

# Durham E-Theses

---

## *A general framework for unbiased tests of gravity using galaxy clusters*

MITCHELL, MYLES,ARTHUR

---

### How to cite:

MITCHELL, MYLES,ARTHUR (2021) *A general framework for unbiased tests of gravity using galaxy clusters*, Durham theses, Durham University. Available at Durham E-Theses Online:  
<http://etheses.dur.ac.uk/14147/>

---

### Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

# A general framework for unbiased tests of gravity using galaxy clusters

Myles Arthur Mitchell

A thesis presented for the degree of  
Doctor of Philosophy



Institute for Computational Cosmology

Department of Physics

Durham University

United Kingdom

October 2021

# A general framework for unbiased tests of gravity using galaxy clusters

Myles Arthur Mitchell

## Abstract

We present a Markov chain Monte Carlo pipeline which can be used for unbiased large-scale tests of gravity using galaxy cluster observations. The pipeline, which currently uses cluster number counts to constrain the present-day background scalar field  $f_{R0}$  of Hu-Sawicki  $f(R)$  gravity, fully accounts for the effects of the fifth force on cluster properties including the dynamical mass, the halo concentration and the observable-mass scaling relations. This is achieved using general models which have been calibrated over a wide and continuous mass range ( $10^{11}M_{\odot} \lesssim M \lesssim 10^{15}M_{\odot}$ ) using a large suite of cosmological simulations, including the first to simultaneously incorporate both screened modified gravity and full baryonic physics. We show, using mock cluster catalogues, that an incomplete treatment of the observable-mass scaling relations in  $f(R)$  gravity, which do not necessarily follow the usual power-law behaviour, can lead to unbiased and imprecise constraints. It is therefore essential to fully account for these effects in future cosmological tests of gravity that will make use of vast cluster catalogues from ongoing and upcoming galaxy surveys. Our constraint framework can be easily extended to other gravity models; to demonstrate this, we have carried out a similar modelling of cluster properties in the normal-branch Dvali-Gabadadze-Porrati model (nDGP), which features a very different screening mechanism. Using our full-physics simulations, we also study the angular power spectra of the thermal and kinetic Sunyaev-Zel'dovich effects in  $f(R)$  gravity and nDGP, and demonstrate the potential for precise constraints of gravity using data from upcoming CMB experiments. Finally, we present a retuned baryonic physics model, based on the IllustrisTNG model, which can be used for full-physics simulations within large cosmological volumes. This can be used to study the properties of galaxy groups and clusters in screened modified gravity over the mass range  $10^{13}M_{\odot} \leq M \lesssim 10^{15}M_{\odot}$ .

Supervisors: Prof. Baojiu Li, Dr. Christian Arnold and Prof. Carlton M. Baugh

---

# Declaration

The research described in this thesis was carried out between 2017 and 2021 while the author was a research student under the supervision of Prof. Baojiu Li, Dr. Christian Arnold and Prof. Carlton M. Baugh at the Institute for Computational Cosmology in the Department of Physics, Durham University. No part of this work has been submitted for any other degree or qualification at Durham University or any other university.

The contents of this thesis have appeared in the following papers:

- **Mitchell M. A.**, He J.-h., Arnold C., Li B., 2018, A general framework to test gravity using galaxy clusters I: Modelling the dynamical mass of haloes in  $f(R)$  gravity, *MNRAS*, **477**, 1133 (**Chapter 2**)
- **Mitchell M. A.**, Arnold C., He J.-h., Li B., 2019, A general framework to test gravity using galaxy clusters II: A universal model for the halo concentration in  $f(R)$  gravity, *MNRAS*, **487**, 1410 (**Chapter 3**)
- **Mitchell M. A.**, Arnold C., Hernández-Aguayo C., Li B., 2021, The impact of modified gravity on the Sunyaev-Zel'dovich effect, *MNRAS*, **501**, 4565 (**Chapter 7**)
- **Mitchell M. A.**, Arnold C., Li B., 2021, A general framework to test gravity using galaxy clusters III: Observable-mass scaling relations in  $f(R)$  gravity, *MNRAS*, **502**, 6101 (**Chapter 4, sections 2 to 5**)
- **Mitchell M. A.**, Hernández-Aguayo C., Arnold C., Li B., 2021, A general framework to test gravity using galaxy clusters IV: Cluster and halo properties in DGP gravity, *accepted by MNRAS*, arXiv:2106.13815 [astro-ph.CO] (**Chapter 6, sections 2 to 4**)
- **Mitchell M. A.**, Arnold C., Li B., 2021, A general framework to test gravity using galaxy clusters V: A self-consistent pipeline for unbiased constraints of  $f(R)$  gravity, *accepted by MNRAS*, arXiv:2107.14224 [astro-ph.CO] (**Chapter 5, sections 2 to 6**)

- **Mitchell M. A.**, Arnold C., Li B., 2021, A general framework to test gravity using galaxy clusters VI: Realistic galaxy formation simulations to study clusters in modified gravity, *submitted to MNRAS*, arXiv:2109.01147 [astro-ph.CO] (**Chapter 8**)

The results presented in this thesis rely on cosmological simulations which have either been run by the author (the simulations presented in Chapter 8) or by collaborators (the simulations described in Chapters 2-7).

**Copyright © 2021 by Myles Arthur Mitchell.**

*The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.*

---

# Acknowledgements

First of all, I am grateful to my supervisory team for giving me the opportunity to be involved in such fascinating and exciting projects during my PhD. I would like to thank Baojiu and Christian for always being so supportive and enthusiastic about my research- it has been a pleasure working with both of you. I also thank Carlton for getting me involved in the CDT programme, which I have thoroughly enjoyed being a part of.

I am also grateful to my collaborators: Jianhua He for his help and guidance during the early stages of my project, and César Hernández-Aguayo for letting me use his simulations. Thank you both for contributing to the work presented in this thesis. I also thank Matteo Cataneo, Weiguang Cui, Lucas Lombriser and Ian McCarthy for their helpful comments and discussions. In particular, I thank the COSMA support team for always being on hand to help.

Last but not least, I am grateful to all of the friends that I have made throughout my time at Durham, particularly my office mates and everyone involved in the CDT programme. A special mention goes to Giorgio- I will miss our afternoon walks and lunchtime banter.

I was supported by a studentship with the Durham Centre for Doctoral Training in Data Intensive Science, funded by the UK Science and Technology Facilities Council (STFC, ST/P006744/1) and Durham University. This thesis used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

---

# Contents

<b>Declaration</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Tables</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 $\Lambda$ CDM: the standard model of cosmology . . . . .	1
1.1.1 Cosmic expansion in $\Lambda$ CDM . . . . .	2
1.1.2 Strengths and weakness of $\Lambda$ CDM . . . . .	3
1.2 Modifications to General Relativity . . . . .	5
1.2.1 $f(R)$ gravity . . . . .	5
1.2.2 Dvali-Gabadadze-Porrati model . . . . .	8
1.3 Testing gravity with large-scale structure . . . . .	10
1.3.1 Previous work . . . . .	11
1.3.2 Current and future observations . . . . .	12
1.3.3 $N$ -body and hydrodynamical simulations . . . . .	14
1.3.4 Markov chain Monte Carlo sampling . . . . .	17
1.4 Thesis overview . . . . .	18

<b>2</b>	<b>Modelling the dynamical mass of haloes in <math>f(R)</math> gravity</b>	<b>21</b>
2.1	Introduction . . . . .	21
2.2	Background . . . . .	24
2.2.1	Thin-shell model . . . . .	24
2.2.2	Dynamical mass and effective mass . . . . .	27
2.3	A framework for gravity tests using clusters . . . . .	28
2.3.1	Cluster abundance tests . . . . .	30
2.3.1.1	Halo abundance . . . . .	32
2.3.1.2	Scaling relations in $f(R)$ gravity . . . . .	33
2.3.1.3	Other issues . . . . .	35
2.3.2	Other observables . . . . .	37
2.4	Simulations and methods . . . . .	38
2.4.1	Simulations . . . . .	38
2.4.2	Measuring the dynamical mass enhancement . . . . .	40
2.4.2.1	tanh function fit to $M_{\text{dyn}}/M_{\text{true}}$ . . . . .	41
2.4.2.2	Fitting of $p_1, p_2$ as functions of $\bar{f}_R(z)/(1+z)$ . . . . .	47
2.5	Results . . . . .	48
2.5.1	Potential implications . . . . .	53
2.6	Summary, discussion and conclusions . . . . .	54
<b>3</b>	<b>A universal model for the halo concentration in <math>f(R)</math> gravity</b>	<b>58</b>
3.1	Introduction . . . . .	58
3.2	Simulations and methods . . . . .	61
3.2.1	Simulations . . . . .	61
3.2.2	Methods . . . . .	63
3.2.2.1	Concentration measurement . . . . .	64
3.2.2.2	Rescaled mass . . . . .	66
3.2.2.3	Concentration enhancement . . . . .	69
3.3	Results . . . . .	73
3.3.1	Concentration measurement comparison . . . . .	73



3.3.2	General model for the concentration enhancement . . . . .	77
3.4	Summary, Discussion and Conclusions . . . . .	82
<b>4</b>	<b>Observable-mass scaling relations in <math>f(R)</math> gravity</b>	<b>85</b>
4.1	Introduction . . . . .	85
4.2	Background . . . . .	86
4.2.1	Effective density approach . . . . .	87
4.2.2	True density approach . . . . .	91
4.3	Simulations and methods . . . . .	92
4.3.1	Simulations . . . . .	93
4.3.2	Group catalogues . . . . .	93
4.4	Results . . . . .	95
4.4.1	Scaling relations . . . . .	95
4.4.1.1	Temperature scaling relations . . . . .	96
4.4.1.2	$Y_{\text{SZ}}$ and $Y_X$ scaling relations . . . . .	100
4.4.1.3	X-ray luminosity scaling relations . . . . .	102
4.4.1.4	Further comments . . . . .	104
4.4.2	Halo mass ratio calculation . . . . .	105
4.4.3	$Y_X$ -temperature scaling relation . . . . .	107
4.5	Summary, Discussion and Conclusions . . . . .	108
<b>5</b>	<b>A self-consistent pipeline for unbiased constraints of <math>f(R)</math> gravity</b>	<b>111</b>
5.1	Introduction . . . . .	111
5.2	Background . . . . .	112
5.2.1	Dynamical mass enhancement and its scatters . . . . .	112
5.2.2	Halo concentration . . . . .	114
5.2.3	Observable-mass scaling relations . . . . .	114
5.2.4	Halo mass function . . . . .	114
5.3	Methods . . . . .	117
5.3.1	Theoretical mass function . . . . .	117

5.3.2	Observable-mass scaling relation . . . . .	119
5.3.3	Mock catalogues . . . . .	120
5.3.4	MCMC sampling . . . . .	123
5.4	Results . . . . .	125
5.4.1	GR pipeline . . . . .	125
5.4.2	$f(R)$ pipeline . . . . .	128
5.5	Potential biases in model constraints . . . . .	132
5.5.1	Constraint pipeline . . . . .	133
5.5.1.1	Power-law scaling relation . . . . .	133
5.5.1.2	Mass ratio scatter . . . . .	134
5.5.2	Cluster sample . . . . .	135
5.5.2.1	Flux threshold . . . . .	135
5.5.2.2	Sample variance . . . . .	138
5.5.3	Tighter priors . . . . .	138
5.6	Summary, Discussion and Conclusions . . . . .	141
<b>6</b>	<b>Cluster and halo properties in DGP gravity</b>	<b>144</b>
6.1	Introduction . . . . .	144
6.2	Simulations and methods . . . . .	145
6.2.1	Simulations . . . . .	145
6.2.2	Halo catalogues . . . . .	147
6.3	Results . . . . .	148
6.3.1	Observable-mass scaling relations . . . . .	148
6.3.2	Concentration-mass relation . . . . .	154
6.3.2.1	Dark-matter-only concentration . . . . .	154
6.3.2.2	Full-physics concentration . . . . .	159
6.3.2.3	Modelling the concentration in nDGP . . . . .	160
6.3.3	Halo mass function . . . . .	162
6.4	Summary, Discussion and Conclusions . . . . .	165

<b>7</b>	<b>The impact of modified gravity on the Sunyaev-Zel'dovich effect</b>	<b>169</b>
7.1	Introduction . . . . .	169
7.2	Simulations and methods . . . . .	171
7.2.1	Simulations . . . . .	171
7.2.2	SZ maps . . . . .	171
7.3	Results . . . . .	176
7.3.1	Electron pressure profiles . . . . .	177
7.3.2	tSZ and kSZ power spectra . . . . .	180
7.3.3	Transverse momentum power spectrum . . . . .	186
7.4	Summary, Discussion and Conclusions . . . . .	189
<b>8</b>	<b>Realistic galaxy formation simulations to study clusters in modified gravity</b>	<b>192</b>
8.1	Introduction . . . . .	192
8.2	Simulations and methods . . . . .	194
8.2.1	Baryonic physics fine-tuning . . . . .	194
8.2.1.1	Gravitational softening . . . . .	196
8.2.1.2	Stellar feedback . . . . .	197
8.2.1.3	Star formation model . . . . .	198
8.2.1.4	Black hole feedback . . . . .	199
8.2.1.5	Summary and further comments . . . . .	201
8.2.2	Large-box simulation . . . . .	202
8.3	Results . . . . .	203
8.3.1	Observable-mass scaling relations . . . . .	203
8.3.1.1	Temperature scaling relation . . . . .	205
8.3.1.2	$Y_{\text{SZ}}$ and $Y_{\text{X}}$ scaling relations . . . . .	206
8.3.1.3	X-ray luminosity scaling relation . . . . .	208
8.3.2	$Y_{\text{X}}$ -temperature and $L_{\text{X}}$ -temperature relations . . . . .	210
8.4	Summary . . . . .	211

<b>9</b>	<b>Conclusions and outlook</b>	<b>214</b>
9.1	Thesis summary . . . . .	214
9.1.1	Modelling the dynamical mass enhancement in $f(R)$ gravity . . . . .	214
9.1.2	Universal model for the halo concentration in $f(R)$ gravity . . . . .	216
9.1.3	Observable-mass scaling relations in $f(R)$ gravity . . . . .	218
9.1.4	Constraint pipeline for unbiased $f(R)$ cluster constraints . . . . .	220
9.1.5	Cluster and halo properties in nDGP . . . . .	221
9.1.6	Sunyaev-Zel'dovich effect in $f(R)$ gravity and nDGP . . . . .	223
9.1.7	Realistic simulations to study clusters in modified gravity . . . . .	225
9.2	Outlook and future work . . . . .	226
<b>Appendix A</b>	<b>Modelling the dynamical mass of haloes in <math>f(R)</math> gravity</b>	<b>229</b>
A.1	Weighted fitting of tanh curve . . . . .	229
A.2	Consistency tests . . . . .	234
<b>Appendix B</b>	<b>A self-consistent pipeline for unbiased constraints of <math>f(R)</math> gravity</b>	<b>238</b>
B.1	Modelling the dynamical mass scatter . . . . .	238
B.2	Mass conversions . . . . .	239
B.3	Test of the constraint pipeline on a stronger $f(R)$ model . . . . .	240
<b>Appendix C</b>	<b>Cluster and halo properties in DGP gravity</b>	<b>242</b>
C.1	Simulation consistency . . . . .	242
C.2	$M_{200}$ mass function . . . . .	245
<b>Appendix D</b>	<b>Realistic galaxy formation simulations to study clusters in modified gravity</b>	<b>247</b>
D.1	L100-N256 simulations . . . . .	247
D.2	L86-N256 simulations . . . . .	248
D.3	L68-N256 and L136-N512 simulations . . . . .	249
	<b>Bibliography</b>	<b>250</b>

---

# List of Figures

1.1	Absolute value of the background scalar field in HS $f(R)$ gravity as a function of redshift. . . . .	8
1.2	Flowchart outlining our general framework for testing gravity using galaxy clusters.	19
2.1	Flowchart outlining our pipeline for constraining $f(R)$ gravity using cluster number counts. . . . .	29
2.2	The ratio of the dynamical mass to the true mass as a function of the true mass for dark matter haloes in HS $f(R)$ gravity. . . . .	42
2.3	Dynamical mass to true mass ratio as a function of the true mass for several present-day field strengths $f_{R0}$ of HS $f(R)$ gravity at various redshifts. . . . .	45
2.4	Parameter $p_2$ in Eq. (2.17) as a function of the background scalar field at redshift $z$ , $\bar{f}_R(z)$ , divided by $(1+z)$ . . . . .	49
2.5	Parameter $p_1$ in Eq. (2.17) as a function of the background scalar field at redshift $z$ , $\bar{f}_R(z)$ , divided by $(1+z)$ . . . . .	51
3.1	Median density profiles of simulated dark matter haloes within four bins of rescaled mass $\log_{10}(M_{500}/10^{p_2})$ . . . . .	68
3.2	Enhancement of the halo concentration in three $f(R)$ models as a function of the halo mass, generated using AREPO data at various redshifts. . . . .	71
3.3	Enhancement of the halo concentration in three $f(R)$ models as a function of the halo mass, generated using ECOSMOG data at various redshifts. . . . .	72

---

3.4	Ratio of the median concentrations in $f(R)$ gravity and GR, computed using three different methods, as a function of the rescaled mass $\log_{10}(M_{500}/10^{p_2})$ . . . . .	74
3.5	Ratio of the median concentrations in $f(R)$ gravity and GR as a function of the rescaled mass, with colouring to indicate redshift, particle resolution, $ \bar{f}_R /(1+z)$ , and the goodness of the NFW fitting compared to GR. . . . .	76
3.6	Ratio of the median concentrations in $f(R)$ gravity and GR as a function of the rescaled mass $\log_{10}(M_{500}/10^{p_2})$ , plotted along with our best-fit universal model. . . . .	79
4.1	Median gas density and temperature profiles of haloes in $f(R)$ gravity and GR, using the full-physics and non-radiative SHYBONE simulations. . . . .	89
4.2	Mass-weighted temperature as a function of mass for haloes in $f(R)$ gravity and GR, using the full-physics and non-radiative SHYBONE simulations. . . . .	97
4.3	$Y_{\text{SZ}}$ parameter as a function of mass for haloes in $f(R)$ gravity and GR, using the full-physics and non-radiative SHYBONE simulations. . . . .	99
4.4	$Y_X$ parameter as a function of mass for haloes in $f(R)$ gravity and GR, using the full-physics and non-radiative SHYBONE simulations. . . . .	100
4.5	X-ray luminosity as a function of mass for haloes in $f(R)$ gravity and GR, using the full-physics and non-radiative SHYBONE simulations. . . . .	102
4.6	Dynamical mass enhancement in $f(R)$ gravity as a function of the true mass for haloes from the non-radiative and full-physics SHYBONE simulations. . . . .	105
4.7	Same as Fig. 4.4 but using Eq. (2.17) for the mapping between the $f(R)$ and GR scaling relations. . . . .	106
4.8	$Y_X$ parameter as a function of the mass-weighted temperature for haloes in $f(R)$ gravity and GR, using the full-physics and non-radiative SHYBONE simulations. . . . .	107
5.1	Root-mean-square scatter of the dynamical mass enhancement in $f(R)$ gravity as a function of the rescaled halo mass $\log_{10}(M_{500}/10^{p_2})$ . . . . .	113
5.2	Predicted halo mass function enhancement in $f(R)$ gravity with respect to GR. . . . .	116
5.3	Halo mass function in GR and F5 for spherical overdensities 500 and $300\Omega_M(z)$ . . . . .	118

5.4	SZ $Y$ -parameter as a function of the halo mass for clusters from a GR mock catalogue. . . . .	122
5.5	Parameter constraints using our GR pipeline for a $\Lambda$ CDM fiducial cosmology. .	126
5.6	Parameter constraints using our full $f(R)$ pipeline for both $\Lambda$ CDM and F5 fiducial cosmologies. . . . .	129
5.7	Parameter constraints from our simplified pipeline with the $f(R)$ effects on the observable-mass scaling relation neglected. . . . .	133
5.8	Effect of the dynamical mass enhancement scatter on the parameter constraints from our $f(R)$ pipeline. . . . .	135
5.9	Parameter constraints generated using our $f(R)$ pipeline for F5 mocks with different flux thresholds. . . . .	136
5.10	Parameter constraints generated using our $f(R)$ pipeline for GR mocks with different flux thresholds. . . . .	137
5.11	Parameter constraints generated by our $f(R)$ pipeline using different priors for $\alpha$ and $\Omega_M$ . . . . .	139
5.12	Predictions of the halo mass function in F6 and F5 for different combinations of values of $\Omega_M$ and $\sigma_8$ . . . . .	140
6.1	Matter power spectra in GR and four nDGP models, generated using the L500 AREPO simulation. . . . .	147
6.2	Mass-weighted temperature as a function of mass for haloes in nDGP and GR, using the full-physics SHYBONE simulations. . . . .	149
6.3	$Y_{SZ}$ parameter as a function of mass for haloes in nDGP and GR, using the full-physics SHYBONE simulations. . . . .	150
6.4	$Y_X$ parameter as a function of mass for haloes in nDGP and GR, using the full-physics SHYBONE simulations. . . . .	151
6.5	Median gas density and temperature profiles of haloes in nDGP and GR, using the full-physics SHYBONE simulations. . . . .	152
6.6	Halo concentration as a function of mass for haloes in GR and four nDGP models at redshifts 0, 1 and 2. . . . .	155

6.7	Halo concentration as a function of redshift for haloes in GR and four nDGP models, for three mass bins. . . . .	156
6.8	Median density profiles of haloes in GR and four nDGP models for a sample of mass bins and redshifts, generated using full-physics and dark-matter-only simulations. . . . .	158
6.9	Relative difference between the median halo concentrations in nDGP and GR as a function of mass, plotted along with the predictions from our general model.	160
6.10	Best-fit values of the parameters $A$ and $B$ of Eq. (6.1) as a function of the $H_0 r_c$ parameter of nDGP. . . . .	161
6.11	Halo mass function and its relative difference in nDGP with respect to GR at redshifts 0, 1 and 2, plotted along with our best-fit model. . . . .	163
7.1	Maps of the thermal SZ effect in GR and its relative enhancement in $f(R)$ gravity, generated using full-physics and non-radiative simulations. . . . .	174
7.2	Maps of the kinetic SZ effect in GR and its relative enhancement in $f(R)$ gravity, generated using full-physics and non-radiative simulations. . . . .	175
7.3	Map of the thermal SZ effect in GR and its relative enhancement in nDGP, generated using full-physics simulations. . . . .	176
7.4	Median electron pressure profiles for haloes in $f(R)$ gravity and GR for a selection of mass bins and redshifts, generated using full-physics and non-radiative simulations. . . . .	177
7.5	Median electron pressure profiles for haloes in nDGP and GR for a selection of mass bins and redshifts, generated using full-physics simulations. . . . .	178
7.6	Angular power spectra of the thermal and kinetic SZ effects in $f(R)$ gravity and GR, generated using full-physics and non-radiative simulations. . . . .	181
7.7	Angular power spectra of the thermal and kinetic SZ effects in nDGP and GR, generated using full-physics simulations. . . . .	184
7.8	Power spectrum of the transverse component of the electron momentum for $f(R)$ gravity and GR, generated using full-physics and non-radiative simulations. . .	186



