linearity of the model. These features have led some to view such models as black boxes. Part of the aim of the sensitivity analysis presented here is to make the behaviour of the model and how it responds to parameter changes more transparent.

In its present form sensitivity analysis can only deal with one-dimensional outputs of a model, which on the one hand means that it cannot be used to resolve correlations in model outputs (such as between the predictions in different bins of the luminosity function or between the luminosity function in different bands; see Benson (2014)), yet on the other hand this feature gives the scientist performing the study unlimited flexibility in choosing and parametrising the outputs they find the most important. Here, we have elected to perform the sensitivity analysis using the model predictions in luminosity bins cast in terms of the difference between the computed and measured K-band luminosity function at $z = 0$. Our motivation for this was that by choosing an established observable with a well understood connection to the underlying physical processes and their description in term of

GALFORM parameters, we could make a convincing case for the usefulness of the sensitivity analysis.

With this in mind, future work on SA might want to examine the variance of the outputs of the semi-analytic model alone, independently of the corresponding measured observable values. There are three main reasons for such an approach: i) using the full dynamic range of the predictions: normalising the model output by observations flattens the dynamic range, and while SA works equally well for small and large values, by only analysing a flat version of the model predictions we effectively take the regions in which the model gives a flat or steep response (for instance, the faint and bright end of the LF respectively) and make them look the same. ii) independence of post-processing: by comparing to data, we had to make a choice about the norm of the discrepancy between the model output and observations: do we retain the sense of the discrepancy or square it? A different SA study could have chosen differently, altering the results. By analysing model outputs independently of the observations these choices are no longer necessary. iii) data independence: SA results could change if a different dataset is used with the same model.

Moreover, the K-band luminosity function is just one possible output and there are many others which a successful semi-analytic model should reproduce accurately. Analysing all of these is outside the scope of this study, but we hope to have shown that SA is a promising avenue of research.

Finally, we note that while correctly estimating model sensitivity can be useful in guiding model optimisation and improving the physical interpretation of the parameters of the galaxy formation models, one must remember that even the most rigorous sensitivity analysis can only provide the answers with regards to the model, not the underlying physical system itself (Taleb & Douady, 2013). Therefore, the relationship between the structure of the model and that of the physical system remains open to discussion.

Chapter 6

Conclusions & Future Work

6.1 Summary

Cosmology's rise as a quantitative science in recent decades has coincided with the emergence of ΛCDM as the concordant cosmological model. Together with GR, it has passed many observational tests, and provided explanations and predictions for numerous astronomical observations. However, there remain observational tensions on small scales, which could be solved by either changing the dark matter paradigm, or invoking baryonic physics. The cosmological constant remains somewhat unpalatable from a theoretical viewpoint, hence the motivation to look at alternatives to ΛCDM .

In this thesis we have presented a body of work on the cosmologies in the presence of modified gravitational field, specifically one postulated by the $f(R)$ theory.

In the first part of the thesis we extended the work on the assembly of haloes in ΛCDM to modified gravity by Ludlow et al. (2013, 2016), in which they argued that the internal structure of dark matter haloes could be connected to their formation histories. This is almost the case in modified gravity, though the relation between different definitions of the formation time is different in $f(R)$ models than it is in GR. Conversely, this difference could be used as a test of the nature of gravity. The requirements on the mass faction of the progenitors used to define the formation time means that, in practice, the halo concentration can only be estimated directly from the merger histories for very well resolved haloes. In any general galaxy catalogue,

most galaxies are predicted to be found in haloes that are far below satisfying the mass condition.

We then incorporated the changes in the properties of dark matter haloes in modified gravity into the GALFORM model. This represents a first, albeit incomplete (see next section), attempt to extend GALFORM beyond GR. Previous work on semi-analytic galaxy formation models in non-standard cosmologies (e.g. Fontanot et al. (2013)) has looked at the impact of modified gravity on galaxy properties that are shaped by a wide range of halo masses, such as the galaxy luminosity function. We decided to focus instead on the halo occupation distribution function of Hα -emitting galaxies. Previous work with GALFORM in GR (Gonzalez-Perez et al., 2018) predicted a peak in the HOD of OII emitters around the mass at which screening is expected to become ineffective, and at which we might expect to see a clearer difference between modified gravity and GR. The changes were nevertheless small and possibly out of the reach of even the next generation of surveys.

Finally, we applied a new technique, sensitivity analysis (SA), to address the question of which parameters matter most for a particular galaxy formation model prediction. The idea behind this project came out of a placement with Atom Bank, the Durham-based first digital-only bank in the UK. The placement was one of the first carried out through the Centre for Doctoral Training in Data Intensive Science. The focus of the placement with Atom Bank was to use the SA to address an audit item on the bank's model for the pricing of mortgages. Much like GALFORM, this was a complex, multi-parameter model. The SA analysis of GALFORM was applied to the K-band luminosity function predicted by GALFORM, allowing us to confirm our physical intuition about the relative importance of parameters, and how it varies with luminosity. This study marks moves away from a "one-at-a-time" parameter variation, improving the transparency of GALFORM results and simplifying interpretation of GALFORM parameters.

6.2 Future Work

Our investigation into the concentration-formation time relation of haloes in modified gravity concluded that the origin of the free parameter of the relation shown in Ludlow et al. (2016) for GR haloes remains unknown, and that the relation in general is not reproduced for $f(R)$ haloes. Therefore, the question about the origin of the connection between haloes' concentrations and formation times remains open. Future studies of modified gravity simulations present a useful avenue of progress, since the changes they introduce make the counterpart GR simulations act as null hypotheses. In particular, running $f(R)$ comparison studies for the F5 and F4 models, and comparing how the concentration-formation time relations react to halo population splitting and parameter variation would be an interesting next step towards understanding what causes the relation to be so well defined in GR in the first place. Furthermore, theoretical studies on defining an environmental proxy able to recover the relation would surely be of interest not only to research of modified gravity, but of halo formation research as well. Finally, while Ludlow et al. (2016) improves upon the original Navarro et al. (1996) formation time definition, the new values are still degenerate in $f(R)$ for different halo mass histories; a revision of halo formation time definition would, very much like a revision of halo concentration definition (Mitchell et al., 2019), be of tremendous value to further modified gravity research.

Here we took the first steps towards a modified gravity version of GALFORM. If we are to fully exploit the measurements from the next generation of galaxy surveys, it is important that we build the best models we can of galaxy formation in gravity models beyond GR. Following our incorporation of the halo merger histories, concentration-mass relation, and the virial velocity-mass relations into galaxy, the next steps towards a full fledged MG GALFORM would be to (i) recalculate hydrostatic equilibrium, to account for the impact of modified gravity on gas cooling, (ii) revise the treatment of galaxy sizes, to take into account MG, and (iii) use stellar population synthesis models that take modified gravity into account for stellar evolution (e.g. Davis et al. (2012); Sakstein (2015)).

We also applied the first sensitivity analysis to a galaxy formation model that attempts to predict the full galaxy population. Lack of transparency is a charge

often levelled at the way in which parameters are set in semi-analytical models, and the SA can partially address it. This study was carried out on a well-studied observable, since both GALFORM and SA itself were being trialed. However, we have managed to successfully reproduce and quantify physical intuition regarding many GALFORM parameters, especially those related to feedback processes. We therefore find SA in general, and variance-based approach in particular, to be a new and promising avenue of research for modelling complex astrophysical processes. We would like to see analyses of models to widen their scope, and to include the sensitivity indices of model parameters alongside their best-fitting values as valuable metric of model performance. Current limitations of SA techniques include the fact that the analysis can only be carried out for one output at a time, e.g. one luminosity bin in the luminosity function, and does not give any indication of what the best-fitting model parameters are any better than a grid search method. Moreover, further research is required to correctly model correlated outputs, as these are currently assumed to be statistically independent variables. Finally, the required number of model evaluation grows quickly for a multi-parameter study, which makes alternative ideas, such as improved sampling or emulation techniques very compelling.