7.9	Derivative of the GR kSZ angular power spectrum as a function of the cosmological scale factor for different values of the angular wavenumber. . . . .	187
8.1	Stellar, gas and black hole properties of haloes in a sample of the L68-N256 baryonic calibration runs. . . . .	195
8.2	Stellar, gas and black hole properties for F5 and GR haloes in the L302-N1136 full-physics simulations. . . . .	204
8.3	Mass-weighted temperature as a function of mass for F5 and GR haloes in the L302 full-physics simulations. . . . .	205
8.4	$Y_{\text{SZ}}$ parameter as a function of mass for F5 and GR haloes in the L302 full-physics simulations. . . . .	207
8.5	$Y_{\text{X}}$ parameter as a function of mass for F5 and GR haloes in the L302 full-physics simulations. . . . .	207
8.6	X-ray luminosity as a function of mass for F5 and GR haloes in the L302 full-physics simulations. . . . .	208
8.7	$Y_{\text{X}}$ parameter as a function of temperature for F5 and GR haloes in the L302 full-physics simulations. . . . .	209
8.8	X-ray luminosity as a function of temperature for F5 and GR haloes in the L302 full-physics simulations. . . . .	210
A.1	Dynamical mass enhancement for haloes in HS $f(R)$ gravity with weighted least squares fits of Eq. (2.17). . . . .	230
A.2	Parameter $p_2$ of Eq. (2.17), measured using weighted least squares fitting, as a function of $\bar{f}_R(z)/(1+z)$ . . . . .	231
A.3	Parameter $p_1$ of Eq. (2.17), measured using weighted least squares fitting, as a function of $\bar{f}_R(z)/(1+z)$ . . . . .	233
A.4	Comparison of measurements of the parameter $p_2$ of Eq. (2.17), using different simulations and datasets. . . . .	235
B.1	Parameter constraints using our full $f(R)$ pipeline for both F4.5 and F5 fiducial cosmologies. . . . .	241

C.1	Comparison of measurements of the median halo concentration in GR and N1 using simulations with different resolutions. . . . .	243
C.2	Comparison of measurements of the halo mass function in GR and N1 using simulations with different resolutions. . . . .	244
C.3	Halo mass function in nDGP and GR, measured and fitted using mass definition $M_{200}$ . . . . .	245

---

## List of Tables

2.1	Specifications of the ECOSMOG simulations used to study the ratio of the dynamical mass to the true mass. . . . .	39
3.1	Specifications of the ECOSMOG and AREPO simulation used to study the $f(R)$ halo concentration. . . . .	61
3.2	Optimal parameter values and errors from the fit of Eq. (3.9) to the concentration enhancement data in Fig. 3.6. . . . .	81
5.1	Parameter constraints using our GR pipeline for a $\Lambda$ CDM fiducial cosmology. .	127
5.2	Parameter constraints using our full $f(R)$ pipeline for both $\Lambda$ CDM and F5 fiducial cosmologies. . . . .	130
6.1	Specifications of the four dark-matter-only simulations used to study cluster and halo properties in nDGP. . . . .	146
D.1	Specifications of the simulations used to tune our baryonic model. . . . .	248

---

# Introduction

## 1.1 $\Lambda$ CDM: the standard model of cosmology

Cosmology is the study of the origin and evolution of the Universe. The primary goal is to understand what happened between the ‘Big Bang’ ( $\sim 14$  billion years ago) and the present-day and to understand the fundamental physics driving this evolution; how and why quantum fluctuations evolved into density perturbations and eventually into galaxies like the Milky Way in which we were formed. The best way to answer these questions is to develop theoretical models whose predictions can be compared with astronomical observations.

The current standard model of cosmology is the ‘ $\Lambda$ -cold-dark-matter’ ( $\Lambda$ CDM) paradigm. This is the simplest model whose predictions can match a number of key observations, including the temperature fluctuations in the cosmic microwave background (CMB) (e.g., [Spergel et al., 2003](#)), the late-time accelerated expansion of the cosmos (e.g., [Riess et al., 1998](#)), and the distribution of galaxies on large scales (e.g., [Eisenstein et al., 2005](#)). The model makes the following assumptions: gravity, which is the dominant force on cosmological scales, obeys Einstein’s General Relativity (GR); matter is primarily made up of a ‘cold dark matter’ component, which is assumed to consist of massive particles which are yet to be detected experimentally; a smaller proportion ( $\sim 15\%$ ) of matter is made up of ‘visible’ matter, including gas and stars (we will refer to this as ‘baryonic’ matter); and the accelerated expansion is brought about by assigning a positive value to the cosmological

constant  $\Lambda$  which appears in the framework of GR.

Throughout this chapter, we will use the unit convention  $c = 1$  for the speed of light in a vacuum. Greek indices can take values 0, 1, 2 and 3, and, unless otherwise stated, a subscript  $_0$  denotes the present-day value of a quantity.

### 1.1.1 Cosmic expansion in $\Lambda$ CDM

The size of the Universe at a given time  $t$  can be parameterised using the cosmic scale factor  $a$ , which is normalised such that  $a(t = t_0) \equiv a_0 = 1$  at the present-day. Because the Universe is expanding, the light from distant galaxies is observed to be redshifted. The redshift  $z$  is related to the scale factor by the relation  $1 + z = 1/a$ , and both  $z$  and  $a$  can be used as alternative coordinates of time.

The expansion of space can be determined using the Einstein field equations which govern GR:

$$G_{\alpha\beta} + \Lambda g_{\alpha\beta} = 8\pi G T_{\alpha\beta}, \quad (1.1)$$

where  $G_{\alpha\beta}$  is the Einstein tensor,  $\Lambda$  is the cosmological constant,  $g_{\alpha\beta}$  is the metric tensor,  $G$  is Newton's gravitational constant and  $T_{\alpha\beta}$  is the stress-energy tensor. This tensor equation describes how the curvature of spacetime (encapsulated in  $G_{\alpha\beta}$ ) is related to the distribution of energy and momentum (encapsulated in  $T_{\alpha\beta}$ ). By using the Friedmann-Lemaître-Robertson-Walker metric for a homogeneous and isotropic expanding universe, the Friedmann equation can be derived from the 00 component of Eq. (1.1):

$$\left(\frac{\dot{a}}{a}\right)^2 \equiv H^2 = \frac{8\pi G}{3}(\rho_M + \rho_r) + \frac{K}{a^2} + \frac{\Lambda}{3}, \quad (1.2)$$

where  $H$  is the Hubble parameter. This equation can determine the cosmic expansion rate (where  $\dot{a}$  is the time-derivative of  $a$ ) as a function of the background densities of matter  $\rho_M$  and radiation  $\rho_r$ , the curvature  $K$  and the cosmological constant. It can be shown that a positive cosmological constant is required for an accelerating expansion: physically, this can be thought of as a vacuum energy with constant density  $\rho_v = \Lambda/(8\pi G)$ .

Observations (e.g., [Aghanim et al., 2020](#)) indicate that the Universe is extremely close to being spatially flat ( $K \approx 0$ ). Therefore, the second term in the right-hand side of

Eq. (1.2) can be neglected and the sum of the energy densities  $\rho_M$ ,  $\rho_r$  and  $\rho_v$  is equal to the critical density for a spatially flat universe:

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}. \quad (1.3)$$

In the  $\Lambda$ CDM framework, the energy densities are often reparameterised in terms of the dimensionless ratios of the present-day densities with respect to the critical density:  $\Omega_M = \rho_{M,0}/\rho_{\text{crit},0}$ ,  $\Omega_r = \rho_{r,0}/\rho_{\text{crit},0}$  and  $\Omega_\Lambda = \rho_v/\rho_{\text{crit},0}$ . The Friedmann equation can be rewritten in terms of these parameters:

$$\frac{H(z)}{H_0} \equiv E(z) = \sqrt{\Omega_M(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda}, \quad (1.4)$$

where  $H_0$  is the present-day value of the Hubble parameter (the ‘Hubble constant’),  $E(z)$  is the time evolution of the Hubble parameter, and we have used the time-dependencies  $\rho_M(z) \propto (1+z)^3$  and  $\rho_r(z) \propto (1+z)^4$  to derive the right-hand side. For a given set of values of the parameters  $H_0$ ,  $\Omega_M$ ,  $\Omega_r$  and  $\Omega_\Lambda$ , the expansion history of the Universe can be inferred using Eq. (1.4). Observations show that the radiation component  $\Omega_r$  ( $\approx 10^{-4}$ ) is very small compared to  $\Omega_M$  ( $\approx 0.3$ ) and  $\Omega_\Lambda$  ( $\approx 0.7$ ) (e.g., [Aghanim et al., 2020](#)). Our work is primarily focused on late times (low  $z$ ), where the radiation contribution in Eq. (1.4) is negligible; we will therefore assume  $\Omega_r = 0$  for the remainder of this thesis.

### 1.1.2 Strengths and weakness of $\Lambda$ CDM

One of the most notable successes of  $\Lambda$ CDM is in accurately predicting the statistical properties of the CMB temperature fluctuations. This is achieved using linear perturbation theory, which uses fluid dynamics to predict the evolution of primordial perturbations in the matter density field (which are seeded by quantum fluctuations) from inflation to the epoch of recombination. One of the key differential equations in perturbation theory, which will appear later in this thesis, is the Poisson equation:

$$\nabla^2 \Phi = 4\pi G \delta \rho_M, \quad (1.5)$$

where  $\nabla^2$  is a (second-order) space derivative rather than a spacetime derivative. This is used to relate the Newtonian gravitational potential  $\Phi$  to the matter density perturbations

$\delta\rho_M$  of the cosmic fluid. By comparing observations of the CMB with the predictions from linear perturbation theory, precise constraints can be made of  $\Lambda$ CDM parameters including  $\Omega_M$ ,  $\Omega_\Lambda$  and  $H_0$  (e.g., [Ade et al., 2014](#)).

At later times, the evolution of the density fluctuations (which start to undergo gravitational collapse) becomes nonlinear and extremely difficult to solve analytically. Instead, computer simulations can be used to evolve the density field from some early time (based on initial conditions that are consistent with CMB observations) to the present-day (e.g., [Boylan-Kolchin et al., 2009](#)). This numerical approach is able to reproduce the large-scale distribution of galaxies at the present-day (the ‘cosmic web’). This close match to galaxy observations is another success of  $\Lambda$ CDM, and comparing predictions and observations of galaxy statistics (for example, the abundance of galaxy clusters) offers another means of constraining the model parameters (e.g., [Anderson et al., 2012](#); [Ade et al., 2016](#)).

Arguably the greatest success of  $\Lambda$ CDM is that it requires only six independent parameters to accurately reproduce these observations and other phenomena such as the late-time accelerated expansion. This simplicity is why it has become the standard working model of cosmology.

Despite these successes,  $\Lambda$ CDM also has a number of weaknesses. For example: dark matter particles have still not been detected by particle physics experiments; numerical simulations predict too many dwarf-size galaxies and Milky Way satellites (e.g., [Weinberg et al., 2015](#)); and estimates of the vacuum energy density from quantum electrodynamics are many orders of magnitude larger than the value derived from cosmological constraints of  $\Lambda$  (e.g., [Martin, 2012](#)). In this thesis, we will be addressing the latter, namely the possibility that a different mechanism is driving the accelerated cosmic expansion. One possibility is that the acceleration is caused by some exotic matter species that has not yet been detected; however, we will be focusing on another idea that the acceleration arises due to departures from GR on large scales.

## 1.2 Modifications to General Relativity

A wide range of modified gravity (MG) models have been used to explain the accelerated cosmic expansion (see, e.g., [Koyama, 2016](#)). We will be focusing on two popular models —  $f(R)$  gravity (e.g., [Sotiriou and Faraoni, 2010](#); [De Felice and Tsujikawa, 2010](#)) and the Dvali-Gabadadze-Porrati (DGP) model ([Dvali et al., 2000](#)) — which both feature an extra scalar field that mediates a force between matter particles. Since there are already four observed fundamental interactions, this is called the ‘fifth force’. When this force is able to act, the total strength of gravity is enhanced; this can speed up the formation of large-scale structure and create observational signatures, such as an enhanced abundance of galaxy clusters, which can be used to probe gravity on large scales. However, tests of gravity within our Solar System have already verified GR to a remarkably high precision (e.g., [Will, 2014](#)), ruling out a strengthened gravitational force within our local environment; to get around this, both models feature ‘screening’ mechanisms which suppress the fifth force in particular regimes, including sufficiently high-density regions or regions where the second derivatives of the scalar field are large, enabling the models to evade Solar System tests.

The following sections will provide some background on these models, including details of the underlying theory, the screening mechanisms and the strength of the fifth force.

### 1.2.1 $f(R)$ gravity

The  $f(R)$  gravity model is an extension of GR which is constructed by adding a nonlinear function of the Ricci scalar curvature,  $f(R)$ , to the  $R$  term in the Einstein-Hilbert action:

$$S = \int d^4x \sqrt{-g} \left[ \frac{R + f(R)}{16\pi G} + \mathcal{L}_M \right], \quad (1.6)$$

where  $g$  is the determinant of the metric and  $\mathcal{L}_M$  is the Lagrangian density for matter fields (as mentioned above, we will mainly focus on late-time behaviour, where matter is non-relativistic). By setting the variation of the action with respect to the metric to zero, we obtain the modified Einstein field equations, which now contain a new tensor  $X_{\alpha\beta}$ :

$$G_{\alpha\beta} + X_{\alpha\beta} = 8\pi G T_{\alpha\beta}, \quad (1.7)$$



where:

$$X_{\alpha\beta} = f_R R_{\alpha\beta} - \left( \frac{f}{2} - \square f_R \right) g_{\alpha\beta} - \nabla_\alpha \nabla_\beta f_R. \quad (1.8)$$

The tensor  $R_{\alpha\beta}$  represents the Ricci curvature tensor,  $\nabla_\alpha$  is the covariant derivative associated with the metric, and  $\square$  is the d'Alembert operator. The derivative  $f_R \equiv df(R)/dR$  represents the extra dynamic degree of freedom in this model, and can be treated as a scalar field whose dynamics is governed by the following equation of motion:

$$\square f_R = \frac{1}{3}(R - f_R R + 2f + 8\pi G \rho_M), \quad (1.9)$$

which is derived from the trace of Eq. (1.7). The scalar field mediates a fifth force whose physical range is set by the Compton wavelength:

$$\lambda_C = a^{-1} \left( 3 \frac{df_R}{dR} \right)^{\frac{1}{2}}. \quad (1.10)$$

The fifth force is an attractive force which is felt by massive particles. In sufficiently low-density environments (for example, cosmic voids and the outer regions of galaxy clusters), it enhances the strength of gravity by up to a factor of 1/3. However, in high-density regions, it is suppressed and gravity behaves according to GR. This is caused by the chameleon screening mechanism (e.g., [Khoury and Weltman, 2004a,b](#); [Mota and Shaw, 2007](#)), which was included in the model to ensure consistency with Solar System tests. The screening is brought about by an environment-dependent effective mass of the scalar field which becomes very heavy in dense regions so that the fifth force becomes very short-ranged and undetectable. Consequently, the fifth force can only act in regions where the gravitational potential well is not too deep.

Structure formation in  $f(R)$  gravity is governed by the modified Poisson equation, which, in the weak-field and quasi-static limits, is given by (e.g., [Li et al., 2012](#)):

$$\nabla^2 \Phi = \frac{16\pi G}{3} \delta \rho_M - \frac{1}{6} \delta R, \quad (1.11)$$

where  $\delta R$  represents the perturbation of the Ricci scalar curvature and  $\Phi$  now represents the modified gravitational potential. The scalar field is related to the curvature and density perturbations as follows:

$$\nabla^2 f_R = \frac{1}{3}(\delta R - 8\pi G \delta \rho_M). \quad (1.12)$$


---

This work will focus on the Hu-Sawicki (HS) model (Hu and Sawicki, 2007) of  $f(R)$  gravity, which uses the following prescription for  $f(R)$ :

$$f(R) = -m^2 \frac{c_1 (-R/m^2)^n}{c_2 (-R/m^2)^n + 1}, \quad (1.13)$$

where  $m^2 \equiv 8\pi G\rho_{M,0}/3 = H_0^2\Omega_M$ . The theory has three model parameters:  $c_1$ ,  $c_2$  and  $n$ . By choosing  $c_1/c_2 = 6\Omega_\Lambda/\Omega_M$  and assuming that the curvature obeys the inequality  $-R \gg m^2$ , it can be shown that the function  $f(R)$  behaves as a cosmological constant in background cosmology (Hu and Sawicki, 2007), giving rise to the observed late-time accelerated expansion. Under these assumptions, and by choosing  $n = 1$  (which is commonly used in  $f(R)$  literature), it follows that:

$$f_R \simeq -\frac{c_1}{c_2^2} \left( \frac{m^2}{-R} \right)^2. \quad (1.14)$$

Therefore, the background value of the scalar field,  $\bar{f}_R$ , can be expressed as:

$$\bar{f}_R(a) = \bar{f}_{R0} \left( \frac{\bar{R}_0}{\bar{R}(a)} \right)^2, \quad (1.15)$$

where, assuming that the background expansion history is almost indistinguishable from that of  $\Lambda$ CDM, the background curvature  $\bar{R}$  is given by:

$$\bar{R} \simeq -3m^2 \left[ a^{-3} + 4 \frac{\Omega_\Lambda}{\Omega_M} \right]. \quad (1.16)$$

We note that  $-\bar{R} \gg m^2$  is a good approximation for a realistic choice of cosmological parameters. Therefore, using realistic approximations, we are able to work with just a single parameter:  $\bar{f}_{R0}$ , the present-day background scalar field (the over-bar of  $\bar{f}_{R0}$  will be omitted for the remainder of this thesis). The amplitude  $|f_{R0}|$  represents the highest value of the scalar field in cosmic history: see Fig. 1.1 for the time evolution between redshifts 0 and 3. A higher value of  $|f_{R0}|$  corresponds to a stronger modification of GR, allowing regions of higher density to be unscreened at a given time. As a convention, we will refer to models of HS  $f(R)$  gravity with  $|f_{R0}| = 10^{-6}$ ,  $|f_{R0}| = 10^{-5.5}$ ,  $|f_{R0}| = 10^{-5}$ , ... as F6, F5.5, F5, etc.

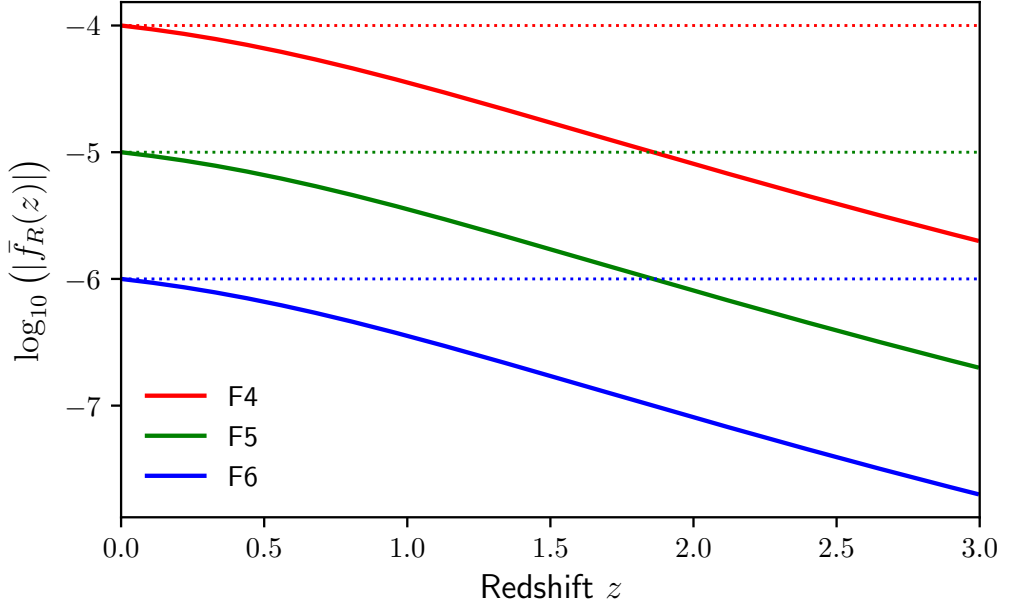


Figure 1.1: Absolute value of the background scalar field in HS  $f(R)$  gravity plotted as a function of redshift for the F4, F5 and F6 models (from top to bottom), assuming the Hu-Sawicki  $f(R)$  model with parameters  $n = 1$  and  $|f_{R0}| = 10^{-4}, 10^{-5}, 10^{-6}$  respectively. Cosmological parameters  $\Omega_\Lambda = 0.719$  and  $\Omega_M = 0.281$  are used.

### 1.2.2 Dvali-Gabadadze-Porrati model

The DGP model consists of two branches: a ‘self-accelerating’ branch (sDGP) and a ‘normal’ branch (nDGP). The former is able to reproduce the late-time cosmic acceleration without requiring any dark energy; however, it is also prone to ghost instabilities which are absent in the latter (e.g., [Koyama, 2007](#)). Because of this, nDGP is the preferred model of gravity. One caveat of this model is that it requires an extra dark energy component in order to be viable; nevertheless, it is still a useful toy model that can be used to search for deviations from GR.