Bibliography

Akaike H., 1974, IEEE Transactions on Automatic Control, 19, 716

Allen S. W., Evrard A. E., Mantz A. B., 2011, Annual Review of Astronomy and Astrophysics, 49, 409

Amendola L., et al., 2013, Living Reviews in Relativity, 16

Angel P. W., Poole G. B., Ludlow A. D., Duffy A. R., Geil P. M., Mutch S. J., Mesinger A., Wyithe J. S. B., 2016,] 10.1093/mnras/stw737, 459, 2106

Antonov I., Saleev V., 1979, USSR Computational Mathematics and Mathematical Physics, 19, 252

Arnold C., Leo M., Li B., 2019a, Nature Astronomy

Arnold C., Leo M., Li B., 2019b, Nature Astronomy, p. 424

Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101

Baugh C. M., 2013, Publications of the Astronomical Society of Australia, 30, e030

Baugh C. M., et al., 2019, Monthly Notices of the Royal Astronomical Society, 483, 4922

Benson A. J., 2010, Physics Reports, 495, 33

Benson A. J., 2014, Monthly Notices of the Royal Astronomical Society, 444, 2599

Benson A. J., Bower R., 2010, Monthly Notices of the Royal Astronomical Society, 405, 15731623

Berlind A. A., et al., 2003, The Astrophysical Journal, 593, 1

Berry C. P. L., Gair J. R., 2011, Physical Review D, 83, 104022

Bett P., Eke V., Frenk C. S., Jenkins A., Helly J., Navarro J., 2007, Monthly Notices of the Royal Astronomical Society, 376, 215

Bigiel F., et al., 2011, The Astrophysical Journal, 730, L13

Blas D., Lesgourgues J., Tram T., 2011, Journal of Cosmology and Astroparticle Physics, 2011, 034

Blitz L., Rosolowsky E., 2006, The Astrophysical Journal, 650, 933

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, The Astrophysical Journal, 379, 440

Bose S., Hellwing W. A., Li B., 2015, Journal of Cosmology and Astroparticle Physics, 2015, 034

Bose S., Hellwing W. A., Frenk C. S., Jenkins A., Lovell M. R., Helly J. C., Li B., 2016, Monthly Notices of the Royal Astronomical Society, 455, 318

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, Monthly Notices of the Royal Astronomical Society, 370, 645

Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, Monthly Notices of the Royal Astronomical Society, 407, 2017

Buchdahl H. A., 1970, Monthly Notices of the Royal Astronomical Society, 150, 1

Bullock J. S., Boylan-Kolchin M., 2017, Annual Review of Astronomy and Astrophysics, 55, 343

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, Monthly Notices of the Royal Astronomical Society, 321, 559

Carroll S. M., Duvvuri V., Trodden M., Turner M. S., 2004, Physical Review D, 70, 043528

Cataneo M., et al., 2015, Physical Review D, 92, 044009

Chan K., Saltelli A., Tarantola S., 1997, in Proceedings of the 29th conference on Winter simulation - WSC '97. ACM Press, doi:10.1145/268437.268489, `https://doi.org/10.1145%2F268437.268489`

Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Physics Reports, 513, 1

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, Monthly Notices of the Royal Astronomical Society, 319, 168

Contreras S., Baugh C. M., Norberg P., Padilla N., 2013, Monthly Notices of the Royal Astronomical Society, 432, 2717

Contreras S., Zehavi I., Padilla N., Baugh C. M., Jiménez E., Lacerna I., 2019, Monthly Notices of the Royal Astronomical Society, 484, 1133

Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, Monthly Notices of the Royal Astronomical Society, 452, 1217

Crain R. A., et al., 2015, Monthly Notices of the Royal Astronomical Society, 450, 1937

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, The Astrophysical Journal, 292, 371

Davis A.-C., Lim E. A., Sakstein J., Shaw D. J., 2012, Physical Review D, 85

De Lucia G., Blaizot J., 2007, Monthly Notices of the Royal Astronomical Society, 375, 2

Diemer B., Joyce M., 2019, The Astrophysical Journal, 871, 168

Diemer B., Kravtsov A. V., 2015, The Astrophysical Journal, 799, 108

Driver S. P., et al., 2012, Monthly Notices of the Royal Astronomical Society, 427, 3244

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008,] 10.1111/j.1745-3933.2008.00537.x, 390, L64

Efstathiou G., Lake G., Negroponte J., 1982, Monthly Notices of the Royal Astronomical Society, 199, 1069

Einasto J., 1965. Trudy Astrophys. Inst., Alma-Ata

Einstein A., et al., 1916, Annalen der Physik, 49, 769

Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, Monthly Notices of the Royal Astronomical Society, 410, 53

Fisher R. A., 1918, Transactions of the Royal Society of Edinburgh, 52, 399433 (1918)

Fontanot F., Puchwein E., Springel V., Bianchi D., 2013, Monthly Notices of the Royal Astronomical Society, 436, 2672

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

Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, Monthly Notices of the Royal Astronomical Society, 387, 536

Genina A., et al., 2017, Monthly Notices of the Royal Astronomical Society, 474, 1398

Golovin D., Solnik B., Moitra S., Kochanski G., Karro J. E., Sculley D., eds, 2017, Google Vizier: A Service for Black-Box Optimization `http://www.kdd.org/kdd2017/papers/view/google-vizier-a-service-for-black-box-optimization`

Gómez F. A., Coleman-Smith C. E., O'Shea B. W., Tumlinson J., Wolpert R. L., 2012, The Astrophysical Journal, 760, 112

Gómez F. A., Coleman-Smith C. E., O'Shea B. W., Tumlinson J., Wolpert R. L., 2014, The Astrophysical Journal, 787, 20

Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, Monthly Notices of the Royal Astronomical Society, 439, 264

Gonzalez-Perez V., et al., 2018, Monthly Notices of the Royal Astronomical Society, 474, 4024

Green J., et al., 2012, arXiv:1208.4012 [astro-ph]

Griffin A. J., Lacey C. G., Gonzalez-Perez V., Lagos C. d. P., Baugh C. M., Fanidakis N., 2019, Monthly Notices of the Royal Astronomical Society, 487, 198

Gunn J. E., Gott III J. R., 1972, The Astrophysical Journal, 176, 1

Guo Q., et al., 2016, Monthly Notices of the Royal Astronomical Society, 461, 3457

Haas M. R., Schaye J., Jeeson-Daniel A., 2012, Monthly Notices of the Royal Astronomical Society, 419, 2133

Halton J. H., 1964, Communications of the ACM, 7, 701

Han J., Jing Y. P., Wang H., Wang W., 2012, Monthly Notices of the Royal Astronomical Society, 427, 2437

Han J., Cole S., Frenk C. S., Benitez-Llambay A., Helly J., 2018, Monthly Notices of the Royal Astronomical Society, 474, 604

He J.-h., Guzzo L., Li B., Baugh C. M., 2018, Nature Astronomy

Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, Monthly Notices of the Royal Astronomical Society, 396, 535

Herman J., Usher W., 2017, The Journal of Open Source Software, 2, 97

Hernández-Aguayo C., Baugh C. M., Li B., 2018, Monthly Notices of the Royal Astronomical Society, 479, 4824

Hinshaw G., et al., 2013, The Astrophysical Journal Supplement Series, 208, 19

Hou J., Lacey C. G., Frenk C. S., 2017, Monthly Notices of the Royal Astronomical Society, 475, 543

Hu W., Sawicki I., 2007, Physical Review D, 76

Hubble E., 1929, Proceedings of the national academy of sciences, 15, 168

Ishigami T., Homma T., 1991, in [1990] Proceedings. First International Symposium on Uncertainty Modeling and Analysis. IEEE Comput. Soc. Press, College Park, MD, USA, pp 398–403, doi:10.1109/ISUMA.1990.151285, `http://ieeexplore.ieee.org/document/151285/`

Jansen M. J., 1999, Computer Physics Communications, 117, 35

Jansen R. A., Franx M., Fabricant D., 2001, The Astrophysical Journal, 551, 825

Jenkins A., 2010, Monthly Notices of the Royal Astronomical Society, 403, 1859

Jenkins A., 2013, Monthly Notices of the Royal Astronomical Society, 434, 2094

Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, Monthly Notices of the Royal Astronomical Society, 440, 2115