In nDGP, the Universe is modelled as a 4-dimensional brane embedded in a 5-dimensional bulk spacetime. The gravitational action of the model consists of two terms:

$$S = \int_{\text{brane}} d^4x \sqrt{-g} \left( \frac{R}{16\pi G} \right) + \int d^5x \sqrt{-g^{(5)}} \left( \frac{R^{(5)}}{16\pi G^{(5)}} \right), \quad (1.17)$$

where the first term is the usual Einstein-Hilbert action of GR and the second term represents the contribution from the 5D bulk (where  $g^{(5)}$ ,  $R^{(5)}$  and  $G^{(5)}$  are analogous to  $g$ ,  $R$

and  $G$ ). A characteristic length scale, known as the cross-over scale  $r_c$ , can be defined:

$$r_c = \frac{1}{2} \frac{G^{(5)}}{G}. \quad (1.18)$$

The second term of Eq. (1.17) will dominate on scales larger than the cross-over scale, where gravity transitions to 5D. Taking the variation of Eq. (1.17) and setting this to zero leads to the modified Friedmann equation:

$$\frac{H(a)}{H_0} = \sqrt{\Omega_M a^{-3} + \Omega_{DE}(a) + \Omega_{rc}} - \sqrt{\Omega_{rc}}, \quad (1.19)$$

where  $\Omega_{DE}(a)$  is the dimensionless density parameter for the additional dark energy component mentioned above, and  $\Omega_{rc}$  is given by:

$$\Omega_{rc} \equiv \frac{1}{4H_0^2 r_c^2}. \quad (1.20)$$

We will assume that  $\Omega_{DE}(a)$  takes a form that makes  $H(a)$  identical to a  $\Lambda$ CDM expansion history, and that the dark energy component has negligible clustering on the sub-horizon scales that are the focus of this work. The quantity  $H_0 r_c$  is often used to quantify deviations from GR, with a larger value representing a smaller departure from GR. As a convention, we will refer to models with  $H_0 r_c = 5, 2, 1$  and  $0.5$  as N5, N2, N1, and N0.5, respectively.

The modified Poisson equation in nDGP, in the weak-field and quasi-static limits, is given by (Koyama and Silva, 2007):

$$\nabla^2 \Phi = 4\pi G a^2 \delta \rho_M + \frac{1}{2} \nabla^2 \varphi, \quad (1.21)$$

where  $\varphi$  is the extra scalar field of the model. This satisfies the following dynamical equation of motion:

$$\nabla^2 \varphi + \frac{r_c^2}{3\beta a^2} \left[ (\nabla^2 \varphi)^2 - (\nabla_i \nabla_j \varphi)(\nabla^i \nabla^j \varphi) \right] = \frac{8\pi G a^2}{3\beta} \delta \rho_M, \quad (1.22)$$

where the function  $\beta$  is given by:

$$\beta(a) = 1 + 2Hr_c \left( 1 + \frac{\dot{H}}{3H^2} \right) = 1 + \frac{\Omega_M a^{-3} + 2\Omega_\Lambda}{2\sqrt{\Omega_{rc}} (\Omega_M a^{-3} + \Omega_\Lambda)}. \quad (1.23)$$

The fifth force can act on sufficiently large scales, where the nonlinear terms in the square bracket of Eq. (1.22) can be ignored. This enhances the strength of gravity by a factor of  $[1 + 1/(3\beta)]$ , which is larger at later times with present-day enhancements of approximately 1.04

---

for N5, 1.08 for N2, 1.12 for N1 and 1.18 for N0.5. On small scales, the nonlinearity of the scalar field causes the fifth force to be suppressed via the Vainshtein screening mechanism (Vainshtein, 1972). This efficiently suppresses the fifth force in regimes including the Solar System, galaxies and galaxy clusters.

### 1.3 Testing gravity with large-scale structure

The fifth forces in  $f(R)$  gravity and nDGP leave numerous observational signatures in large-scale cosmic structure. For example, the sped-up structure formation that results from a strengthened gravity can lead to a greater number of galaxy clusters being formed by the present-day. Galaxy clusters are the largest objects to have been observed in the Universe: they are gravitationally-bound groups of thousands of galaxies that are found within vast dark matter haloes whose mass ranges from  $\sim 10^{14}M_\odot$  to  $\sim 10^{16}M_\odot$ . Galaxy clusters are therefore thought to trace the highest peaks of the primordial density field, and are therefore highly sensitive to the values of cosmological parameters that affect cosmic structure formation, including  $f_{R0}$  and  $\Omega_{rc}$  which control the effect of the fifth forces in  $f(R)$  gravity and nDGP, respectively. Clusters can also be detected using a number of means, including the X-ray emission from the hot intra-cluster gas and the Sunyaev-Zel'dovich (SZ) effect, which is caused by inverse Compton scattering of CMB photons off high-energy electrons in the intra-cluster gas. By comparing observational data with theoretical predictions, statistics including cluster number counts can be used to probe the strength of gravity on the largest scales.

The following sections are laid out as follows: Sec. 1.3.1 will summarise previous tests of  $f(R)$  gravity and nDGP using large scale structure; Sec. 1.3.2 will outline ongoing and future galaxy cluster surveys and discuss some key considerations for future gravity tests; Sec. 1.3.3 will provide an outline of  $N$ -body and hydrodynamical numerical simulations, which provide the best means of modelling the effects of a strengthened gravity on the properties of galaxy clusters; finally, Sec. 1.3.4 will provide an introduction to Markov chain Monte Carlo (MCMC) sampling, which is a popular method for inferring parameter constraints using observations and theoretical predictions.

### 1.3.1 Previous work

Previous works have tested  $f(R)$  gravity using cluster number counts (e.g., [Cataneo et al., 2015](#); [Liu et al., 2016](#); [Peirone et al., 2017](#)) and the clustering of clusters ([Arnalte-Mur et al., 2017](#)), which are both enhanced by the sped-up structure formation. Meanwhile, the effects of the fifth force on galaxy velocities have been tested using redshift-space distortions (e.g., [Bose and Koyama, 2017](#); [He et al., 2018](#); [Hernández-Aguayo et al., 2019](#)). The temperature of the intra-cluster gas, which is closely correlated with the total gravitational potential, is also enhanced by the fifth force; the model can therefore be tested using the temperature-mass relation (see, e.g., [Hammami and Mota, 2017](#); [Del Popolo et al., 2019](#)). A raised temperature will produce greater X-ray emission from high-energy electrons and a higher level of inverse Compton scattering of CMB photons. The  $f(R)$  model can therefore be tested using the cluster SZ profile ([De Martino, 2016](#)), the cluster gas mass fraction probed through the gas temperature and X-ray luminosity (e.g., [Li et al., 2016](#)), and by comparing weak lensing measurements of clusters (which are unaffected by the fifth force for realistic models) with X-ray and SZ observations (e.g., [Terukina et al., 2014](#); [Wilcox et al., 2015](#)). The model has also been tested using weak lensing by cosmic voids ([Cautun et al., 2018](#)), which, owing to their extremely low density, are largely unaffected by the chameleon screening.

The Vainshtein mechanism of nDGP is more efficient than the chameleon mechanism of  $f(R)$  gravity at screening out the fifth force within galaxy clusters. However, the nDGP fifth force can still speed up the formation of clusters, and in recent years the model has been tested using cluster number counts (e.g., [Schmidt, 2009](#); [von Braun-Bates and Devriendt, 2018](#)) and redshift-space distortions (e.g., [Barreira et al., 2016](#); [Hernández-Aguayo et al., 2019](#)). Again, cosmic voids are largely unaffected by the screening and can be used to test the model (e.g., [Falck et al., 2018](#); [Paillas et al., 2019](#)). Other models which feature Vainshtein screening have also been tested by, for example, comparing weak lensing data with SZ and X-ray cluster observations ([Terukina et al., 2015](#)).

### 1.3.2 Current and future observations

It is currently an exciting time for cluster cosmology: various ongoing and upcoming astronomical surveys are generating vast cluster catalogues using all means of detection, including the clustering of galaxies in optical/infrared surveys (e.g., [Lawrence et al., 2007](#); [Ivezic et al., 2008](#); [Laureijs et al., 2011](#); [Levi et al., 2013](#)), distortions of the CMB by the SZ effect (e.g., [Hasselfield et al., 2013](#); [Ade et al., 2016](#); [Abazajian et al., 2016](#); [Ade et al., 2019](#)), and X-ray emission from the hot intra-cluster gas ([Weisskopf et al., 2000](#); [Jansen et al., 2001](#); [Merloni et al., 2012](#)). These will be many times larger than previous catalogues, and have the potential to significantly advance our understanding of gravity at the largest scales and the mechanisms driving the accelerated cosmic expansion.

Current X-ray surveys include XMM-Newton ([Jansen et al., 2001](#)) and Chandra ([Weisskopf et al., 2000](#)), which together provide coverage of  $\sim 1000$  clusters. The ongoing X-ray survey eROSITA ([Merloni et al., 2012](#)) will detect  $\sim 125,000$  objects with mass  $M > 10^{13} h^{-1} M_{\odot}$  over redshifts  $z < 1$ , most of which will be galaxy groups. The cluster number count data from eROSITA is expected to produce very tight ( $\sim 1\%$ ) constraints on  $\Omega_M$  and the root-mean-squared linear matter density fluctuation  $\sigma_8$ . Another planned X-ray survey, Athena ([Barcons et al., 2017](#)), will launch in the early 2030s and detect clusters with mass  $M > 5 \times 10^{13} M_{\odot}$ , including distant objects at  $z > 2$ .

Current CMB experiments include Planck (e.g., [Ade et al., 2016](#)), the South Pole Telescope (SPT, e.g., [Bocquet et al., 2019](#)) and the Atacama Cosmology Telescope (ACT, e.g., [Hasselfield et al., 2013](#)). The Planck survey has detected over 400 clusters spanning  $2 \times 10^{14} M_{\odot}$  to  $10^{15} M_{\odot}$  over redshifts  $z < 1$ . The Advanced ACTPol (AdvACT) receiver has detected over 4000 clusters since it launched in 2016 ([Hilton et al., 2021](#)). The upcoming SPT-3G survey ([Benson et al., 2014](#)) will extend the work of the existing SPT-SZ survey and detect  $\sim 5000$  clusters with mass  $M \gtrsim 10^{14} M_{\odot}$  by 2023. Further upcoming ground-based experiments include the Simons Observatory (2022, [Ade et al., 2019](#)), which will produce a legacy catalogue of  $\sim 16,000$  clusters, and CMB-S4 (2029, [Abazajian et al., 2016](#)), which is expected to detect  $\sim 140,000$  clusters with mass exceeding  $6 - 8 \times 10^{13} M_{\odot}$ .

Existing optical/infrared surveys include the ongoing Dark Energy Survey (DES, [Ab-](#)

bott et al., 2016), which will detect  $\sim 10^5$  clusters spanning  $z \lesssim 1$ , and the Dark Energy Spectroscopic Instrument (DESI, Levi et al., 2013). Upcoming surveys include the space-based survey Euclid (e.g., Adam et al., 2019), and the ground-based Vera Rubin Observatory (Ivezic et al., 2008) and 4MOST (e.g., Finoguenov et al., 2019). Together, these surveys will be capable of detecting 100,000s of clusters out to high redshifts and an unprecedented number of galaxy groups. For example, Euclid (starting 2022) will detect  $> 10^5$  clusters out to redshift  $z \sim 2$  and down to mass  $10^{13.5} M_\odot$ , while the Vera Rubin Observatory (starting 2022-2023) expects to obtain a nearly complete sample out to  $z \sim 1.2$ , with  $\sim 10^5$  massive clusters and  $\sim 10^6$  galaxy groups. Meanwhile, 4MOST will provide spectroscopic data for  $\sim 40,000$  groups and clusters detected by eROSITA.

Synergies between these surveys will be vital for precise constraints. For example, a major challenge in cluster cosmology is precise calibration of the cluster mass. In the most recent cluster abundance constraints from Planck (Ade et al., 2016), it was noted that the cluster mass calibration, estimated to have a precision of 10 – 15%, was the single largest source of uncertainty. Recent works have combined weak lensing data, which can be used to estimate the absolute mass, with SZ cluster data in order to reduce systematics related to mass calibration (e.g., Bocquet et al., 2019). The weak lensing mass calibration used by the upcoming optical/infrared surveys listed above is expected to achieve 1 – 2% precision, so crossovers with SZ and X-ray selected catalogues will be very helpful. For example, it is planned that the DES cluster sample will overlap with 1000s of SZ and X-ray selected clusters. The upcoming CMB surveys described above will also use CMB lensing as a novel mass calibration technique, which will help to further reduce these systematics.

In order to make the best use of this wealth of observational data for MG tests, it is now vital to prepare accurate and robust theoretical predictions which can be safely combined with the observations to make precise and unbiased constraints. As mentioned in Sec. 1.3.1, the fifth force can affect the total gravitational potential and the thermal properties of clusters, and it can also affect the density profile. If these effects are not fully accounted for in tests that use, for example, cluster number counts, the measurements of the cluster mass could be biased.



For example, previous works (e.g., [He and Li, 2016](#)) have shown that the  $f(R)$  fifth force can alter the amplitude of the intra-cluster gas temperature and related observables, including the integrated SZ flux, by up to 33% (i.e., the maximum enhancement of the total gravitational potential). For the case of the integrated SZ flux, which according to [Ade et al. \(2016\)](#) varies with the cluster mass as  $\sim M^{1.79}$ , failing to account for the effect of the fifth force could lead to an estimate of the mass that is biased by up to  $\sim 15\%$ : this is similar to the 10 – 15% precision of the mass calibration used by Planck. Therefore, as weak lensing and novel CMB lensing mass calibrations continue to approach the 1 – 2% level targeted by upcoming surveys, it will be vital to conduct a full modelling of the effects of the fifth force on cluster properties in order to avoid biased constraints in future tests.

Another exciting prospect is the inclusion of data from the galaxy group regime ( $10^{13}M_{\odot} \lesssim M \lesssim 10^{14}M_{\odot}$ ) in cluster catalogues for the first time (e.g., [Lovisari et al., 2021](#)). This has important implications for gravity tests. For example, the screening of the fifth force in  $f(R)$  gravity is typically less efficient for lower-mass objects, making departures from GR more prevalent in low-mass clusters and groups than in massive clusters. The detections of galaxy groups down to  $10^{13}M_{\odot}$  will consequently make it possible to rule out weaker models of  $f(R)$  gravity.

### 1.3.3 *N-body and hydrodynamical simulations*

The best way to model and understand cluster properties is to use numerical cosmological simulations which can accurately reproduce the formation of large-scale structure. The simplest examples are  $N$ -body simulations which assume that the matter content of the Universe is entirely in the form of collisionless dark matter. These consist of massive ‘particles’ which only interact via gravity, where the gravitational potential is evaluated at each timestep using the Poisson equation, which for GR is given by Eq. (1.5). The simulation is started at some early time using initial conditions which are consistent with, for example, cosmological constraints from CMB temperature fluctuations. Over the course of the simulation, the dark matter particles will attract and cluster together to form gravitationally-bound structures called ‘haloes’ and ‘subhaloes’: these are the theoretical counterparts of

galaxy groups and clusters and galaxies, respectively.

Dark-matter-only (DMO) simulations have been widely used in past works: for example, they can successfully reproduce the web-like structure of the Universe on large scales (e.g., [Boylan-Kolchin et al., 2009](#)), and they can be used for accurate theoretical predictions of the cluster abundance (e.g., [Tinker et al., 2008](#)), which is quantified using the halo mass function (HMF). However, a notable limitation of these simulations is the absence of baryonic matter, which behaves very differently to collisionless dark matter. This can be incorporated in the form of gas, which, in addition to interacting via gravity, must also obey the hydrodynamical fluid equations.

There are a number of ways to include gas. For example, some simulation codes use smoothed particle hydrodynamics (SPH), where the continuum of the gas fluid is approximated using particles with a local resolution that automatically follows the mass flow (e.g., [Springel, 2005](#)). Another example is the AREPO code ([Springel, 2010](#)), which has been used to run all of the hydrodynamical simulations presented in this thesis. This tracks the gas using a moving, unstructured Voronoi mesh, which is made up of gas cells that adaptively refine (split) and derefine (merge) such that the mass of any cell does not differ by more than a factor of two from the mean. The AREPO code can also account for the presence of a magnetic field which dynamically couples to the gas through magnetic pressure. This is achieved by solving the ideal magneto-hydrodynamics equations (see [Pakmor et al., 2011](#); [Pakmor and Springel, 2013](#)).

In order to develop an accurate model for galaxy formation, it is necessary to account for additional effects such as radiative gas cooling, star formation, and stellar and black hole feedback. Significant progress has been made in recent years in the development of sub-resolution models for these additional baryonic processes (e.g., [Schaye et al., 2015](#); [Weinberger et al., 2017](#); [Pillepich et al., 2018c](#)). Throughout this thesis, we will refer to these as ‘full-physics’ models, since they provide the most complete treatments of the underlying baryonic physics that are currently available.

For example, the IllustrisTNG model ([Weinberger et al., 2017](#); [Pillepich et al., 2018c](#)), which is implemented in AREPO, employs a sub-resolution scheme for star formation which

is based on the [Springel and Hernquist \(2003\)](#) model: for gas cells which exceed a particular threshold density, a fraction of the gas mass is converted to mass in star particles at each simulation timestep according to the Kennicutt-Schmidt law. A portion of the gas mass is also converted into wind particles which are launched in random directions: these represent supernova-driven galactic winds. Having travelled outside their local dense interstellar medium, these wind particles will recouple to the gas, transferring their thermal energy and metal content. Gas also undergoes radiative cooling, which is modulated by a time-dependent ultraviolet background. The model also accounts for supermassive black holes, which are seeded at the centre of haloes which exceed a particular mass threshold. These grow through a combination of black hole mergers and Eddington-limited Bondi gas accretion. In high accretion states, a thermal feedback model is employed which heats up the surrounding gas, while in low accretion states a kinetic feedback model is employed which produces black hole-driven winds (see [Weinberger et al., 2017](#); [Vogelsberger et al., 2013](#)).

Significant advances have been made in recent years in the development of numerical simulations of screened modified gravity (see, e.g., [Winther et al., 2015](#)). For example, the MG solver implemented by AREPO can be used to run  $N$ -body and hydrodynamical simulations of  $f(R)$  gravity and nDGP, where structure formation is now governed by the modified Poisson equations, which are given by Eqs. (1.11) and (1.21), respectively. Both models feature a highly nonlinear scalar field which is calculated on an adaptively refining mesh, ensuring accurate calculations in high-density regions. The fifth force is then computed on this mesh and interpolated to the simulation particles and gas cells. This calculation is carried out less frequently in high-density regions where the fifth force is screened out, making the code highly efficient. For the first time, it is now possible to incorporate full-physics baryonic models into these MG simulations (e.g., [Arnold et al., 2019b](#); [Hernández-Aguayo et al., 2020](#)), and this will be invaluable in helping to fully understand the effects of a strengthened gravitational force on the properties of clusters.

As a convention throughout this thesis, when studying  $N$ -body and hydrodynamical simulations, we will define the halo mass as the total mass contained within the sphere

that is centred on the gravitational potential minimum of the halo and encloses an average density of  $\Delta$  times the critical density of the Universe at the halo redshift. We will typically consider overdensities  $\Delta = 200, 500$  or  $300\Omega_M(z)$ , where the corresponding halo mass (radius) is labelled  $M_{200c}$  ( $R_{200c}$ ),  $M_{500c}$  ( $R_{500c}$ ) and  $M_{300m}$  ( $R_{300m}$ ), respectively (we will often neglect the ‘c’ in the subscripts). We note that for  $\Delta = 300\Omega_M(z)$ , the average enclosed density is equal to 300 times the mean matter density of the Universe.

### 1.3.4 Markov chain Monte Carlo sampling

Cosmological parameter constraints can be obtained by combining observations with simulation-based theoretical predictions using MCMC sampling. This is an iterative random sampling technique based on Bayes’ theorem, which states that the conditional probability of the parameter values  $\theta \equiv \{\theta_i\}$ , given the observed data  $D$ , is given by:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \propto P(D|\theta)P(\theta), \quad (1.24)$$

where  $P(D|\theta)$  is the conditional probability of the data given the parameter values (the ‘likelihood’  $\mathcal{L}$ ),  $P(\theta)$  is the *a priori* probability of the parameter values (the ‘prior’) and  $P(D)$  is the probability of the observed data (the latter is a constant, since the data is fixed during the sampling). Because sums are easier to work with than products, it is common to use logarithms:

$$\ln P(\theta|D) = \ln \mathcal{L} + \ln P(\theta) + \text{const} \quad (1.25)$$

The sampling works as follows: starting with an initial point in the parameter space, a new point is generated by perturbing the initial point by a small, random amount. This new point is either accepted or rejected based on the effect of the parameter change on the *a posteriori* probability  $P(\theta|D)$  (the ‘posterior’). An example is the Metropolis-Hastings algorithm: this accepts the new point if the new posterior probability is higher, and it may still accept it if the new posterior is lower by a small amount. This process is repeated such that a chain of points is generated which eventually converges on a small region in the parameter space containing the maximum of the posterior. Typically, multiple chains (‘walkers’) are used, which are generated using different initial points, to ensure that the global maximum of the posterior distribution is successfully located.

This process requires two ingredients: the likelihood and the prior. For the case of cluster number counts, the Poisson likelihood can be used, which gives the probability of counting the observed number of clusters given the parameter-dependent theoretical prediction. For the parameter priors, either a flat (constant) prior is assumed (i.e., no prior knowledge of the parameters is included) or a Gaussian prior can be used which is based on a previous constraint using different observables. For example, the Planck 2015 cluster counts analysis (Ade et al., 2016) adopted Gaussian priors, which are based on previous constraints from CMB temperature fluctuations, Big Bang nucleosynthesis and baryon acoustic oscillations, for parameters including  $H_0$  and the dimensionless baryonic density parameter  $\Omega_b$ , while flat priors can be used for  $\Omega_M$  and  $\sigma_8$  since these are well-constrained by cluster number counts. The prior is then given by:

$$\ln P(\theta) = - \sum_i \frac{(\theta_i - \mu_i)^2}{2\sigma_i^2} + \text{const}, \quad (1.26)$$

where  $\mu_i$  and  $\sigma_i$  are the mean and standard deviation inferred by the previous constraints of parameter  $i$ , and the sum only carries over parameters with a Gaussian prior.

## 1.4 Thesis overview

The primary focus of our work is in developing a general framework for unbiased cluster tests of gravity. Fig. 1.2 provides a broad overview of this framework for the case cluster number count constraints of modified gravity. Most of the chapters of this thesis will focus on different parts of this flowchart.

In Chapter 2, we will provide a more detailed overview of our framework and describe our modelling of the  $f(R)$  enhancement of the dynamical mass of dark matter haloes using  $f(R)$  gravity simulations. The dynamical mass is defined as the mass that is felt by a nearby massive test particle, and its enhancement is equivalent to the gravitational force enhancement. This model forms a core part of the framework for  $f(R)$  gravity constraints and will be referred to in many of the subsequent chapters.