Joyce A., Jain B., Khoury J., Trodden M., 2015, Physics Reports, 568, 1

Kaiser N., 1984, The Astrophysical Journal, 284, L9

Kauffmann G., Nusser A., Steinmetz M., 1997, Monthly Notices of the Royal Astronomical Society, 286, 795

Khoury J., Weltman A., 2004, Physical Review Letters, 93, 171104

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, Monthly Notices of the Royal Astronomical Society, 457, 4340

Knebe A., et al., 2015, Monthly Notices of the Royal Astronomical Society, 451, 4029

Koyama K., 2016, Reports on Progress in Physics, 79, 046902

Lacey C., Cole S., 1993, Monthly Notices of the Royal Astronomical Society, 262, 627

Lacey C. G., et al., 2016, Monthly Notices of the Royal Astronomical Society, 462, 3854

Lagos C. d. P., Lacey C. G., Baugh C. M., Bower R. G., Benson A. J., 2011, Monthly Notices of the Royal Astronomical Society, 416, 1566

Lagos C. d. P., Lacey C. G., Baugh C. M., 2013, Monthly Notices of the Royal Astronomical Society, 436, 1787

Laureijs R., et al., 2011, arXiv:1110.3193 [astro-ph]

Lemaître G., 1927, in Annales de la Société scientifique de Bruxelles. pp 49–59

Levi M., et al., 2013

Levitan Y., Markovich N., Rozin S., Sobol' I., 1988, USSR Computational Mathematics and Mathematical Physics, 28, 88

Li B., Mota D. F., Barrow J. D., 2011, The Astrophysical Journal, 728, 109

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

Lombriser L., 2014, Annalen der Physik, 526, 259

Lombriser L., Koyama K., Li B., 2014, Journal of Cosmology and Astroparticle Physics, 2014, 021

Lu Y., Mo H. J., Weinberg M. D., Katz N., 2011, Monthly Notices of the Royal Astronomical Society, 416, 1949

Lu Y., Mo H. J., Katz N., Weinberg M. D., 2012, Monthly Notices of the Royal Astronomical Society, 421, 1779

Ludlow A. D., Angulo R. E., 2016, Monthly Notices of the Royal Astronomical Society: Letters, 465, L84

Ludlow A. D., Navarro J. F., Springel V., Vogelsberger M., Wang J., White S. D. M., Jenkins A., Frenk C. S., 2010, Monthly Notices of the Royal Astronomical Society, 406, 137

Ludlow A. D., et al., 2013, Monthly Notices of the Royal Astronomical Society, 432, 1103

Ludlow A. D., Navarro J. F., Angulo R. E., Boylan-Kolchin M., Springel V., Frenk C., White S. D. M., 2014a, Monthly Notices of the Royal Astronomical Society, 441, 378

Ludlow A. D., Borzyszkowski M., Porciani C., 2014b,] 10.1093/mnras/stu2021, 445, 4110

Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, Monthly Notices of the Royal Astronomical Society, 460, 1214

Ludlow A. D., Schaye J., Bower R., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3663

Maccio A. V., Dutton A. A., Van Den Bosch F. C., Moore B., Potter D., Stadel J., 2007, Monthly Notices of the Royal Astronomical Society, 378, 55

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, Monthly Notices of the Royal Astronomical Society, 391, 1940

Macciò A. V., Ruchayskiy O., Boyarsky A., Muñoz-Cuartas J. C., 2013,] 10.1093/mnras/sts078, 428, 882

Markevitch M., Gonzalez A. H., Clowe D., Vikhlinin A., Forman W., Jones C., Murray S., Tucker W., 2004, The Astrophysical Journal, 606, 819

Martindale H., Thomas P. A., Henriques B. M., Loveday J., 2017, Monthly Notices of the Royal Astronomical Society, 472, 1981

Massey R., et al., 2015, Monthly Notices of the Royal Astronomical Society, 449, 3393