Chapter 3 will focus on our modelling of the enhancement of the halo concentration in  $f(R)$  gravity, which can be used to model the density profiles of haloes. This will be

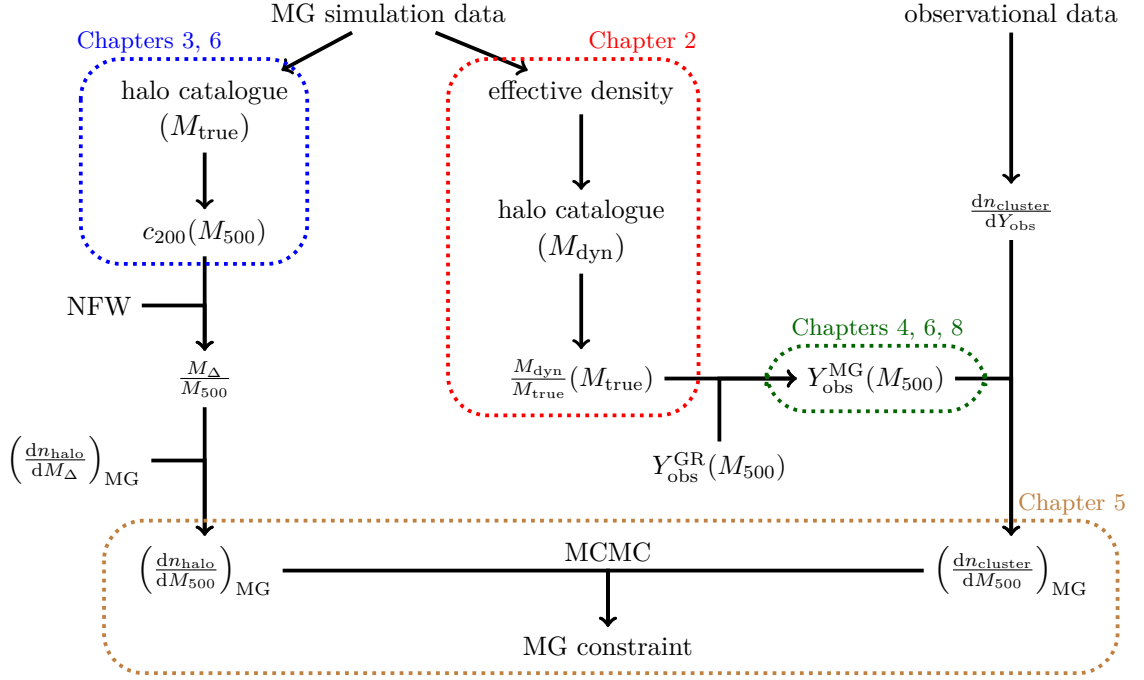


Figure 1.2: Flowchart outlining our general framework for testing gravity using cluster number counts. Using MG simulation data, we can calibrate models for the MG enhancements of the halo concentration (*blue dotted box*) and the dynamical mass (*red dotted box*). The concentration model can be used for converting between different halo mass definitions, which is required if the theoretical predictions and observations are initially defined using different spherical overdensities. The dynamical mass model can be used to convert the GR observable-mass scaling relation into a form that is consistent with the MG model of interest (*green dotted box*). The latter is used to relate the observational mass function,  $dn/dY_{\text{obs}}$  to the theoretical form  $dn/dM$ . Finally, MCMC sampling is used to constrain the MG model parameters using the theoretical predictions and observations (*brown dotted box*). These components will be described for  $f(R)$  gravity and nDGP in the annotated chapters. Chapter 7, which is not shown here, will focus on the thermal and kinetic SZ power spectra, which can potentially be used as alternative probes of gravity, and used in combination with the cluster number count constraint shown here.

required for converting between different halo mass definitions. Then, in Chapter 4, we will demonstrate the effects of the fifth force on the scaling relations between the cluster mass and a number of cluster observables using  $f(R)$  gravity simulations which include a full treatment of baryonic physics. Here, we will show how our model for the dynamical mass enhancement can be used to calibrate observable-mass scaling relations in  $f(R)$  gravity using their GR counterparts. This is required in order to relate the cluster observable (provided in cluster catalogues) to the cluster mass.

In Chapter 5, we will present our full  $f(R)$  constraint pipeline which can be used to constrain  $f_{R0}$  using cluster number counts. The pipeline uses MCMC sampling and it incorporates all of the fifth force effects described above.

The above chapters focus on HS  $f(R)$  gravity, however our framework is designed to be easily extended to other gravity models and observables. In Chapter 6, we will use nDGP simulations (again including a full baryonic treatment) to study and model the effect of the strengthened gravity in this model on observable-mass scaling relations, the halo concentration and the HMF. Then, in Chapter 7, we will demonstrate the impact of  $f(R)$  gravity and nDGP on the angular power spectrum of the SZ effect. This can potentially be used for precision constraints of gravity on large scales using data from future CMB experiments.

In Chapter 8, we will present a retuned baryonic model, which we have calibrated by running over 200 test simulations, that can be used for full-physics simulations of screened modified gravity with much larger cosmological volumes. We have used this model to run a set of large-box simulations in  $f(R)$  gravity and to revisit the  $f(R)$  gravity observable-mass scaling relations using a mass range extending to much higher cluster masses.

Finally, in Chapter 9, we will provide a summary of the main results of each chapter along with some considerations for future work.

# Modelling the dynamical mass of haloes in $f(R)$ gravity

## 2.1 Introduction

In observations, it is generally difficult to directly measure the masses of clusters. This is particularly the case for distant clusters, for which the required exposure time is prohibitively expensive. Instead, one often has to infer them using mass proxies such as the X-ray temperature, luminosity and the SZ Compton  $Y$ -parameter. This, however, can lead to various sources of bias and uncertainty. For example, this can stem from the calibration procedures used to find the scaling relations linking these proxies to the masses, where observational uncertainty and various assumptions can lead to uncertain and possibly biased estimates of the mass. Unless these scaling relations are re-calibrated for any new cosmological models to be studied to remove any sources of bias, these will carry through to the predictions of properties that are dependent on the mass, such as the cluster abundance and the cluster gas fraction, which will therefore lead to biased constraints of the cosmological models and parameters.

In practice, the calibration of the scaling relations can be achieved through different approaches. One way is to use full physics hydrodynamical simulations including radiative processes (e.g., [Fabjan et al., 2011](#); [Nagai et al., 2007](#)). [Fabjan et al. \(2011\)](#) employ this ap-



proach to calibrate the relations for three X-ray proxies. Another way is to use subsamples of a complete data-set as, e.g., in [Vikhlinin et al. \(2009\)](#), where Chandra observations are used to calibrate relations for X-ray proxies that can be cross-checked with weak lensing data. A third option is self-calibration, where the calibration is achieved with additional observables, for instance the clustering of clusters ([Schuecker et al., 2003](#); [Majumdar and Mohr, 2004](#)). In addition to these external calibrations, one can also calibrate data internally, e.g., by simultaneously constraining the scaling relations and cosmological models via a joint likelihood analysis (e.g., [Mantz et al., 2010, 2014b](#)).

The situation becomes even more complicated and largely unexplored when it comes to testing MG theories using galaxy cluster observations. A common effect of MG models which feature a fifth force is to enhance the dynamical mass of a galaxy cluster so that it becomes larger than the true (lensing) mass. This results from the additional gravitational forces. Tests which aim to measure both the dynamical and lensing masses to check for a disparity include recent works by [Terukina et al. \(2014\)](#); [Wilcox et al. \(2015, 2016\)](#); [Pizzuti et al. \(2017\)](#), which utilise actual measurements of the profiles of these two masses for massive clusters. Other probes include the cluster gas fraction ([Li et al., 2016](#)), the clustering of clusters ([Arnalte-Mur et al., 2017](#)) and weak lensing (e.g., [Barreira et al., 2015](#)) by clusters. The resulting weak lensing masses are only modified in some but not all MG models ([Arnold et al., 2014](#)).

While earlier studies have pointed to a strong power of cluster observations in the tests of gravity, one potential issue that has so far not been given detailed attention is that the inferred cluster abundance, and other mass-dependent quantities, can change as a result of the enhancement of the dynamical mass with respect to the true mass, depending on which mass proxy is being used. If this enhancement is not accurately taken into account, the inferred abundance could be biased. In particular, scaling relations that are used to determine the cluster mass should first be calibrated in the contexts of specific MG models in order to incorporate this effect. Furthermore, these scaling relations are often derived using multiple probes, for example X-ray emission and weak lensing, which are affected by MG in different ways even in the same model. This adds more complexity and

challenges for cosmological constraints. The main purpose of this chapter is to consider these complications and propose a suitable calibration method that is straightforward to implement in MG model tests.

As discussed in Chapter 1, a primary focus of this thesis is to develop a framework to incorporate the various effects of MG on galaxy clusters in a self-consistent way. The aim is to have fully calibrated models which incorporates these effects into model predictions and allows for detailed MCMC searches of the parameter space to produce de-biased constraints of gravity. Of particular importance in this framework is the requirement to be able to make reliable model predictions for arbitrary model parameter values, as opposed to a very small number of model parameters that have been studied in detail in previous  $N$ -body simulations of MG (which are therefore not allowing for a continuous search of the large parameter space). To achieve this we will provide various simulation-calibrated fitting formulae that are essential for model predictions. In this chapter, we will focus on the relationship of the lensing and dynamical masses of galaxy clusters in  $f(R)$  gravity. In this model, massive particles feel an extra force (the fifth force) mediated by an additional scalar field  $f_R$ . This field is redshift dependent, and, as described in Chapter 1, its present day background value,  $f_{R0}$ , can be chosen as a model parameter. The enhancement of the dynamical mass therefore depends on the redshift and the background field strength at  $z = 0$ .

Previous works analysing the dynamical mass and lensing mass in  $f(R)$  gravity include Schmidt (2010); Zhao et al. (2011b); Arnold et al. (2014). The studies were model specific, and they did not give a general formula that can be applied to arbitrary values of model parameters and redshifts. For example, the focus may only be on a particular present-day field strength at  $z = 0$ : these results can be used for a qualitative understanding of particular models, but we really need a generic formula that is applicable to general models at all redshifts. In this chapter, we propose such a generic fitting formula which is based on a simple analytical model, the spherical thin-shell model (Khoury and Weltman, 2004a). We check this fitting formula against simulations with different resolutions and find it to work very well across all tested field strengths. Although we use a specific choice of  $f(R)$

gravity as our example, as discussed below, the results are expected to be applicable to or have useful implications for general chameleon gravity theories (e.g., [Gronke et al., 2015](#)).

The chapter is organised as follows: Sec. 2.2 discusses the key results of the thin-shell model, and defines the effective mass, which can be used interchangeably with the dynamical mass in simulations; Sec. 2.3 discusses the background behind the use of galaxy clusters in constraining cosmological models, presents an outline of our proposed framework for  $f(R)$  constraints, and proposes a method to account for the dynamical mass enhancement in scaling relations; Sec. 2.4 summarises the properties of the simulations that are used and how we make use of them in our analyses, presents our fitting formula for the enhancement, and illustrates the method used to test this model; Sec. 2.5 presents the main results of our tests, including key formulae that have been fitted to the simulation data; and finally, Sec. 2.6 summarises the key insights from this investigation and the implications for future work. Also, in Appendix A, we summarise the results obtained from using an alternative fitting procedure and show consistency tests to check for dispersions between the various data-sets used.

## 2.2 Background

In Sec. 2.2.1, we describe the thin shell model, which provides a way of modelling the chameleon screening in  $f(R)$  gravity. Then, in Sec. 2.2.2, we define the dynamical and effective mass of dark matter haloes, and describe how to measure these from simulations.

### 2.2.1 Thin-shell model

A useful way to model chameleon screening is via thin-shell modelling, which was first proposed in [Khoury and Weltman \(2004b\)](#) and has been used extensively in theoretical modelling, (e.g., [Li and Efstathiou, 2012](#); [Lombriser et al., 2013, 2014](#)). Consider a constant spherically symmetric top-hat matter density,  $\rho_{\text{in}}$ , within a radius,  $r_{\text{th}}$ , where  $\phi_{\text{in}}$  and  $\phi_{\text{out}}$  represent the scalar field inside and outside of  $r_{\text{th}}$  respectively. Given this setup, one can

make the following approximation:

$$\frac{\Delta r}{r_{\text{th}}} \approx (3 + 2\omega) \frac{\phi_{\text{in}} - \phi_{\text{out}}}{6\Psi_{\text{N}}} \approx -\frac{\phi_{\text{out}}}{2\Psi_{\text{N}}}, \quad (2.1)$$

where  $\Delta r$  is the distance (from the boundary of the top-hat density distribution) necessary for the scalar field,  $\phi$ , to settle from  $\phi_{\text{out}}$  to  $\phi_{\text{in}}$ , which to a good approximation is  $\phi_{\text{in}} \approx 0$ .  $\omega$  is the Brans-Dicke parameter, equal to zero for the  $f(R)$  model under consideration. One can furthermore identify  $\phi$  with  $f_R$ , and  $\phi_{\text{out}}$  with the background value  $\bar{f}_R(z)$  for a given model and redshift. The depth of the Newtonian potential at the boundary,  $\Psi_{\text{N}}$ , is given by,

$$\Psi_{\text{N}} = \frac{GM}{r_{\text{th}}}, \quad (2.2)$$

with  $M$  the mass enclosed in the spherical top-hat. Using,

$$M \equiv \frac{4\pi}{3} \rho_{\text{in}} r_{\text{th}}^3, \quad (2.3)$$

we find that  $\Psi_{\text{N}} \propto M^{\frac{2}{3}}$  for a fixed density.

In this chapter, we will focus on dark matter haloes found from  $N$ -body simulations. To make a connection between these haloes and the spherical top-hat densities described above which are used for thin-shell modelling, we make two approximations. First, dark matter structures in real simulations are not spherically symmetric, but we approximate them as spherical. Second, the radial density distribution of dark matter haloes are known to satisfy a Navarro-Frenk-White (Navarro et al., 1997, NFW) profile,

$$\rho(r) = \frac{\rho_0}{(r/R_s)(1 + r/R_s)^2}, \quad (2.4)$$

where  $\rho_0$  is a parameter with the same unit as density, and  $R_s$  is the scale radius.  $\rho(r)$  scales like  $r^{-1}$  ( $r^{-3}$ ) in the inner (outer) part of a halo, and is not a constant within the halo radius,  $R_{\Delta\text{c}}$ , which is determined as the distance from the halo centre within which the mean density is  $\Delta$  times the critical density of the Universe,  $\rho_{\text{crit}}$ , at the halo redshift. In our modelling, we treat the haloes as top-hats with density equal to  $M_{\Delta\text{c}} / \left(\frac{4}{3}\pi R_{\Delta\text{c}}^3\right)$ , where  $M_{\Delta\text{c}}$  is the halo mass, i.e., the mass enclosed in  $R_{\Delta\text{c}}$ <sup>\*</sup>. It is furthermore shown in Arnold

---

<sup>\*</sup>For a more detailed and realistic modelling of chameleon screening, see, e.g., Lombriser et al. (2012, 2014); Cataneo et al. (2016). However, as we show below, our simpler treatment works well and its predictions are in excellent agreement with simulations.

et al. (2016), that the above scaling approach also works for ideal NFW haloes, validating our second assumption. The top-hat radius is given by  $r_{\text{th}} = R_{\Delta c}$ .

With the above approximations, we have

$$\Psi_N = \frac{\frac{4\pi G}{3} \rho_{\text{crit},0} \Delta (1+z)^3 \frac{r_{\text{th}}^3}{(1+z)^3}}{\frac{r_{\text{th}}}{1+z}} = \frac{GM}{r_{\text{th}}} (1+z) \propto M^{\frac{2}{3}} (1+z), \quad (2.5)$$

where  $\rho_{\text{crit},0}$  is the critical density today, and so  $\rho_{\text{crit},0} \Delta$  is the mean matter density in the halo today; the factor  $(1+z)^3$  multiplying the density guarantees that we are using the physical density at redshift  $z$ , and the  $(1+z)$  factors associated with  $r_{\text{th}}$  ensures that we use the physical radius (note that  $R_{\Delta c} = r_{\text{th}}$  is the comoving radius of a halo).

With this setup, a qualitative argument can be made (e.g., Li and Efstathiou, 2012) that gravity is enhanced by the maximum factor  $4/3$  when  $\Delta r \geq \frac{r_{\text{th}}}{3}$ . On the other hand, a small positive constant  $\epsilon \ll 1$  can be defined such that one can assume no deviation from GR when  $\Delta r \leq \epsilon r_{\text{th}}$ .

From the theoretical arguments discussed above, it is expected that the dynamical mass of a halo in  $f(R)$  gravity varies in a range  $M_{\text{true}} \leq M_{\text{dyn}} \leq \frac{4}{3} M_{\text{true}}$  (Schmidt, 2010; Zhao et al., 2011b). One can define the smallest true halo mass,  $M_1$ , for which there is no deviation from GR ( $M_{\text{dyn}} = M_{\text{true}}$ ), and the highest true halo mass,  $M_2$ , for which there is no chameleon suppression of the scalar field ( $M_{\text{dyn}} = \frac{4}{3} M_{\text{true}}$ ). From Eqs. (2.1) and (2.5) and using the definitions for  $M_1$  and  $M_2$ , these are respectively given by

$$M_1 = \kappa_1 \left( \frac{1}{\epsilon} \frac{\bar{f}_R(z)}{1+z} \right)^{\frac{3}{2}} \propto \left( \frac{\bar{f}_R(z)}{1+z} \right)^{\frac{3}{2}}, \quad (2.6)$$

and

$$M_2 = \kappa_2 \left( 3 \frac{\bar{f}_R(z)}{1+z} \right)^{\frac{3}{2}} \propto \left( \frac{\bar{f}_R(z)}{1+z} \right)^{\frac{3}{2}}, \quad (2.7)$$

where the constants  $\kappa_1$  and  $\kappa_2$  enclose Newton's gravitational constant along with some other constant factors from Eqs. (2.1, 2.2, 2.3):

$$\kappa_1 = \kappa_2 = (2GH_0)^{-1} \Delta^{-1/2}. \quad (2.8)$$

Both masses display power law fits as functions of  $\frac{\bar{f}_R(z)}{1+z}$ , and this is an important observation in this chapter: when comparing thin-shell model predictions against  $N$ -body simulations,

both of them should be expressed as a function of  $\bar{f}_R(z)/(1+z)$ . An additional advantage is that this makes the dependence on the model parameter  $f_{R0}$  implicit: two models, A and B, with different  $f_{R0}$  values, should have the same  $\bar{f}_R(z)/(1+z)$  value at some different redshifts  $z_A$  and  $z_B$ . If the thin-shell model is generic enough, its predictions for model A at  $z_A$  and model B at  $z_B$  should be the same, irrespective of the fact that these are two different models. We shall show below that this is indeed the case, and so promises a way to constrain general  $f(R)$  models.

In reality, chameleon screening comes not only from a halo's own mass, but also from the matter that surrounds it. This can be considered as environmental screening. This is more important in F6 than in F4 and F5, because in the former the weak scalar field is more easily suppressed, occasionally resulting in total suppression of the field inside a low-mass halo if it is within a larger scale high-density environment. This means that the background field value at a halo,  $\bar{f}_R(z)$ , evaluated by Eq. (1.15), may often be incorrect if there is a surrounding high-density environment. Therefore a better approximation for the thin-shell modelling would be to replace  $\Psi_N$  in Eq. (2.1) with  $\Psi_N + \Psi_{\text{env}}$  with  $\Psi_{\text{env}}$  the average Newtonian potential caused by the environment at the location of the halo (He et al., 2014; Shi et al., 2017), which can be read from the simulation data. For the time being this will not be included in the modelling in this investigation, as it is not necessary to achieve such accuracy in the statistical treatment we aim for. Our approach will cover haloes which live in different environments so that the effects of  $\Psi_{\text{env}}$  largely cancel when looking at the median of all haloes (see below for further comments on this point).

### 2.2.2 Dynamical mass and effective mass

The dynamical mass of a cluster or halo is the mass that massive test particles (e.g., stars or nearby galaxies) feel. It can be measured using the relationship between the gravitational potential energy and the kinetic energy of all of the constituent parts. In simulations it can be calculated for each halo, detected from the density field created by the dark matter particles.

The formation of large-scale structures in  $f(R)$  gravity is largely determined by the

modified Poisson equation, which recall is given by Eq. (1.11). An effective density field,  $\delta\rho_{\text{eff}}$  (He et al., 2015), can be defined such that Eq. (1.11) can be cast into a form that is similar to Eq. (1.5):

$$\nabla^2\Phi = 4\pi G\delta\rho_{\text{eff}}, \quad (2.9)$$

where  $\delta\rho_{\text{eff}}$  and  $\delta\rho$  are related via:

$$\delta\rho_{\text{eff}} \equiv \left( \frac{4}{3} - \frac{\delta R}{24\pi G\delta\rho} \right) \delta\rho. \quad (2.10)$$

The effective haloes are then identified from the effective density field, which is not necessarily the same as the true density field. In GR the two are seemingly the same but in MG they are different. It has been suggested in previous work by He et al. (2015) that using the effective density field to describe haloes allows us to view the dynamical properties of haloes in an  $f(R)$  model as in a  $\Lambda$ CDM cosmology. In this sense, calculations of dynamical properties, such as the circular velocity of the halo, can be done assuming GR regardless of the model ( $f(R)$  gravity or GR) that the simulation is actually run for, as long as the effective mass of a halo is known. Therefore the effective mass can be used as a proxy for  $M_{\text{dyn}}$ . As is evident from Eq. (2.10), the maximum enhancement to the true density field is  $4/3$ . Thus both the effective and the dynamical mass vary between  $M_{\text{true}}$  and  $\frac{4}{3}M_{\text{true}}$ . In what follows we shall use the effective mass and dynamical mass interchangeably, regardless of the (minor) differences between them (He et al., 2015).

## 2.3 A framework for gravity tests using clusters

In this section, we focus on tests of HS  $f(R)$  gravity using the galaxy cluster abundance (see, e.g., Schmidt et al., 2009; Mak et al., 2012; Cataneo et al., 2015, for earlier works along this direction). This is a specific case of the general framework shown in Fig. 1.2, which is intended to be used for testing a range of MG models. Our proposed  $f(R)$  framework is sketched in Fig. 2.1.

A fitting formula for the HMF is required to predict the halo abundance, and this can be obtained by using semi-analytical models calibrated by simulations. In this thesis, we adopt the model which has been proposed and calibrated by Cataneo et al. (2016),

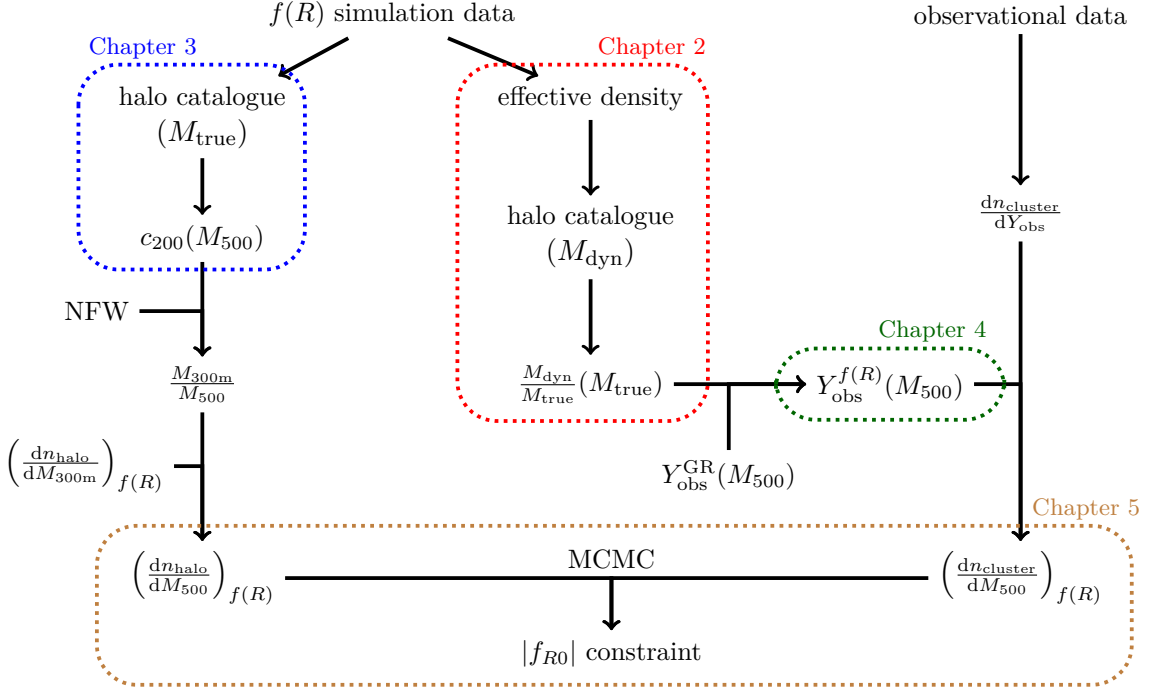


Figure 2.1: Flowchart outlining our framework for constraining the present-day value of the background scalar field,  $f_{R0}$ , of HS  $f(R)$  gravity. Using  $f(R)$  simulation data, we have calibrated models for the  $f(R)$  enhancements of the halo concentration (Chapter 3, blue dotted box) and the dynamical mass (this chapter, red dotted box). The concentration model is used to convert the mass definition of the theoretical HMF into a form that is consistent with the observational data. The dynamical mass model is used to convert the GR observable-mass scaling relation into a form that is consistent with  $f(R)$  gravity (Chapter 4, green dotted box). The latter is used for relating the observational mass function,  $dn/dY_{\text{obs}}$  to the theoretical form  $dn/dM_{500}$ . Finally, MCMC sampling is used to constrain  $f_{R0}$  using the theoretical and observational predictions (Chapter 5, brown dotted box).

which itself is built upon earlier works (Li and Efstathiou, 2012; Li and Lam, 2012; Lam and Li, 2012; Lombriser et al., 2013, 2014) motivated by excursion set theory (Bond et al., 1991); this will be discussed in Sec. 2.3.1.1. The Cataneo et al. (2016) HMF has been calibrated using the halo mass definition  $M_{300\text{m}}$ . To ensure generality, we will also require a mass conversion,  $M_{300\text{m}}(M_{\Delta})$ , to allow conversions to arbitrary mass definitions, which will require a concentration-mass relation (e.g.,  $c_{200}(M_{\Delta})$ ), of dark matter haloes in  $f(R)$  gravity. This is discussed in Sec. 2.3.1.3, in addition to other work to be carried out. These ingredients will enable us to predict a theoretical cluster abundance for generic  $f(R)$  models and mass definitions.



On the observational side, a key observable to be used in our test framework is the cluster abundance derived from SZ and X-ray surveys, such as Planck’s SZ cluster abundance (Ade et al., 2016). As discussed in Sec. 2.3.1.2, converting from cluster observables to the cluster mass typically involves the use of a scaling relation, however the most accurate scaling relations that are currently available are observational and/or derived for  $\Lambda$ CDM. We propose a method for converting these relations from  $\Lambda$ CDM to  $f(R)$  gravity, based on the findings of He and Li (2016). We discuss this point in more detail in Sec. 2.3.1.2. The conversion requires a formula for the ratio  $M_{\text{dyn}}/M_{\text{true}}$ , which is the focus of this chapter. Our procedure to measure  $M_{\text{dyn}}/M_{\text{true}}$  as a function of  $M_{\text{true}}$ ,  $z$  and  $\bar{f}_R$  is discussed in Sec. 2.4 and our results are presented in Sec. 2.5. We show that a simple fitting formula for  $M_{\text{dyn}}/M_{\text{true}}$  motivated by the theoretical modelling of Sec. 2.2.1 works very well in describing the results of a large suite of simulations. The simulations are introduced in Sec. 2.4.1.

Following the corrections described above, the predicted and observed abundances can be combined to constrain  $|f_{R0}|$  by confronting theoretical predictions for models with an arbitrary value of  $f_{R0}$  with observations. A continuous parameter space search can be carried out using techniques such as MCMC, which accounts for relevant covariances between data. The fitting formulae for various quantities, with corresponding errors, can be used to construct mock cluster catalogues to validate the model constraint pipeline. In Sec. 2.3.2, we will also mention some other possible observables which can be included in this framework and which will also require a knowledge of  $M_{\text{dyn}}/M_{\text{true}}$  which we focus on in this chapter.

### 2.3.1 Cluster abundance tests

One of the frequently used probes of cosmological models and the underlying theory of gravity is the cluster abundance, defined as the number density of galaxy clusters per unit mass interval,  $\frac{dn_{\text{cluster}}}{d\log_{10} M}$ . This depends sensitively on the cluster mass,  $M$ , which means that model tests using the cluster abundance require an accurate measurement of the cluster mass. We have seen that the term ‘mass’ can be ambiguous in MG theories because different

observables depend on different masses, e.g., dynamical versus lensing mass. Therefore, any effects of  $f(R)$  gravity on the mass should be accounted for to prevent a biased prediction of the abundance.

The theoretical counterparts of galaxy clusters in  $N$ -body simulations are massive dark matter haloes ( $\gtrsim 10^{14} h^{-1} M_{\odot}$ ). A prediction of the cluster abundance can be obtained by measuring the abundance of haloes. Some efforts must also be made to account for the limitations of an observational survey, for example the blocking of many clusters by foreground stars and the galactic plane, and the rejection of low signal-to-noise sources. These effects are specific to the survey under consideration. In summary, the following quantities are required:

- An HMF which evaluates the number density of dark matter haloes per unit mass interval;
- A scaling relation to predict the cluster observable, given the mass of the dark matter halo;
- The selection function of the survey, which evaluates the probability of a cluster being detected and included in the resulting data-set, as a function of the observable flux, redshift, etc.;
- The likelihood of the measurements, which would be produced along with the observed data itself.

These corrections will ensure that the prediction of the cluster abundance is consistent with measurements taken in the real Universe using detectors with finite precision. However, the HMF and scaling relations are generally more challenging to implement in  $f(R)$  gravity tests without inducing sources of bias. This can stem from effects like the chameleon screening mechanism and the enhancement of the dynamical mass, which are complicated to model exactly. Secs. 2.3.1.1-2.3.1.2 illustrate our proposed methods to tackle these difficulties, and Sec. 2.3.1.3 discusses other current issues in using the cluster abundance to test  $f(R)$  gravity which we seek to address in this thesis.

### 2.3.1.1 Halo abundance

The abundance of dark matter haloes can be predicted using semi-analytical models, such as excursion set theory (Bond et al., 1991), which generally show reasonable qualitative agreement with simulations. These models connect high peaks in the initial density field to the late-time massive dark matter haloes by assuming spherical collapse. However quantitative agreements with simulations are not great, which has motivated models with more physical assumptions, such as the ellipsoidal collapse model (Sheth and Tormen, 1999, 2002; Sheth et al., 2001) which gives up the sphericity assumption above. These efforts have led to various fitting formulae of the HMF in standard  $\Lambda$ CDM, whose parameters can be calibrated using simulations (e.g., Jenkins et al., 2001; Warren et al., 2006; Reed et al., 2007; Tinker et al., 2008).

In MG theories, excursion set theory still applies but the connection between initial density peaks and late-time dark matter haloes becomes more complicated. In some scenarios, such as the Galileon model (e.g., Nicolis et al., 2009; Deffayet et al., 2009), as in  $\Lambda$ CDM, the spherical collapse of an initial top-hat overdensity does not depend on the environment, and analytical solutions can be obtained for their HMFs (Schmidt et al., 2010; Barreira et al., 2013, 2014). In  $f(R)$  gravity and general chameleon models, however, the behaviour of the fifth force is more complicated and the spherical collapse becomes environment-dependent. Theoretical models of HMFs in these theories have been studied in Li and Efstathiou (2012); Li and Lam (2012); Lam and Li (2012); Lombriser et al. (2013, 2014); Kopp et al. (2013), and qualitative agreement with simulations is reasonable.

In this thesis, we adopt the HMF as proposed in Cataneo et al. (2016), which is based on an extension of the theoretical modelling described in Lombriser et al. (2013, 2014) by adding free parameters to the theoretical HMF to account for the chameleon screening mechanism and allow a better match with simulations. These parameters have been fitted using a subset (Crystal, see Sec. 2.4.1) of our DMO  $f(R)$  gravity simulations which have been run for F4, F5 and F6, but they work for general values of  $|f_{R0}|$  within  $[10^{-6}, 10^{-4}]$ . Cataneo et al. (2016) show that their HMF fitting formula agrees with simulation results to within 5%.

We note that the HMF fitting formula is an independent ingredient in our framework as depicted in Fig. 2.1, by virtue of which we can always use the latest and most accurate in our analysis.

### 2.3.1.2 *Scaling relations in $f(R)$ gravity*

The cluster mass is difficult to measure via direct observations, and a scaling relation is usually used to connect the cluster mass to some more readily observable quantities, such as the average temperature,  $T_{\text{gas}}$ , of the intra-cluster gas. This relates to the total mass,  $M$ , via the virial theorem which leads to:

$$\frac{GM}{R} = \frac{3}{2} \frac{k_B T_{\text{gas}}}{\mu m_p}, \quad (2.11)$$

where  $R$  is the cluster radius,  $m_p$  is the proton mass,  $k_B$  is the Boltzmann constant and  $\mu$  is the molecular weight.

We are interested in cluster abundances measured from X-ray emission, the SZ effect and weak lensing. The X-ray radiation by a cluster is generated by the bremsstrahlung process, and the SZ effect is due to the inverse-Compton scattering of cosmic microwave background photons off electrons in the intra-cluster medium. Both of these effects depend on  $T_{\text{gas}}$ . Therefore, several related and easily observable quantities can be used as mass proxies, such as the integrated SZ Compton  $Y$ -parameter,  $Y_{\text{SZ}}$ , the X-ray equivalent of the integrated SZ flux,  $Y_X$ , and the X-ray luminosity,  $L_X$ . For each of these observables the cluster mass can be inferred through a scaling relation.

In  $\Lambda$ CDM, such scaling relations can be obtained in different ways, such as by using hydrodynamical simulations (e.g., Fabjan et al., 2011; Nagai et al., 2007) or from subsets of observed clusters whose masses can be measured in other means, e.g., weak lensing (Vikhlinin et al., 2009). An example is the  $Y_{\text{SZ}} - M$  scaling relation calibrated by the Planck Collaboration (Ade et al., 2016), which incorporates the results from various observational surveys and simulations, and where rigorous methods have been used to prevent various sources of bias, including Malmquist bias and hydrostatic equilibrium bias.

In  $f(R)$  gravity, and in general for any new gravity theory, the scaling relations calibrated for  $\Lambda$ CDM are unlikely to still apply. It is impractical to calibrate these relations by

using hydrodynamical simulations, since they are expensive even for a single specific  $f(R)$  model, let alone the whole  $f_{R0}$  parameter range. Calibrations using a subset of data or using other observables should be treated with caution as well. For example, the scaling relations may be different between the subset of data and the whole sample, due to the environmental dependence of the MG effect, and different observables are proxies of different masses in  $f(R)$  gravity, and so the combined use of different observations is tricky. It is therefore highly desirable to have a physically motivated model for obtaining (certain) scaling relations for arbitrary values of the  $f(R)$  parameter  $f_{R0}$  with good precision and minimal effort.

Along this line and based on the use of the so-called effective mass (Sec. 2.2.2), a procedure for correcting for the effect of MG on the physical properties of clusters, such as their various observable-mass scaling relations, has been proposed by [He and Li \(2016\)](#). This method avoids direct calibration of the cluster mass using full hydrodynamical simulations in the  $f(R)$  model, and instead calculates the scaling relations in  $f(R)$  gravity by using the corresponding ones in standard  $\Lambda$ CDM (which are better known) with a rescaled baryon-to-total mass ratio. Its results are found to agree very well with  $f(R)$  simulations.

[He and Li \(2016\)](#) discussed the cluster mass proxies  $L_X$ ,  $Y_{SZ}$  and  $Y_X$ , and here we describe the result for  $Y_{SZ}$  as an example. Using a non-radiative approximation, in which the baryonic content of the hydrodynamical simulations behaves as an ideal gas satisfying Eq. (2.11),  $Y_{SZ}$  is given by:

$$Y_{SZ} = \frac{\sigma_T}{m_e c^2} \int_0^r dr 4\pi r^2 P_e, \quad (2.12)$$

where  $\sigma_T$  is the Thomson cross section and  $m_e$  is the electron mass. The electron pressure,  $P_e$ , is given by  $P_e = \frac{2+\mu}{5} n_{\text{gas}} k_B T_{\text{gas}}$ , where  $n_{\text{gas}}$  is the number density of gas particles. From the simulations it was found that the  $T_{\text{gas}}-M$  relations for the effective haloes in  $f(R)$  gravity and the haloes in  $\Lambda$ CDM agree very well:

$$T_{\text{gas}}^{f(R)} \left( M_{\text{dyn}}^{f(R)} \right) = T_{\text{gas}}^{\Lambda\text{CDM}} \left( M^{\Lambda\text{CDM}} \right). \quad (2.13)$$

This is as expected given that the temperature and the gravitational potential of a halo are intrinsically linked through the virial theorem.

Using a suite of non-radiative hydrodynamical simulations, it was found that outside the core regions, the profiles of effective haloes in  $f(R)$  gravity closely resemble those in  $\Lambda$ CDM, with a rescaled gas mass fraction:

$$\rho_{\text{gas}}^{f(R)}(r) \approx \frac{M^{f(R)}}{M_{\text{dyn}}^{f(R)}} \rho_{\text{gas}}^{\Lambda\text{CDM}}(r) \propto \frac{M^{f(R)}}{M_{\text{dyn}}^{f(R)}} \frac{\Omega_{\text{b}}}{\Omega_{\text{m}}} \left(r^2 + r_{\text{core}}^2\right)^{-\frac{3\beta}{2}}, \quad (2.14)$$

where  $r_{\text{core}}$  is the core radius and  $\beta$  is the ratio between the specific kinetic energy (kinetic energy per unit mass) of cold dark matter and the specific internal energy (internal energy per unit mass) of gas. For an effective halo in  $f(R)$  gravity with an effective mass that is equal to the true mass of a  $\Lambda$ CDM halo,  $M_{\text{dyn}}^{f(R)} = M^{\Lambda\text{CDM}}$ , it follows from Eqs. (2.13) and (2.14) that:

$$\begin{aligned} & \int_0^r dr 4\pi r^2 \left(\rho_{\text{gas}}^{f(R)}\right)^a \left(T_{\text{gas}}^{f(R)}\right)^b \\ & \approx \left(\frac{M^{f(R)}}{M_{\text{dyn}}^{f(R)}}\right)^a \int_0^r dr 4\pi r^2 \left(\rho_{\text{gas}}^{\Lambda\text{CDM}}\right)^a \left(T_{\text{gas}}^{\Lambda\text{CDM}}\right)^b, \end{aligned} \quad (2.15)$$

where  $a$  and  $b$  are indices of power. By combining this result with Eq. (2.12) it follows that the  $Y_{\text{SZ}}-M$  scaling relations in these two models can be related by:

$$\frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} Y_{\text{SZ}}^{f(R)} \left(M_{\text{dyn}}^{f(R)}\right) \approx Y_{\text{SZ}}^{\Lambda\text{CDM}} \left(M^{\Lambda\text{CDM}} = M_{\text{dyn}}^{f(R)}\right). \quad (2.16)$$

As mentioned previously, this relation has been verified by a suite of non-radiative hydrodynamical simulations. Similar results have been obtained and verified for the other two proxies ( $Y_{\text{X}}$  and  $L_{\text{X}}$ ) as well, and are particularly accurate for  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  with the error just slightly over 3%.

As the scaling relations in  $\Lambda$ CDM are much better understood than in  $f(R)$  gravity, Eq. (2.16) can potentially be used to re-calibrate a scaling relation obtained for  $\Lambda$ CDM, into a form linking  $Y_{\text{SZ}}$  to the cluster dynamical mass in  $f(R)$  gravity.

### 2.3.1.3 *Other issues*

The mass of a galaxy cluster or dark matter halo is usually defined as the mass enclosed in some radius centred around the cluster or halo centre. This is the radius in which the average matter density is  $\Delta$  times the mean matter density or the critical density at the halo redshift. In the literature, different values of  $\Delta$  are commonly used, and so it is essential

to be able to convert between them. As an example, [Cataneo et al. \(2016\)](#), whose  $f(R)$  gravity HMF fitting formula we use by default in our framework, work with  $M_{300\text{m}}$ . As another example, in the literature  $M_{200c}$  is very commonly used.

It is straightforward to convert between the different masses by noting that the different definitions only differ in where the halo boundary lies. Therefore, all we need is the density profile  $\rho(r)$  of a halo. In  $\Lambda\text{CDM}$ , dark matter haloes are well described by the NFW density profile given by Eq. (2.4), which has two free parameters,  $\rho_0$  and  $R_s$ . The NFW profile has also been shown to work well for haloes in  $f(R)$  gravity ([Lombriser et al., 2012](#); [Shi et al., 2015](#)). Of the two NFW parameters, the scale radius,  $R_s$ , can be expressed by using the halo concentration,  $c_\Delta \equiv R_\Delta/R_s$ , and  $\rho_0$  can be further fixed using the halo mass,  $M_\Delta \equiv M(\leq R_\Delta)$ . Therefore, to convert between the different mass definitions requires an understanding of the concentration-mass relation,  $c_\Delta(M_\Delta)$ . We note that, as long as the concentration is known for one overdensity  $\Delta$ , it can be inferred for any other overdensity (for example, see Appendix B.2). We have studied the  $c_{200}(M_{500})$  concentration-mass relation in both screened and unscreened regimes, using data from various  $f(R)$  simulations, and the results will be presented in Chapter 3.

Another issue that merits further investigation is a check of the method by [He and Li \(2016\)](#) against full-physics hydrodynamical simulations including baryonic feedback processes, which go beyond the non-radiative approximations originally used. Studies in  $\Lambda\text{CDM}$  (e.g., [Fabjan et al., 2011](#)) have found that, for certain quantities such as  $Y_X$ , the resulting scaling relation is insensitive to baryonic processes, such as cooling, star formation and black hole feedback, in galaxy formation if the data from the very inner part of a cluster is excluded. We expect the same to apply in  $f(R)$  gravity, but in order to be certain we have conducted an analysis using full-physics hydrodynamical simulations for HS  $f(R)$  gravity, and will present the results in Chapter 4.

Such simulations will also be useful to better understand the impact of galaxy formation on the HMF in  $f(R)$  gravity, though we expect it to be small. We also note that the fitting formula by [Cataneo et al. \(2016\)](#), which has a 3-5% accuracy with the simulation data for F4-F6 and halo masses above  $10^{13}h^{-1}M_\odot$ , was calibrated using DMO simulations

(Crystal, see Sec. 2.4.1).

### 2.3.2 Other observables

As mentioned above, the focus of the remainder of this chapter is a fitting function for the relationship between the dynamical and true masses of dark matter haloes, which would be useful for deriving cluster scaling relations in  $f(R)$  gravity. But the use of this relation is certainly not restricted to this.

A direct use of the  $M_{\text{dyn}}/M_{\text{true}}$  relation is to constrain the fifth force by comparing measurements of  $M_{\text{dyn}}$  and  $M_{\text{true}}$ . In observations, the profiles of these masses can be obtained using the X-ray surface brightness profile and lensing tangential shear profile of a cluster respectively. The measurements can be done for massive clusters for which high-quality X-ray and lensing data are available. [Terukina et al. \(2014\)](#); [Wilcox et al. \(2015, 2016\)](#) performed the first analyses using this method and found constraints on general chameleon gravity theories. A more recent analysis can be found in [Pizzuti et al. \(2017\)](#). The dynamical mass or potential can also be inferred from the escape velocity edges in the radius/velocity phase space, which can be compared with the lensing-inferred mass profile, or the gravitational potential profiles for samples of low- and high-mass haloes, which would feel different effects of gravity due to the chameleon screening, can be compared ([Stark et al., 2016](#)).

Another potentially powerful probe in cluster cosmology is the cluster gas fraction (e.g., [Mantz et al., 2014a](#)),  $f_{\text{gas}} = M_{\text{gas}}/M_{\text{halo}}$ , where  $M_{\text{gas}}$  is the mass of baryons (or hot gas) in the intra-cluster medium and  $M_{\text{halo}}$  is the total halo mass. In massive clusters, the mass of the hot intra-cluster gas dominates over that in cold gas and stars, and thus  $f_{\text{gas}}$  is expected to approximately match the cosmic baryon fraction,  $\Omega_{\text{b}}/\Omega_{\text{M}}$ . However, measurements of  $f_{\text{gas}}$  involve measuring  $M_{\text{halo}}$ , which is the dynamical rather than the true mass of the halo. Constraints from  $f_{\text{gas}}$  on  $f(R)$  gravity are therefore likely to be biased ([Li et al., 2016](#)). To make amends for this we will require a general formula for the ratio  $M_{\text{dyn}}/M_{\text{true}}$ , which is presented in Sec. 2.5.