Mathewson D. S., Ford V. L., Buchhorn M., 1992, The Astrophysical Journal Supplement Series, 81, 413

Merritt D., Graham A. W., Moore B., Diemand J., Terzić B., 2006, The Astronomical Journal, 132, 2685

Merson A., Smith A., Benson A., Wang Y., Baugh C., 2019, arXiv:1903.02030 [astro-ph]

Mitchell P. D., Lacey C. G., Baugh C. M., Cole S., 2013, Monthly Notices of the Royal Astronomical Society, 435, 87

Mitchell P. D., et al., 2018, Monthly Notices of the Royal Astronomical Society, 474, 492

Mitchell M. A., Arnold C., He J.-h., Li B., 2019, arXiv:1901.06392 [astro-ph]

Mo H. J., White S. D. M., 1996, Monthly Notices of the Royal Astronomical Society, 282, 347

Mo H., van den Bosch F., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press, `http://www.astro.yale.edu/vdbosch/book.html`

Moore B., Governato F., Quinn T., Stadel J., Lake G., 1998, The Astrophysical Journal Letters, 499, L5

Morris M. D., 1991, Technometrics, 33, 161

Mutch S. J., Poole G. B., Croton D. J., 2013, Monthly Notices of the Royal Astronomical Society, 428, 2001

Navarro J. F., Frenk C. S., White S. D. M., 1996, The Astrophysical Journal, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, The Astrophysical Journal, 490, 493

Neto A. F., et al., 2007, Monthly Notices of the Royal Astronomical Society, 381, 1450

Nunes R. C., Pan S., Saridakis E. N., Abreu E. M. C., 2017, Journal of Cosmology and Astroparticle Physics, 2017, 005

Oleśkiewicz P., Baugh C. M., Ludlow A., 2019, arXiv preprint arXiv:1903.07608

Ostriker J. P., McKee C. F., 1988, Reviews of Modern Physics, 60, 1

Parkinson H., Cole S., Helly J., 2007, Monthly Notices of the Royal Astronomical Society, 383, 557

Peccei R. D., Quinn H. R., 1977, Physical Review Letters, 38, 1440

Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Series in Physics, Princeton University Press, Princeton, N.J

Peebles P. J. E., 1982, The Astrophysical Journal, 263, L1

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

Pillepich A., et al., 2018, Monthly Notices of the Royal Astronomical Society, 473, 4077

Planck Collaboration et al., 2014, Astronomy & Astrophysics, 571, A16

Planck Collaboration et al., 2016a, Astronomy & Astrophysics, 594, A1

Planck Collaboration et al., 2016b, Astronomy & Astrophysics, 594, A13

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, Monthly Notices of the Royal Astronomical Society, 338, 14

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012,] 10.1111/j.1365-2966.2012.21007.x, 423, 3018

Press W. H., Schechter P., 1974, The Astrophysical Journal, 187, 425

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2007, Numerical Recipes: The Art of Scientific Computing, 3. ed. edn. Cambridge Univ. Press, Cambridge

Puchwein E., Baldi M., Springel V., 2013, Monthly Notices of the Royal Astronomical Society, 436, 348

Rachen J. P., Gahlings U. G., 2013, Journal of Comparative Irrelevance (Letters), 23, 966

Refregier A., 2003, Annual Review of Astronomy and Astrophysics, 41, 645

Riess A. G., et al., 1998, The Astrophysical Journal, 116, 1009

Robertson A., Massey R., Eke V., 2016, Monthly Notices of the Royal Astronomical Society, 465, 569

Rodrigues L. F. S., Vernon I., Bower R. G., 2017, Monthly Notices of the Royal Astronomical Society, 466, 2418

Rubin V. C., W. Kent J. F., 1970, The Astrophysical Journal, 159, 379

Ruiz A. N., et al., 2015, The Astrophysical Journal, 801, 139

Sakstein J., 2015, Physical Review Letters, 115

Salpeter E. E., 1955, The Astrophysical Journal, 121, 161

Saltelli A., Annoni P., Azzini I., Campolongo F., Ratto M., Tarantola S., 2010, Computer Physics Communications, 181, 259