Our framework is sufficiently flexible to include these, among other, observables in



the ultimate cluster constraints, though certain generalisations may be needed, such as the concentration-mass relations for not only the true but also the effective haloes.

## 2.4 Simulations and methods

The specifications of the  $f(R)$  gravity simulations used in this chapter are presented in Sec. 2.4.1. The procedure to measure the dynamical mass enhancement from this data is discussed in Sec. 2.4.2, along with the details for the modelling of this enhancement and its parameters.

### 2.4.1 Simulations

Our collisionless simulations are run using the ECOSMOG code (Li et al., 2012), a code based on the publicly-available  $N$ -body and hydrodynamical code RAMSES (Teyssier, 2002), and which can be used to run  $N$ -body simulations for a wide range of MG and dynamical dark energy scenarios. The code is efficiently parallelised, and uses adaptive mesh refinement to ensure accuracy of the fifth force solution in high-density regions. In order to reliably fit the dynamical mass enhancement as a function of the halo mass, an appropriate range of halo true mass which covers the transition between  $M_{\text{dyn}} = M_{\text{true}}$  and  $M_{\text{dyn}} = \frac{4}{3}M_{\text{true}}$  would be required. For this reason, three different simulations of varying resolutions were utilised. For the purposes of clarification, these are listed as the *Crystal*, *Jade* and *Diamond* simulations with increasing resolutions.

The parameters and technical specifications of the simulations are listed in Table 2.1. The Hubble expansion rate,  $H_0$ , is set to  $69.7 \text{ kms}^{-1}\text{Mpc}^{-1}$ . Diamond is the highest resolution simulation, and its small particle mass allows lower-mass haloes to be investigated. While Crystal is the lowest resolution, its large volume and particle number mean that higher-mass haloes can be included. Jade is needed in order to provide bridging halo mass regimes with both Crystal and Diamond to ensure that a complete range of masses is tested and to verify that the different simulations agree well in the overlapping regions (see Appendix A.2). Because the results of this investigation are intended to be used with the Planck 2015 data, which only covers up to redshift  $z = 1$ , only simulation snapshots with

Table 2.1: Specifications of the three ECOSMOG simulations used to study the  $f(R)$  enhancement of the dynamical mass, labelled Diamond, Jade and Crystal for convenience. The gold data is defined as having been generated by  $f(R)$  gravity simulations, whereas the silver data comes from effective density data generated from  $\Lambda$ CDM simulations. The Hubble constant,  $H_0$ , is set to  $69.7 \text{ kms}^{-1}\text{Mpc}^{-1}$  in all simulations.

Parameters and data types	Simulations		
	Diamond	Jade	Crystal
box size / $h^{-1}\text{Mpc}$	64	450	1024
particle number	$512^3$	$1024^3$	$1024^3$
particle mass / $h^{-1}M_\odot$	$1.52 \times 10^8$	$6.60 \times 10^9$	$7.80 \times 10^{10}$
$\Omega_M$	0.281	0.282	0.281
$\Omega_\Lambda = 1 - \Omega_M$	0.719	0.718	0.719
gold	F6	F5	F4, F5, F6
silver	F5.5, F6.5	F4.5, F5.5, F6.5	F4.5, F5.5

$z < 1$  are used. This includes 19 snapshots from both Crystal and Diamond, and 33 from Jade. The use of data from only  $z < 1$  also means that we can avoid using high- $z$  data from the Crystal simulations, which suffer from poor resolutions.

Halo catalogues for these simulations are constructed in two steps. First a modified ECOSMOG code is run to generate effective density data from the particle data for all of the snapshots. After that, AHF (Gill et al., 2004; Knollmann and Knebe, 2009), a halo finder which is properly modified to read the effective density data, is run to identify effective haloes. AHF is run with the  $M_{500c}$  mass definition, and the outputted halo catalogues include the ratio  $M_{\text{dyn}}/M_{\text{true}}$  for each halo, as well as the lensing mass which can be treated as  $M_{\text{true}}$ .

Given the expensive cost of full MG simulations, our  $f(R)$  simulation suite only includes a limited number of models. The Crystal simulations have only been run for F4, F5 and F6, Jade has been run for F5 only and Diamond for F6 only. From Fig. 1.1, we can see that up to  $z = 1$  (the redshift limit in the simulation data for our analysis) the three simulated models — F4, F5, F6 — do not cover all possible values of  $\bar{f}_R(z)$  continuously but leave gaps in between. In order to test the proposed model for the dynamical mass enhancement over the greatest possible range of field values, without making too much effort in running full  $f(R)$  simulations for other  $f_{R0}$  values, we propose a simpler

approach. At any desired redshift  $z$ , the MG solver in the ECOSMOG code was run on the particle data of  $\Lambda$ CDM simulations to generate further effective density data by assuming these were actually  $f(R)$  gravity calculations with strengths F4.5 ( $|f_{R0}| = 10^{-4.5}$ ), F5.5 ( $|f_{R0}| = 10^{-5.5}$ ) and F6.5 ( $|f_{R0}| = 10^{-6.5}$ ). Because these calculations involve running the ECOSMOG only for one step (for each  $f_{R0}$  and  $z$ ), they are much less expensive than a full simulation which means that we can afford to run many of them. Indeed, we could repeat this for any other values of  $f_{R0}$ , but found that the above three additional values already give decent overlapping in the halo mass ranges (see below).

AHF effective halo catalogues were then generated for the additional  $f_{R0}$  values using the effective density field from these ‘approximate simulations’, the latter neglecting effects from the different structure formations under these models which could lead to different internal structure and large-scale environments of haloes. For this reason, this additional data is labelled ‘silver’ data, and it was used in addition to the ‘gold’ data which was generated from the actual full  $f(R)$  gravity simulations. We justify the use of silver data by noticing that our thin-shell modelling (see above) treats haloes as spherical top-hats by averaging the mass distribution within  $R_{500c}$  (the same can be done for other halo mass definitions, although in this study we use  $M_{500c}$  when studying  $M_{\text{dyn}}/M_{\text{true}}$ ) and therefore is not sensitive to the actual subtle differences in the halo density profiles from the full and approximate simulations. In addition, we have checked the validity of using silver data by doing the same analysis for  $|f_{R0}| = 10^{-5}$ , for which we have gold data to compare to: as is shown in Appendix A.2, in this case the gold and silver data of F5 are in excellent agreement.

### 2.4.2 Measuring the dynamical mass enhancement

The ratio of the dynamical mass to the true mass of a halo depends on the mass of the halo, the background scalar field of the Universe and the redshift. Because the field is a redshift-dependent quantity, the different snapshots for a given model all have different field values with which to investigate the dynamical mass enhancement. The ratio  $M_{\text{dyn}}/M_{\text{true}}$  is described by two parameters  $p_1, p_2$  (as will be discussed below), which vary with the

background field value  $\bar{f}_R(z)$  and redshift  $z$ . In Sec. 2.2.1 it was shown that, according to our thin-shell modelling, the screening effect can be described by a specific combination of  $\bar{f}_R(z)$  and  $z$ ,  $\bar{f}_R(z)/(1+z)$ , and so we expect that both  $p_1$  and  $p_2$  can be fitted as functions of  $\bar{f}_R(z)/(1+z)$  using their values at the snapshots. In this subsection we describe how this fitting process was carried out in our analysis.

#### 2.4.2.1 *tanh function fit to $M_{\text{dyn}}/M_{\text{true}}$*

In this step, the AHF halo catalogues were first sifted to keep only haloes made up of a sufficient number of dark matter particles and to exclude sub-haloes. The mass criteria for the sifting of Crystal, Jade and Diamond was, respectively,  $M_{500} > (4 \times 10^{13}, 3 \times 10^{12}, 6.5 \times 10^{10}) h^{-1} M_{\odot}$ , which correspond to a minimum number of particles per halo of 513, 454 and 428. These numbers were chosen conservatively to ensure that the  $\Lambda$ CDM halo catalogues are complete down to those masses, which in practice was done by requiring that the HMF is in good agreement with the [Tinker et al. \(2010\)](#) analytical fitting formula.

Three plots of the mass ratio  $M_{\text{dyn}}/M_{\text{true}}$  as a function of the halo mass  $M_{\text{true}}$  are shown in Fig. 2.2, for the sifted Crystal F4, Jade F5 and Diamond F6.5 data for redshifts 0, 0.43 and 0.95 respectively. These include the extremes in both field strength and redshift. Each black data point corresponds to an individual halo. In each plot a majority of the haloes lie along a dark band of points that is asymptotic at ratios 4/3 and 1. The asymptote at ratio 1 corresponds to  $M_{\text{dyn}} = M_{\text{true}}$ , which holds for higher-mass haloes whose self-screening is sufficient to completely remove the enhancement due to the fifth force. The asymptote at 4/3 represents the maximum possible enhancement to  $M_{\text{dyn}}$ , and therefore results for haloes in a relatively empty environment and with mass low enough that there is effectively no self-screening of the fifth force.

For F5 many points are found below the dark band. These correspond to haloes that have most likely experienced environmental screening due to nearby more massive haloes, such that chameleon suppression of the fifth force is active even though the halo mass itself might not be great enough for self-screening. The effect of environmental screening in F5 is weak enough that the dark band of data only traces haloes for which self-screening

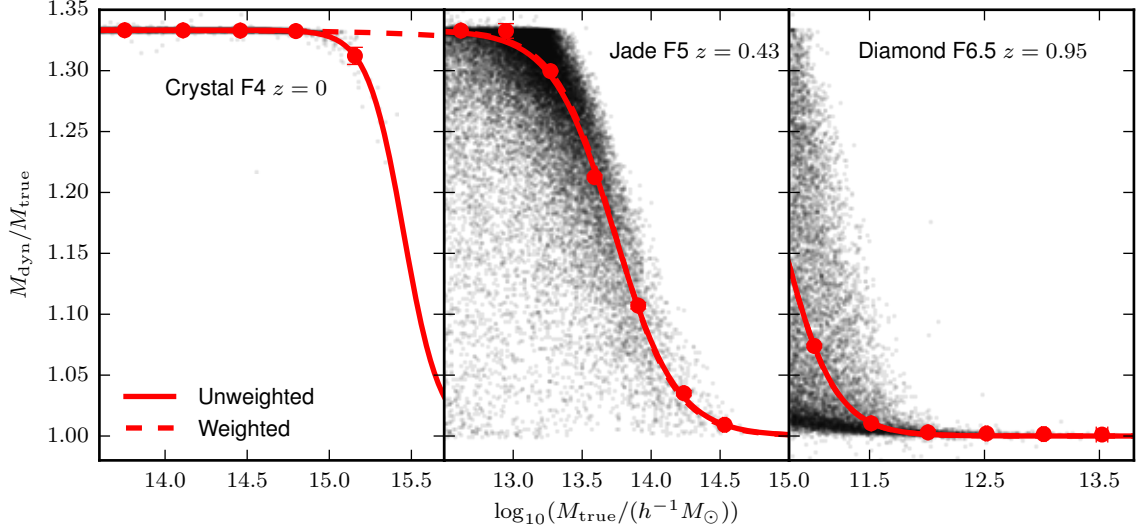


Figure 2.2: The ratio of the dynamical mass to the lensing mass versus the lensing mass for the dark matter haloes generated from  $N$ -body simulations run with modified ECOSMOG simulations for HS  $f(R)$  gravity. *From left to right:* Crystal simulation with  $z = 0$  for F4; Jade simulation with  $z = 0.43$  for F5; Diamond simulation with  $z = 0.95$  for F6.5. The simulation specifications are provided in Table 2.1. Unweighted (*solid*) and weighted (*dashed*) least squares fits of Eq. (2.17) are plotted over the data. These are generated using mass bins represented by the mean mass and the median ratio, shown by the red points. These points and their one standard deviation error bars are produced using jackknife resampling. For jackknife errors less than  $10^{-4}$ , we replace these with half of the 68% width of the data, between the 16th and 84th percentiles (see main text, below).

dominates over environmental screening. In F4, few data points are observed below the band because environmental screening is less effective in stronger background fields. For F4 and F5, apart from numerical noise, no data points are found to lie above  $4/3$  which is the maximally-allowed dynamical mass enhancement in  $f(R)$  gravity. In F6.5 the dark band of data is observed to have lower enhancement, with many data points found above it, particularly at  $M_{\text{true}} \leq 10^{11.5} h^{-1} M_{\odot}$ . With such low field values and halo masses in this mass range in F6.5, environmental screening is now able to begin to dominate over self-screening, which means the dark band of data no longer traces the haloes with self-screening only, as it did for F4 and F5. This is why it is now possible to find haloes above the main trend, as these simply correspond to haloes in emptier environments. Note that the upper bound of  $4/3$  applies also in this case.

In order to extract a trend for this data, the haloes are grouped into a set of equally-spaced logarithmic mass bins, which effectively cover the full range of halo masses under

consideration for a given model and snapshot. For each bin, the mean halo mass is measured along with the median ratio  $M_{\text{dyn}}/M_{\text{true}}$  among all haloes. The data in each bin approximately follows a lognormal distribution, and the median is expected to yield an appropriate ratio from within the main band of data. While we do not provide a detailed modelling of the distribution of  $M_{\text{dyn}}/M_{\text{true}}$  here, we have provided a simple model for the root-mean-square scatter in Chapter 5.

In the absence of multiple realisations of the data, the errors on the mean halo mass and median  $M_{\text{dyn}}/M_{\text{true}}$  in the bins are evaluated using jackknife resampling, in which the data is randomly split into 150 sub-volumes at each snapshot. By systematically excluding one sub-volume at a time, 150 resamples are created. For each resample, the haloes are split into the same set of mass bins, and 150 median ratios  $M_{\text{dyn}}/M_{\text{true}}$  and 150 mean masses are measured for each bin. Following the procedure outlined by [Norberg et al. \(2009\)](#), the errors in the median ratio and mean mass are generated by taking the square root of the variance of the 150 values, which has to be rescaled by a factor 149 to account for the lack of independence of the resamples.

The mass ratio data is quoted to 4 decimal places in the AHF output. Such precision can result in zero, or an unphysically small, variance being measured by the jackknife method. This can happen in unscreened or completely screened regimes where most of the data in the bin spans only a small range of ratios. Using the argument that the ratio errors must at least equal  $10^{-4}$ , any errors generated by jackknife which are less than this value are replaced with half of the width of the 68% range (in the bin under consideration), which spans from the 16th percentile to the 84th percentile. The percentile spread is most often used for lower-mass bins in strongly unscreened regimes, where the ratio data spans only a very small range. This ensures that the errors for these bins become a reasonable size relative to the errors of the other bins, which are estimated by jackknife, though rigorously speaking the 68% range is more of a description of the spread of the mass ratio rather than sample variation of the median ratio as jackknife gives. As discussed below, in the main results of this chapter we do not use the error bars estimated using this combination of jackknife and the 68% range.

The results for these bins are shown in Fig. 2.2, plotted over the raw data. To account for the asymptotic nature of the data, we fit the following tanh curve:

$$\frac{M_{\text{dyn}}}{M_{\text{true}}} = \frac{7}{6} - \frac{1}{6} \tanh \left( p_1 \left[ \log_{10} \left( M_{\text{true}} M_{\odot}^{-1} h \right) - p_2 \right] \right). \quad (2.17)$$

The two constants  $7/6$  and  $1/6$  are used to ensure the function remains between fixed asymptotes at ratios  $4/3$  and  $1$ . The parameters  $p_1$  and  $p_2$  represent, respectively, the inverse width of the mass transition and the mass logarithm at the centre of the transition.

For the dashed line, the parameters have been optimised through weighted least squares: the minimisation of the sum of the squared normalised residuals, where the normalisation is equal to the size of the error bars. For F5 and F6.5, this fit of Eq. (2.17) shows excellent agreement with the bin data, however for F4 the fit shows poor agreement with the result for the highest mass bin. This is because the error bar of this bin is substantially greater than those of the lower-mass bins, and it contributes very little weight in the optimisation. Weighted least squares therefore over-estimates the value of  $p_2$  for this snapshot, as the tanh curve starts to drop at a higher mass than the raw data. In contrast, the data point in the second highest mass bin has a much smaller error and it slightly overshoots a tanh curve which would perfectly go through the highest mass data point (the solid line, see below). Note that the same happens to the second and third lowest-mass data points for F5 (the middle panel of Fig. 2.2), but in this case there are four other data points at higher masses which dominate the optimisation, resulting in a good visual agreement between the dashed curve and the data points. This indeed highlights the importance of having data points which cover the full transition of the tanh curve in order to fit  $p_1$  and  $p_2$  accurately. Furthermore, the observation that the second lowest mass point for F5 lies above the tanh curve is quite generic and happens in most other plots where the curve starts to deviate from  $4/3$ , implying a slight insufficiency in the tanh fitting (we will comment on how this affects the fitted values of  $p_1$  and  $p_2$  below).

On the other hand, for the solid line in Fig. 2.2 the parameters have been optimised via unweighted least squares: the minimisation of the sum of the squared residuals, which have equal weights for all bins now. Since it does not suffer from the same issues as described above for the weighted fitting, this fit shows better agreement with the data point of the

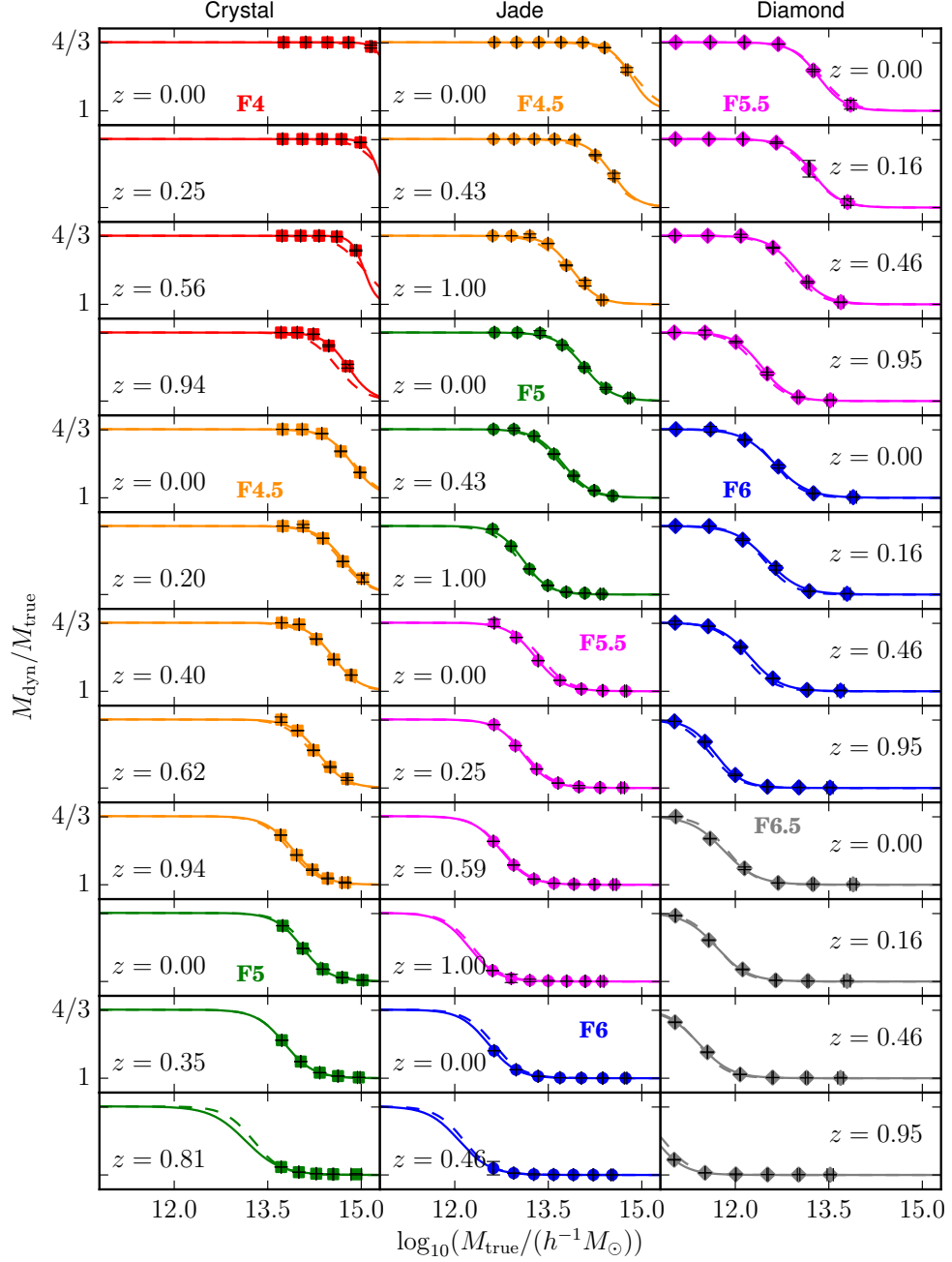


Figure 2.3: Dynamical mass to lensing mass ratio as a function of the lensing mass for F4 (red), F4.5 (orange), F5 (green), F5.5 (magenta), F6 (blue) and F6.5 (grey) at various redshifts as annotated. The halo data comes from the Crystal (left column), Jade (middle column) and Diamond (right column) simulations (see Table 2.1). The data points, corresponding to mass bins represented by their median ratio and mean mass, and their one standard deviation error bars are produced using jackknife resampling. Jackknife errors less than  $10^{-4}$  are replaced with half of the range between the 16th and 84th percentiles. *Solid line*: Eq. (2.17) with  $p_1$  and  $p_2$  determined by unweighted least squares fitting for the given snapshot; *Dashed line*: Eq. (2.17) with best-fit constant  $p_1$  result ( $p_1 = 2.21$ ) and linear  $p_2$  result (Eq. (2.20)) from Fig. 2.5 and Fig. 2.4, respectively.



highest mass bin of F4, while elsewhere shows equally good agreement as weighted least squares.