Saltelli A., Aleksankina K., Becker W., Fennell P., Ferretti F., Holst N., Li S., Wu Q., 2019, Environmental Modelling & Software, 114, 29

Schaye J., et al., 2014, Monthly Notices of the Royal Astronomical Society, 446, 521

Schmidt F., Lima M., Oyaizu H., Hu W., 2009, Physical Review D, 79

Schwarz G., 1978, The Annals of Statistics, 6, 461

Shi D., Li B., Han J., Gao L., Hellwing W. A., 2015, Monthly Notices of the Royal Astronomical Society, 452, 3179

Shi D., Li B., Han J., 2017, Monthly Notices of the Royal Astronomical Society, 469, 705

Silva L., Granato G. L., Bressan A., Danese L., 1998, The Astrophysical Journal, 509, 103

Slipher V. M., 1915, Popular astronomy, 23, 21

Smith A., Cole S., Baugh C., Zheng Z., Angulo R., Norberg P., Zehavi I., 2017, Monthly Notices of the Royal Astronomical Society, 470, 4646

Sobol' I. M., 1967, Ž. Vyčisl. Mat. i Mat. Fiz., 7, 784

Sobol' I. M., 1993, Math. Modeling Comput. Experiment, 1, 407

Sobol' I., 2001, Mathematics and Computers in Simulation, 55, 271

Somerville R. S., Davé R., 2015, Annual Review of Astronomy and Astrophysics, 53, 51

Springel V., 2005, Monthly Notices of the Royal Astronomical Society, 364, 1105

Springel V., 2010, Monthly Notices of the Royal Astronomical Society, 401, 791

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, Monthly Notices of the Royal Astronomical Society, 328, 726

Springel V., et al., 2005, Nature, 435, 629

Stasinska G., 1990, Astronomy and Astrophysics, 83, 501

Stein M., 1987, Technometrics, 29, 143

Taleb N. N., Douady R., 2013, Quantitative Finance, 13, 1677

Terukina A., Lombriser L., Yamamoto K., Bacon D., Koyama K., Nichol R. C., 2014, Journal of Cosmology and Astroparticle Physics, 2014, 013

Teyssier R., 2002, Astronomy & Astrophysics, 385, 337

Thomas P. A., Muanwong O., Pearce F. R., Couchman H. M. P., Edge A. C., Jenkins A., Onuora L., 2001, Monthly Notices of the Royal Astronomical Society, 324, 450

Thomas P. A., et al., 2015, arXiv:1508.05388 [astro-ph]

Tumlinson J., 2009, The Astrophysical Journal, 708, 1398

Ulam S., 1960, Problems in Modern Mathematics. John Wiley & Sons, New York

Upadhye A., Steffen J. H., 2013, arXiv:1306.6113 [astro-ph]

Vale A., Ostriker J. P., 2004, Monthly Notices of the Royal Astronomical Society, 353, 189

Vale A., Ostriker J. P., 2006, Monthly Notices of the Royal Astronomical Society, 371, 1173

Vogelsberger M., et al., 2014, Monthly Notices of the Royal Astronomical Society, 444, 1518

Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, The Astrophysical Journal, 568, 52

Weinberg D. H., Bullock J. S., Governato F., de Naray R. K., Peter A. H. G., 2015, Proceedings of the National Academy of Sciences, 112, 12249

White S. D. M., Rees M. J., 1978, Monthly Notices of the Royal Astronomical Society, 183, 341

Zehavi I., et al., 2011, The Astrophysical Journal, 736, 59

Zhao D. H., Jing Y. P., Mo H. J., Brner G., 2003, The Astrophysical Journal, 597, L9

Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2009, The Astrophysical Journal, 707, 354

Zhao G.-B., Li B., Koyama K., 2011, Physical Review Letters, 107

Zheng Z., et al., 2005, The Astrophysical Journal, 633, 791

Zuntz J., Zlosnik T. G., Bourliot F., Ferreira P. G., Starkman G. D., 2010, Physical Review D, 81

Zwicky F., 1933, Helvetica Physica Acta, 6, 110