As discussed above, in the completely screened or unscreened regimes there is very little variation of the mass ratio and therefore the resulting uncertainties — by using either Jackknife resampling or the 68% range — for mass bins in those regimes are extremely small. Together with the facts that in many snapshots (e.g., the left panel of Fig. 2.2) the data points only cover part of the transition of the tanh curve and that the lower-mass bins can contain around three orders of magnitude more haloes than the higher-mass bins, this makes it challenging to find a consistent way to estimate uncertainties in all mass bins across all models/snapshots. Since the inhomogeneous sizes of error bars in the data points can lead to clearly unphysical fitting results, as shown in the dashed lines of the left panel of Fig. 2.2, the main results of this chapter shall be given using the unweighted least squares approach. We have tried a number of different ways to assign data error bars, including setting a lower limit such as  $10^{-4}$  to the individual errors, which all involve certain degrees of arbitrariness (for example, the 68% range to get error bars in Fig. 2.2 is really a characterisation of the spread of the data rather than an uncertainty of the median, and it is used solely to avoid very small uncertainties for some mass bins). Perhaps more importantly, the different ways of estimating uncertainties for the weighted least squares approach that we have tried all lead to similar fitting results of  $p_1, p_2$  as functions of  $\bar{f}_R(z)/(1+z)$  (the topic of the next sub-subsection), and the situation depicted in the left panel of Fig. 2.2 happens only for a few snapshots. As an example for reassurance, in Appendix A.1 we present fitting results of  $p_1$  and  $p_2$  using the weighted least squares approach with the error bars estimated as in Fig. 2.2, which confirms that this different approach does not significantly affect the final result.

For each snapshot in the investigation, five mass bins were used for Crystal, seven for Jade and six for Diamond, as these are the maximum possible numbers of bins such that there are a minimum of five haloes in almost all bins. We have checked different bin numbers, and this combination of bin numbers was also found to yield the smoothest results.

### 2.4.2.2 Fitting of $p_1, p_2$ as functions of $\bar{f}_R(z)/(1+z)$

By carrying out a fitting of Eq. (2.17) for all snapshots of all models, the field and redshift dependence of  $p_1$  and  $p_2$  can be tested. To understand what should be plotted, Sec. 2.2 and in particular the approximations for  $M_1$  and  $M_2$ , given by Eqs. (2.6) and (2.7), are used. From the way that  $p_1$  and  $p_2$  have been defined, the following can be shown:

$$p_1(z, \bar{f}_R) \propto \frac{1}{\log_{10}(M_1) - \log_{10}(M_2)} = \text{const}; \quad (2.18)$$

$$p_2(z, \bar{f}_R) = \frac{\log_{10}(M_1) + \log_{10}(M_2)}{2} = \frac{3}{2} \log_{10} \left( \frac{|\bar{f}_R|}{1+z} \right) + \text{const}. \quad (2.19)$$

Eqs. (2.6) and (2.7) have been used to bring in the  $z$  and  $\bar{f}_R(z)$  dependences. Eq. (2.19) implies  $p_2$  should have a linear trend as a function of  $\log_{10} \left( \frac{\bar{f}_R}{1+z} \right)$  with a slope of 1.5. This comes from the power 3/2 in Eqs. (2.6) and (2.7), where it in turn stems from the 2/3 power in  $\Psi_N \propto M^{\frac{2}{3}}$  for the Newtonian potential given by Eq. (2.2). On the other hand Eq. (2.18) implies  $p_1$  has no dependence on  $z$  and  $\bar{f}_R$  apart from through higher order effects, such as the non-sphericity of haloes, non-uniformity of the mass distributions within haloes, environmental screening, etc. Due to the simplicity of our thin-shell modelling, here we shall not attempt to include these higher-order effects. Indeed, under the thin-shell approximation, using Eqs. (2.6, 2.7, 2.8), it is found that the intercept of  $p_2$  in Eq. (2.19) only depends on  $\epsilon$ ,  $G$  (there is no dependence on  $H_0 = 100h \text{ kms}^{-1}\text{Mpc}^{-1}$  since the  $h$  is absorbed into the unit of  $10^{p_2}$ ,  $h^{-1}M_\odot$ ) and  $\Delta$ , and  $p_1$  depends only on  $\epsilon$ ; neither depends on the cosmological parameters, whose effects are completely in determining  $\bar{f}_R(z)$ . We will find later that  $p_1$  is indeed very weakly dependent on  $\bar{f}_R(z)/(1+z)$ . We also show that this dependency can be safely ignored without significantly affecting the value of the ratio  $M_{\text{dyn}}/M_{\text{true}}$ .

A potential issue arises from the limitations of the mass range covered by a particular set of data. As can be seen from Fig. 2.3, the mass bins are located almost entirely in the unscreened regime for F4 at low redshifts, while for high redshift Crystal F5, Jade F6 and Diamond F6.5 the mass bins are mostly found in the completely screened regime. As will be discussed in Figs. 2.4 and 2.5 of Sec. 2.5, the latter can result in under-estimation of

the  $p_1$  and  $p_2$  values, and we have already seen in Fig. 2.2 how, depending on the choice of fitting procedure,  $p_2$  can be over-estimated for F4 at low redshift.

To understand why the parameters are affected in such a manner, consider the scenario where all mass bins are located at ratio  $4/3$ . As can be seen in the F4 and F5 panels of Fig. 2.2, the median ratio data from the simulations in this regime is almost completely flat, so a tanh fit will predict a turning point at a mass higher than is actually the case, and so  $p_2$  will be over-estimated. This flatness of the raw data in the unscreened regime is particularly evident in the F5 panel, where the second data point from the left ends up above the trend line, despite having a negligible error, at the same height as the first data point (this suggests that this region of the data cannot be fitted perfectly by a tanh curve). On the other hand, for mass bins at high-redshift snapshots and for low field strengths (F6.5 - F5), where almost all of the data points lie at a ratio of 1, because the data here is flatter than predicted by Eq. (2.17) the turning point at ratio 1 will thus be predicted at lower mass, leading to an under-estimation of  $p_2$ . The effect on  $p_1$  turns out to be similar to  $p_2$ , but is even more sensitive to these limitations.

The issues presented here were the main motivation for using data from simulations with differing resolutions. To prevent such dubious estimations of  $p_1$  and  $p_2$  from adversely affecting the main results, a strict criterion is enforced: we only trust  $p_1$  and  $p_2$  values that have been calculated using snapshots for which the mass bins enclose at least half of the height of the mass ratio transition (a median ratio range of  $1/6$  or greater).

## 2.5 Results

As mentioned above, a fitting function for the ratio  $M_{\text{dyn}}/M_{\text{true}}$  that works for general scalar field strength  $f_{R0}$  and redshift  $z$  should be calibrated and validated against full numerical simulations with a large dynamical range of halo masses in order to maximally cover the transition between screened and unscreened regimes, which itself varies strongly with  $z$  and  $f_{R0}$ . However,  $N$ -body simulations are known to have a limited dynamical range and it is also too expensive to run full simulations for too many  $f_{R0}$  values. Our recipe to tackle the former challenge is to combine a suite of simulations with varying resolutions

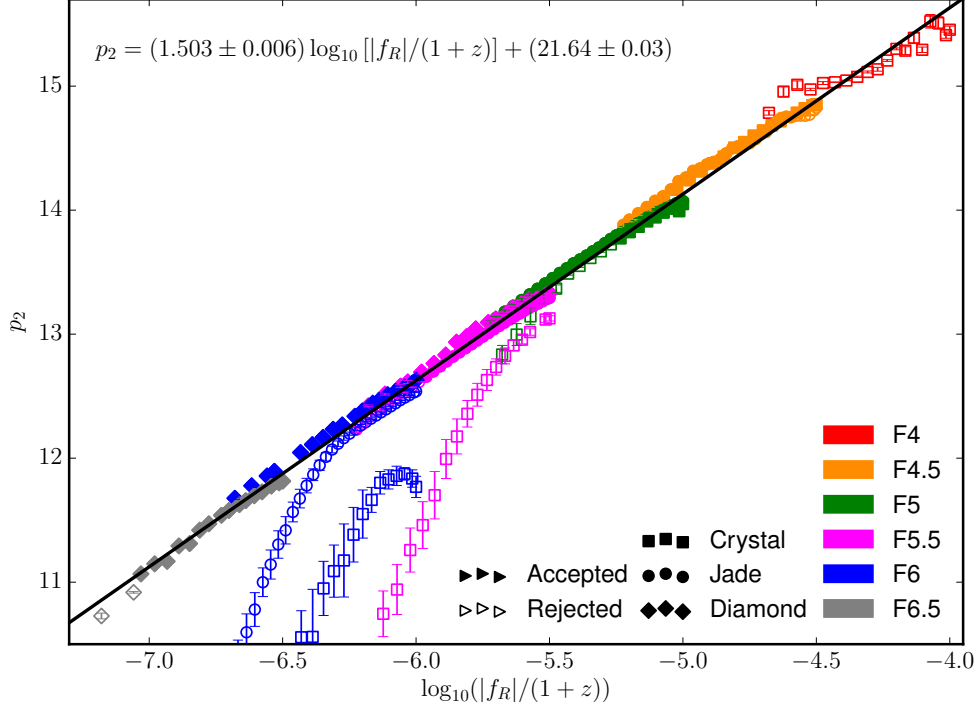


Figure 2.4: Parameter  $p_2$  in Eq. (2.17) plotted as a function of the background scalar field at redshift  $z$ ,  $\bar{f}_R(z)$ , divided by  $(1+z)$ , for several present day field strengths  $f_{R0}$  (see legends) of HS  $f(R)$  gravity.  $p_2$  is measured via an unweighted least squares optimization of Eq. (2.17) to data from modified ECOSMOG simulations, described by Table 2.1, at simulation snapshots with redshift  $z < 1$ .  $\bar{f}_R(z)$  is calculated for each snapshot using Eq. (1.15). The trend line has been produced via a weighted least squares linear fit, using the one standard deviation error bars, of the solid data points, which correspond to snapshots for which the mass bins contain at least half of the median mass ratio range 1 to 4/3. The hollow data does not meet this criterion, so is deemed unreliable and neglected from the fit, which is given by Eq. (2.20).

(Crystal, Jade and Diamond) to increase the halo mass range, while for the latter issue we have introduced the low-cost ‘silver’ simulations (see Sec. 2.4.1). Both approaches need to be explicitly checked to guarantee validity and consistency. Furthermore, in Sec. 2.4.2 we have discussed subtleties in the tanh curve fitting such as the weighted and unweighted least squares approaches. In this section we give the main results on  $p_1$  and  $p_2$  from using this methodology, for unweighted least squares, and leave various consistency checks to the Appendices. In Appendix A.1 we compare with results from using the weighted least squares approach as a double check, and in Appendix A.2 we check the use of ‘silver’ data and the combination of the Crystal, Jade and Diamond simulations.

A plot of  $p_2$  as a function of  $\log_{10} \left( \frac{|\bar{f}_R|}{1+z} \right)$  is shown in Fig. 2.4. A linear trend is fitted using the filled data points, which correspond to snapshots for which the mass bins enclose a median ratio range of 1/6 or greater. The motivation for this criterion is discussed in Sec. 2.4.2. The filled data points are expected to give a reasonable estimate for the logarithm of the mass at the centre of the transition, and they all turn out to lie along a clear linear trend in Fig. 2.4. The result of the linear fit, found using the one standard deviation error bars, is:

$$p_2 = (1.503 \pm 0.006) \log_{10} \left( \frac{|\bar{f}_R|}{1+z} \right) + (21.64 \pm 0.03). \quad (2.20)$$

The gradient of  $1.503 \pm 0.006$  shows excellent agreement with the theoretical prediction of 1.5 from Eq. (2.19).

Many of the hollow data points are observed to be peeling off the trend, particularly in the F6 and F5.5 models. These snapshots correspond to cases in which all mass bins are found in the totally screened regime, resulting in an under-estimation of the centre of the transition as discussed in the previous section. This behaviour provides no useful information about the dynamical mass enhancement, but rather it tells us that a higher resolution simulation, with lower-mass particles to probe haloes of lower mass, is required. For F5.5 the peeling-off corresponds to Crystal data, whereas the higher resolution Jade and Diamond simulations produce linear data. For F6 both the Crystal and Jade data peel off from the linear trend, as only Diamond has a high enough resolution to probe unscreened haloes in F6. Diamond turns out to have a sufficient resolution to effectively examine F6.5 as well, although a couple of high redshift snapshots do not get used in the linear fit, suggesting these are on the boundary between reliable and untrustworthy data. F6.5 nevertheless agrees with the linear behaviour of the rest of the filled data.

A relatively noisy trend is observed in the F4 data (though the data points all reasonably follow the linear trend), probably because each snapshot only has one or two mass bins lying within the mass range where the ratio  $M_{\text{dyn}}/M_{\text{true}}$  undergoes a transition between 1 and 4/3. Most bins lie in the unscreened regime, such that none of the snapshots in F4 satisfy the selection criterion to be included in the linear fit – all data points for F4 are hollow in Fig. 2.4. An improvement of this result would require a simulation with a

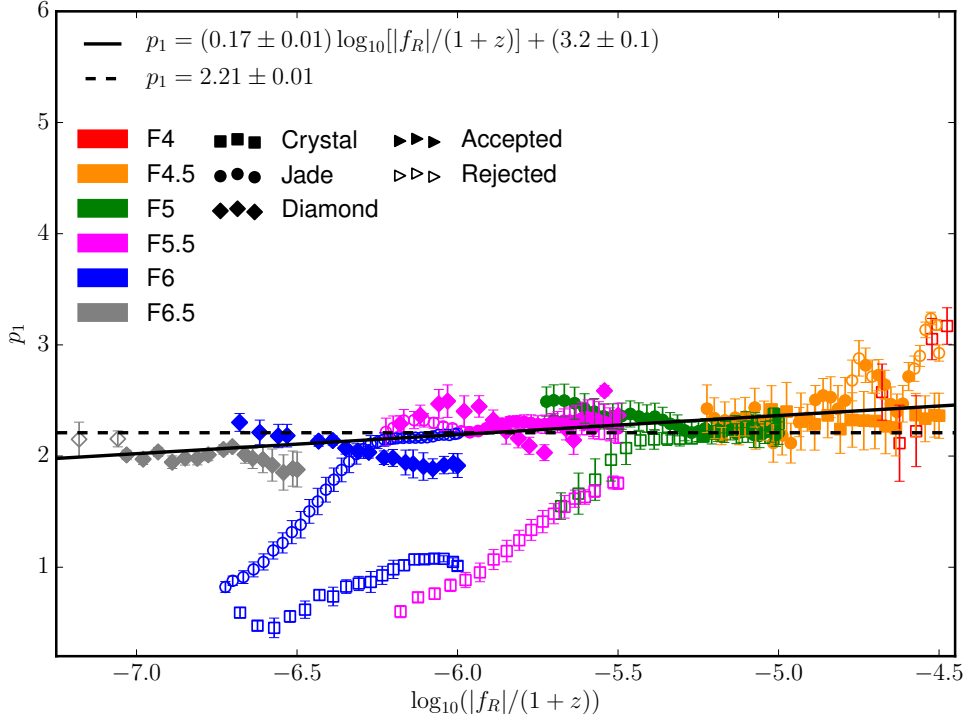


Figure 2.5: Parameter  $p_1$  in Eq. (2.17) plotted as a function of the background scalar field at redshift  $z$ ,  $\bar{f}_R(z)$ , divided by  $(1+z)$ , for several present day field strengths  $f_{R0}$  (see legends) of HS  $f(R)$  gravity.  $p_1$  is measured via an unweighted least squares optimisation of Eq. (2.17) to data from modified ECOSMOG simulations, described by Table 2.1, at simulation snapshots with redshift  $z < 1$ .  $\bar{f}_R(z)$  is calculated for each snapshot using Eq. (1.15). Weighted least squares linear (*solid line*) and constant (*dashed line*) fits, using the one standard deviation error bars, of the solid data points, which correspond to snapshots for which the mass bins contain at least half of the median mass ratio range 1 to 4/3, are shown. The hollow data points do not meet this selection criterion, and therefore are deemed unreliable and neglected from the fits, which are given by Eq. (2.21) and  $p_1 = (2.21 \pm 0.01)$ , respectively.

sufficiently large box size to include more haloes at the higher masses necessary to properly examine screening in F4.

The corresponding plot for  $p_1$  is shown in Fig. 2.5. The trend is more complicated than that of  $p_2$ , partly because the thin-shell model result described in Eq. (2.18) predicts no dependence of  $p_1$  on  $\bar{f}_R$  and  $z$ , while dependence can still be introduced through effects such as environmental screening which are harder to model. However, we expect that these effects have a relatively small impact, and indeed, an approximately flat trend of  $p_1$  is observed. The results are noisier here than in Fig. 2.4 for  $p_2$ , because the width of the mass transition requires a greater range of halo masses for a tanh fit to be reliable. The

criterion for selecting snapshots in the fit of  $p_1$  is the same as for  $p_2$ , and again only the solid data points which satisfy this criterion are fitted. This rules out all of the data from the F4 model (which produces a wild trend here that is left out of the plot area), and several snapshots from other models.

The result for the constant  $p_1$  fit, as predicted by Eq. (2.18), is  $p_1 = (2.21 \pm 0.01)$  and is shown by the dashed line in Fig. 2.5. A linear model was also fitted, shown by the solid line, yielding the following result:

$$p_1 = (0.17 \pm 0.01) \log_{10} \left( \frac{|\bar{f}_R|}{1+z} \right) + (3.2 \pm 0.1). \quad (2.21)$$

These trends have been fitted using the one standard deviation error bars. The gradient of  $0.17 \pm 0.01$  is small, though not in agreement with the prediction of a flat trend. With a theoretical modelling which neglects effects such as environmental screening, a small gradient nevertheless seems like a reasonable result. Being able to accurately predict the width of the mass transition is not as important as being able to predict the central mass of the transition, because the tanh curve is less sensitive to  $p_1$  than to  $p_2$  (which can be easily checked). Almost all the data points observed to be significantly peeling off from the horizontal band of data in Fig. 2.5 (including Jade and Crystal F6, Crystal F5.5 and some of Jade F4.5) fail to satisfy the selection criterion. This is further evidence that these particular trends are indeed caused by the limitations of the simulation resolution. Also, a comparison of Figs. 2.5 and A.3 shows that the use of an unweighted approach to measure  $p_1$  produces the smoother trend in the  $p_1$  data.

The quality of the above fits for  $p_1$  and  $p_2$  as well as the validity of the theoretical predictions, given by Eqs. (2.18) and (2.19), can be assessed by examining Fig. 2.3. The solid lines represent the exact fits produced in the unweighted least squares optimisation of Eq. (2.17) to each snapshot of data. The dashed lines are plotted using Eq. (2.17) and the  $p_1$  and  $p_2$  values that are predicted using the constant fit of Fig. 2.5 (dashed line) and the linear fit of Fig. 2.4 (solid line) respectively. Noticeable disparities between the dashed line and solid line fits are observed in the F4 data, resulting from the relatively flat trend produced by the raw data in unscreened regimes and the limited number of haloes in Crystal covering the high masses necessary for properly examining the transition to complete screening in

F4. The agreement between the dashed and solid lines in Fig. 2.3 generally improves if one uses the linear fit predictions for  $p_1$ , although we only use the constant fit here, which is motivated by our theoretical modelling. Nevertheless, in general the dashed line fits show excellent agreement with the simulation data over the full range of redshifts and models that are plotted in Fig. 2.3, implying that Eq. (2.17) can be treated as a general formula when using our constant and linear fits of  $p_1$  and  $p_2$  respectively.

### 2.5.1 Potential implications

Although they are not directly related to the preparation for cluster constraints, we make the following interesting observations in the results of this section, mainly Fig. 2.4.

First, the solid straight line in Fig. 2.4 represents the logarithm of the halo mass,  $\log_{10} M_{\text{true}}$ , at the centre of the transition of the median of  $M_{\text{dyn}}/M_{\text{true}}$ , and it roughly separates the haloes into two parts – a screened sample ( $\log_{10} M_{\text{true}}$  well above the line) and an unscreened sample ( $\log_{10} M_{\text{true}}$  well below the line). From Figs. 2.2 and 2.4 we notice that even at  $|\bar{f}_R(z)|/(1+z) = 10^{-7}$ , corresponding to a strongly screened model, about half of the haloes (with high ratio  $M_{\text{dyn}}/M_{\text{true}}$ ) with mass  $M_{\text{true}} \sim 10^{11} h^{-1} M_{\odot}$  are unscreened, and these are haloes which are likely to reside in under-dense regions. The other half of these haloes (with low ratio  $M_{\text{dyn}}/M_{\text{true}}$ ) are screened, aided by their environments, implying the importance of environmental screening. It would certainly be interesting to see if this linear trend goes to even smaller values of  $|f_R(z)|/(1+z)$ , which will tell us whether dwarf galaxy haloes can be environmentally screened for those field values. This will be relevant for astrophysical tests of  $f(R)$  gravity (e.g., Jain et al., 2013; Vikram et al., 2013; Sakstein et al., 2014).

Second, it is interesting that the screening of haloes in models with different  $f_{R0}$  can be well described by a single parameter:  $\bar{f}_R(z)/(1+z)$ . This implies that the theoretical modelling of various other properties in  $f(R)$  gravity can perhaps be simplified into a one-parameter family of description and therefore may have profound theoretical and practical implications.



## 2.6 Summary, discussion and conclusions

The global properties of galaxy clusters, such as their abundance and clustering on large scales, are sensitive to the strength of gravity and can be predicted accurately using cosmological simulations. They therefore offer a powerful means of testing alternative models of gravity, including  $f(R)$  gravity, on large scales. In order to utilise the wealth of information being made available through current and upcoming galaxy cluster surveys, it is important to ensure that numerical predictions are prepared that can be directly confronted to the observational data. This includes accounting for various sources of theoretical bias, such as the enhancement of the dynamical mass of galaxy clusters resulting from the presence of the fifth force in unscreened  $f(R)$  gravity. This effect is currently not included in the derivations of scaling relations used to determine the cluster mass. The best means of correcting this would be through a re-calibration of the scaling relations which are better understood in  $\Lambda$ CDM, and make them work in the context of MG, which requires an understanding of the relationship between the dynamical mass and lensing mass. However, previous studies of this relationship in the literature are specific and do not include a general formula that can be applied to arbitrary model parameters and redshifts.

We have found a simple model to describe the relationship between the dynamical mass and lensing mass of dark matter haloes in the HS  $f(R)$  model. As shown by the solid line fits of Fig. 2.3, the tanh fitting formula of Eq. (2.17) has generally shown excellent agreement with AHF halo data, for  $z < 1$ , from three ECOSMOG DMO simulations, which are summarised in Table 2.1. By taking advantage of the variety of resolutions offered by these simulations, and using  $\Lambda$ CDM simulations to produce approximate data for field strengths not covered by the  $f(R)$  gravity simulations, the validity of Eq. (2.17) has been probed vigorously across a wide and continuous range of field values that cover  $10^{-6.5} < |f_{R0}| < 10^{-4}$  within  $z < 1$ .

In addition, we have used a simple thin-shell model (Sec. 2.2.1) to predict the behaviours of free parameters  $p_1$  and  $p_2$  in Eq. (2.17), which characterise the inverse width and the central logarithmic mass of the tanh-like transition respectively. The predictions,

which neglect the effects of environmental screening due to nearby dark matter haloes, are given by Eqs. (2.18, 2.19). Using a stringent criterion to exclude unreliable snapshots in the fitting, the result for  $p_2$ , shown in Fig. 2.4, is given by Eq. (2.20). The slope value of  $1.503 \pm 0.006$  shows excellent agreement with the prediction of 1.5 by Eq. (2.19), and the data of Fig. 2.4 shows a clear linear trend as predicted. As shown by Fig. 2.5, the  $p_1$  data is more scattered, but given the size of the one standard deviation error bars, the constant trend predicted by Eq. (2.18) is not unreasonable, resulting in  $p_1 = (2.21 \pm 0.01)$ . As shown by the dashed line fits of Fig. 2.3, these results for  $p_1$  and  $p_2$  show good agreement with the simulation data across the full range of field values and redshifts. We have also repeated the analysis using a different approach to utilise the errors in the simulation data, and the results, shown in Appendix A.1, also agree with the thin-shell model prediction very well. In Appendix A.2 we further argue that the results in this chapter apply to models with different cosmological parameters such as  $\sigma_8$  and  $\Omega_M$ .

On the other hand, although we make a very specific choice of  $f(R)$  gravity in this chapter, the theoretical model and the procedure we followed to calibrate it are expected to be applicable to general chameleon gravity theories (Gronke et al., 2015, 2016). As discussed briefly in Appendix A.2, in other  $f(R)$  models the transition between screened and unscreened regimes can be different from the Hu and Sawicki (2007) model with  $n = 1$ , which may cause the exact fitted values of  $p_i$  to differ from what we presented in the above. Therefore, other  $f(R)$  models may require a re-calibration based on simulations. However, given that all  $f(R)$  models are phenomenological, it is perhaps more sensible to focus on a representative example, such as that by Hu and Sawicki (2007), to make precise observational constraints. The pipeline and methodology can then be applied to any other models following general parameterisation schemes (e.g., Brax et al., 2012b,a; Lombriser, 2016), which are useful for capturing the essential features of large classes of models using a few parameters. Should a preferred one emerge, the conclusion for the HS model can serve as a rough guideline as to what level future cluster observations can constrain scalar-tensor-type screened theories. For this reason we decide not to explore other forms of  $f(R)$  in this thesis.

A generic fitting function for the relationship between the dynamical and lensing masses of dark matter haloes is an essential ingredient of the new framework proposed in this thesis, to carry out cosmological tests of gravity in an unbiased way. Taking Eq. (2.16) as an example, our general formula for the dynamical mass enhancement allows us to incorporate this particular effect of  $f(R)$  gravity into galaxy cluster scaling relations in a self-consistent way. A key benefit of a fitting function is that it allows a continuous search through the model parameter space without having to run full simulations for every parameter point sampled in MCMC. The results will also be useful for other cluster tests of gravity that employ the difference between dynamical and lensing masses, such as by comparing cluster dynamical and lensing mass profiles, or by looking at measured cluster gas fractions.

The results presented in this chapter indicate that a simple model sometimes works surprisingly well despite the greatly simplified treatment of the complicated nonlinear physics of (modified) gravity. It naturally raises the following question: can other theoretical or observational properties of dark matter haloes also be modelled accurately, based on a simplified physical picture and calibrated by numerical simulations? An example is the relationship between the masses and density profiles of haloes, as mentioned in Sec. 2.3. This concentration-mass relation is critical for converting between the different halo mass definitions commonly used in different communities, and a great deal of effort has been made to explain it in the standard  $\Lambda$ CDM model, while in MG models, such as  $f(R)$  gravity, the understanding is still purely numerical and confined to a limited few cases. We will explore this issue in Chapter 3.

Throughout the analysis in this chapter, we used DMO simulations. The method to rescale the  $\Lambda$ CDM cluster scaling relations to get scaling relations that apply to MG (He and Li, 2016), has been tested and validated using non-radiative hydrodynamical simulations. In Sec. 2.3 we argued that adding the full baryonic physics in the simulations will not substantially change the conclusion, based on previous work on  $\Lambda$ CDM full physics simulations. We will check this using full hydrodynamical simulations for HS  $f(R)$  gravity in Chapter 4.

Finally, we note again that a key ingredient of any test of gravity using the cluster abundance is the ability to predict the HMF for arbitrary model parameters. In this thesis, we have used the recently-developed HMF fitting formula by [Cataneo et al. \(2016\)](#), which was calibrated using a subset of simulations (Crystal) used in this chapter. This formula has 3-5% accuracy for a range of  $f_{R0}$  values between F4 and F6 and for halo masses above  $10^{13}h^{-1}M_{\odot}$ , making it ideal for comparing with observed cluster abundances. A full hydrodynamical simulation can also be useful in understanding how the predicted abundance of dark matter haloes can change with the inclusion of baryonic physics.

We will test the above-mentioned framework, which incorporates these effects into model predictions and allows for detailed MCMC searches of the parameter space, in Chapter 5. Here, the fitting function for  $M_{\text{dyn}}/M_{\text{true}}$  is also useful for constructing mock observational data that are used to validate the MCMC model constraint pipelines.

# A universal model for the halo concentration in $f(R)$ gravity

## 3.1 Introduction

In Chapter 2, we presented a general analytical model for the enhancement of the dynamical mass. This is an important component of our framework for  $f(R)$  gravity constraints using cluster number counts (Fig. 2.1): it is used to predict the dynamical masses of dark matter haloes in mock catalogues and to convert cluster observable-mass scaling relations that have been calibrated in  $\Lambda$ CDM into a form that works in  $f(R)$  gravity. Another important aspect of our framework is the ability to make conversions between halo masses corresponding to different overdensities  $\Delta$ . For example, in order to constrain  $f(R)$  gravity using the cluster abundance, we require a model-dependent calibration of the HMF, which quantifies the number density of dark matter haloes per unit mass interval,  $dn_{\text{halo}}/dM_{\Delta}$ . Our choice of HMF (Cataneo et al., 2016) has been calibrated for overdensity  $\Delta = 300\Omega_{\text{M}}(z)$ , while overdensity  $\Delta = 500$  is more generally used in cluster surveys. Therefore in order to make constraints using observational data that has been calibrated for overdensity  $\Delta = 500$ , it will be necessary for us to apply the conversion  $M_{300\text{m}} \rightarrow M_{500}$  to the HMF. It is also likely that conversions to other overdensities will be required. For example  $\Delta = 2500$  is also sometimes used in observational surveys, and  $\Delta = 200$  is often used in theoretical studies.

Therefore, a prediction for the conversion between halo masses corresponding to arbitrary values of  $\Delta$  is essential.

The halo mass measured for different overdensities corresponds to the total mass enclosed by different halo radii  $R_\Delta$ , where  $R_\Delta$  is larger for lower values of  $\Delta$ . Therefore conversions between the halo mass at different overdensities can be estimated if the density profile of dark matter haloes can be predicted. Typically the universal NFW density profile (Navarro et al., 1997) is assumed. This is a 2-parameter profile, but can be written with one parameter if the mass (or radius) for a particular overdensity is known. This parameter can be the halo concentration, and predicting it in  $\Lambda$ CDM as a function of the cluster mass and redshift has been the subject of much work over the two decades since it was first introduced (e.g., Bullock et al., 2001; Neto et al., 2007; Duffy et al., 2008; Dutton and Macciò, 2014; Ludlow et al., 2014).

In addition to facilitating mass conversions the concentration is also important in studies of the non-linear matter power spectrum (e.g., Brax and Valageas, 2013; Lombriser et al., 2014; Achitouv et al., 2016; Hu et al., 2018; Cataneo et al., 2018), which, like the cluster abundance, can also be used to probe dark energy and MG theories. The large-scale part of the matter power spectrum can, for example, be predicted using linear perturbation theory by incorporating the linear halo bias. On the other hand, the small-scale part of the matter power spectrum can be assembled using the HMF and the halo concentration. The concentration is necessary in order to predict the density profile, which is required, for example, in order to model the size of haloes.

In  $f(R)$  gravity, the concentration can become enhanced due to the effects of the fifth force on the density profile. For example, for an unscreened halo the in-falling particles experience a greater acceleration due to the stronger gravitational force, and this can alter the profile such that the density is raised at the inner regions and lowered at the outer regions. Therefore, the  $\Lambda$ CDM predictions of the concentration are unlikely to apply for lower-mass unscreened haloes in  $f(R)$  gravity. Yet there is no general quantitative model for the concentration in  $f(R)$  gravity that is discussed in the literature, which instead tends to focus on a more qualitative understanding of the effects of the fifth force on the

concentration and on the density profile (e.g., Zhao et al., 2011a; Lombriser et al., 2012; Shi et al., 2015; Arnold et al., 2016; Arnold et al., 2019a). Therefore, a  $\Lambda$ CDM relation for the concentration is often used in the literature. For example, the modellings of the non-linear matter power spectrum in HS  $f(R)$  gravity by Brax and Valageas (2013); Hu et al. (2018); Cataneo et al. (2018) use prescriptions for the concentration-mass relation that have been calibrated in  $\Lambda$ CDM. Also, due to the large scatter of the concentration-mass relation, some works (e.g., Cataneo et al., 2015) argue that it is fine to assume a fixed concentration for a sample of clusters that covers a sufficiently narrow mass range.

In order to prevent potential biases resulting from a simplified treatment of the concentration, the focus of this chapter is to produce a general model for the concentration in HS  $f(R)$  gravity (blue dotted box of Fig. 2.1) that may be applied in future studies. Rather than calibrating a relation for the absolute concentration, we decided to focus on finding a universal model for the enhancement of the concentration as a function of the halo mass and redshift. This has been achieved using data from a suite of DMO  $N$ -body simulations run for three models of HS  $f(R)$  gravity. Note that we define the enhancement as the ratio of the  $f(R)$  concentration to the concentration in GR. This means that one can select a  $\Lambda$ CDM concentration-mass-redshift relation from the literature (e.g., Bullock et al., 2001; Neto et al., 2007; Duffy et al., 2008; Dutton and Macciò, 2014; Ludlow et al., 2014) that they wish to use, then this can be converted into a form in HS  $f(R)$  gravity. Our model includes a dependence on the cosmological density parameters  $\Omega_M$  and  $\Omega_\Lambda$ , so any  $\Lambda$ CDM relation can be used regardless of the values of these parameters. Our model depends on the particular combination  $\bar{f}_R(z)/(1+z)$  where  $\bar{f}_R(z)$  is the background scalar field at redshift  $z$ , and does not explicitly depend on the model parameter  $f_{R0}$ , i.e., the present-day background scalar field value, as one would naively expect. This has the implication that predictions may be made for arbitrary values of  $f_{R0}$  and  $z$  (as long as the above combination is within the range of validity of our fitting). This generality of our model was achieved by combining data from the different  $f(R)$  gravity models by applying a simple transformation to the halo mass using the  $p_2$  parameter defined in Chapter 2, where  $M_{500} = 10^{p_2} h^{-1} M_\odot$  can be considered as the mass above which haloes are screened and below which haloes are

Parameters and models	Simulations			
	Diamond	Jade	Crystal	AREPO
box size / $h^{-1}\text{Mpc}$	64	450	1024	62
particle number	$512^3$	$1024^3$	$1024^3$	$512^3$
particle mass / $h^{-1}M_\odot$	$1.52 \times 10^8$	$6.64 \times 10^9$	$7.78 \times 10^{10}$	$1.52 \times 10^8$
$\Omega_M$	0.281	0.2819	0.281	0.3089
$\Omega_\Lambda = 1 - \Omega_M$	0.719	0.7181	0.719	0.6911
$h$	0.697	0.697	0.697	0.6774
$f(R)$ models	F6	F5	F4, F5, F6	F4, F5, F6

Table 3.1: Specifications of the three ECOSMOG simulations and the AREPO simulation used in this investigation. The ECOSMOG simulations are labelled Diamond, Jade and Crystal for convenience. All simulations have been run for  $\Lambda$ CDM in addition to the HS  $f(R)$  gravity models listed. The Hubble constant,  $H_0$ , is equal to  $100h \text{ kms}^{-1}\text{Mpc}^{-1}$  for each simulation.

unscreened.

This chapter is arranged as follows: Sec. 3.2 provides an overview of the DMO  $N$ -body simulations that are used in this chapter, along with an outline of the methods used to measure the concentration and its enhancement; Sec. 3.3 discusses the results of this chapter, including the general model for the concentration enhancement; and, finally, Sec. 3.4 summarises the main conclusions from this chapter, and outlines the next steps of our framework.

## 3.2 Simulations and methods

The simulations that we have used to study the concentration are presented in Sec. 3.2.1, along with the methods that we use to extract their halo data. Our methods to measure the concentration and its enhancement and a useful technique of rescaling the halo mass are discussed in Sec. 3.2.2.

### 3.2.1 Simulations

Our DMO simulations are shown in Table 3.1. Three of these — the Crystal, Jade and Diamond simulations — were also used in Chapter 2. We also include a simulation that has been run using the AREPO code (Springel, 2010) with its MG solver.



The cosmological parameters and technical specifications of the simulations are listed in Table 3.1. The AREPO simulation, which is the DM-only subset of the SHYBONE simulation suite (Arnold et al., 2019b), has the same mass resolution as the Diamond simulation, with particle mass  $1.52 \times 10^8 h^{-1} M_{\odot}$ , as well as a similar box size of  $62 h^{-1} \text{Mpc}$ . Together, the four simulations cover a wide range of halo masses, which is essential in order to comprehensively study the halo concentration across the full transition between the screened and the unscreened regimes. The AREPO data is particularly useful because it has been run for all three  $f(R)$  gravity models examined in this chapter. Its low particle mass allows low-mass, unscreened haloes to be studied in all three models, ensuring a more detailed exploration of the transition between the screened and unscreened regimes. In addition to this, the similar resolutions of the AREPO simulation and Diamond allow a consistency test of the ECOSMOG and AREPO simulations, which is necessary due to the potential disparities between the results from these two codes which employ different algorithms and assume different cosmological parameters.

The simulation data covers redshifts up to at least  $z = 1$  for all simulations. However redshifts  $z < 2$  and  $z < 3$  have been included for AREPO F5 and F4, respectively, as otherwise the data from these models would only cover the unscreened regime. The dataset consists of 19 snapshots from Crystal, 33 from Jade and 44 from Diamond. For AREPO there are 46, 37 and 24 snapshots from F4, F5 and F6, respectively.

The halo catalogues that we construct consist of dark matter haloes identified using the SUBFIND code (Springel et al., 2001) implemented in AREPO. This employs a standard friends-of-friends (FOF) algorithm to identify FOF groups (haloes) and a gravitational unbinding method to locate the bound substructures (subhaloes) within each group. The mass and radii of the haloes have been measured for overdensities  $\Delta = 500$  and 200 with respect to the critical density of the Universe. Two methods that can be used to calculate the halo concentration (see Sec. 3.2.2.1) also require measurements of the maximum circular velocity,  $V_{\text{max}}$ , and the corresponding orbital radius,  $R_{\text{max}}$ . The SUBFIND code calculates these quantities using just the bound particles of the central, dominant subhalo. To a good approximation, these measurements can be used to represent the maximum circular

velocity and the corresponding orbital radius for the entire FOF group. We have checked that within  $R_{\max}$  almost all particles are bound regardless of whether the fifth force is felt or not.

In order to accurately measure the halo concentration, it is important that the halo consists of enough particles so that it is well-resolved at both the inner and outer regions. Therefore in all of our analyses in this chapter we only use haloes with more than 1000 particles contained within  $R_{500}$ . This corresponds to minimum halo masses of  $M_{500} = (1.52 \times 10^{11}, 6.64 \times 10^{12}, 7.78 \times 10^{13}, 1.52 \times 10^{11}) h^{-1} M_{\odot}$  for Diamond, Jade, Crystal and AREPO respectively.

We bin our haloes by  $M_{500}$  (see Sec. 3.2.2.3), and our model for the enhancement of the concentration is designed to predict the concentration in  $f(R)$  gravity as a function of  $M_{500}$  (see Sec. 3.3.2). The reason for choosing overdensity  $\Delta = 500$  is to be consistent with Chapter 2, in which  $M_{500}$  was used to study the enhancement of the dynamical mass and, crucially, to define the parameter  $p_2$ . As will be discussed later (Sec. 3.2.2.2), the halo mass can be rescaled by this parameter in order to combine data from snapshots with different values of  $|\bar{f}_R|/(1+z)$ .

Various works in the literature which study the halo evolution and aim to model the concentration as a function of redshift and mass often use relaxed samples (e.g., Neto et al., 2007). For example, haloes which have undergone recent mergers are unlikely to give reliable estimates of the concentration. However, we have decided to include all haloes (that satisfy the above mass criteria) since one of the applications of our results will include matter power spectrum predictions, which requires the use of all haloes.

### 3.2.2 Methods

In this section we present the three approaches for measuring the halo concentration that are used in this chapter (Sec. 3.2.2.1), a useful rescaling of the halo mass (Sec. 3.2.2.2) and our method of binning the concentration and evaluating its enhancement (Sec. 3.2.2.3).

### 3.2.2.1 Concentration measurement

Three methods were considered for the measurement of the halo concentration. The most accurate is to directly fit the profiles of the haloes using the NFW profile (Navarro et al., 1997):

$$\rho(r) = \frac{\rho_s}{(r/R_s)(1 + r/R_s)^2}, \quad (3.1)$$

where  $\rho_s$  is the characteristic density and  $R_s$  is the scale radius. The concentration is defined as  $c_{200} = R_{200}/R_s$ . Note that the convention to define the concentration with respect to overdensity 200 is frequently used by the literature. Therefore we elected to use this definition, even though our haloes are required to be binned by  $M_{500}$  (see above). This approach is still consistent because as long as the concentration can be predicted for one overdensity, it can also be predicted for overdensity 500.

Radial bins that are equally spaced in logarithmic distance from the halo centre were used to ensure that both the inner and outer regions were equally well-fitted. As shown by Neto et al. (2007), the choice of radial range over which the profile is fitted can be important. Resolution effects can occur at the outer less-dense regions of the halo or the innermost regions where, due to the limited number of particles, the density can be underestimated. In order to avoid these effects, we chose to calculate the densities of 20 logarithmic radial bins, spanning distances  $0.05R_{200}$  to  $R_{200}$  from the halo centre, which is consistent with the range used by Neto et al. (2007). These densities were fitted using the formula,

$$\log_{10}(\rho) = \log_{10}(\rho_s) - \log_{10}(xc_{200}) - 2\log_{10}(1 + xc_{200}), \quad (3.2)$$

where  $x = r/R_{200}$ . This was achieved via unweighted least squares, where  $\rho_s$  and  $c_{200}$  are allowed to vary independently. The halo concentration was set equal to the optimal value of  $c_{200}$ .

The concentration was originally defined by Navarro et al. (1997) as a parameter of the NFW profile. Therefore, fitting this profile to individual haloes is the only means of accurately measuring the concentration in a way that is true to its definition\*. Even for an

---

\*We note that for the full NFW fitting one can choose to fit the mass profiles instead of the density profiles of haloes (e.g., Kwan et al., 2013). We consider both cases as full NFW fitting, because they make use of the whole range of halo radius (neglecting certain regions excluded from the fitting).

unscreened halo in  $f(R)$  gravity, where the density profile can in principle deviate from an NFW profile, the concentration should still be measured in the same way. However there are a number of simplified methods that have been adopted in the literature, including those presented by [Prada et al. \(2012\)](#) (P12) and [Springel et al. \(2008\)](#) (S08), which we also consider in this chapter. Both methods assume that the halo is well-characterised by the NFW profile without performing a direct fitting. This allows the concentration to be predicted with more limited information, which can save time.

The P12 method uses the relation between  $V_{\max}$  and the circular velocity at the halo radius  $R_{200}$ :

$$V_{200} = \left( \frac{GM_{200}}{R_{200}} \right)^{1/2}. \quad (3.3)$$

For the NFW profile the ratio  $V_{\max}/V_{200}$  is directly related to the halo concentration,  $c_{200}$ , by,

$$\frac{V_{\max}}{V_{200}} = \left( \frac{0.216c_{200}}{f(c_{200})} \right)^{1/2}, \quad (3.4)$$

where the function  $f(c)$  is given by the following:

$$f(c) = \ln(1+c) - \frac{c}{(1+c)}. \quad (3.5)$$

Using only the measurements of  $V_{\max}$  and  $R_{200}$  (or  $M_{200}$ ), Eqs. (3.3)-(3.5) can be combined and solved numerically to estimate the concentration.

One challenge of this method that we encountered was that for some haloes in our sample,  $V_{200} > V_{\max}$ . Our estimate of  $V_{\max}$  includes only particles bound to the central, dominant subhalo of the FOF group, while  $V_{200}$  has been calculated using all particles, bound or unbound, contained within  $R_{200}$ . Therefore it is possible that some large substructure found in the outer regions of the halo or some additional unbound particles flying past the halo could result in this inequality. We did not include P12 measurements of the concentration for haloes with  $V_{200} > V_{\max}$ .

Meanwhile the S08 method involves measuring the mean overdensity within  $R_{\max}$  in units of the present-day critical density,  $\rho_{\text{crit}}$ :

$$\delta_V = \frac{\bar{\rho}(R_{\max})}{\rho_{\text{crit}}} = 2 \left( \frac{V_{\max}}{H_0 R_{\max}} \right)^2. \quad (3.6)$$

The characteristic NFW overdensity,  $\delta_c$ , is then calculated using,

$$\delta_c = \frac{\rho_s}{\rho_{\text{crit}}} = 7.213\delta_V, \quad (3.7)$$

which is related to  $c_{200}$  as follows:

$$\delta_c = \frac{200}{3} \frac{c_{200}^3}{f(c_{200})}. \quad (3.8)$$

By starting with a measurement of  $V_{\text{max}}$  (or  $R_{\text{max}}$ ), the concentration can be estimated by combining Eqs. (3.6)-(3.8). Note that this method can be used for all haloes in our sample, including haloes with  $V_{200} > V_{\text{max}}$ .

The main distinction between these two simplified methods is that the S08 method can give a more stable result which is independent of deviations from an NFW profile. This can be useful for simulations of a low resolution, since the S08 method is only sensitive to density changes around  $R_{\text{max}}$  and can therefore reproduce the density profile without too much noise. On the other hand, the P12 method uses information at  $R_{200}$  as well as  $R_{\text{max}}$ , so if the density profile deviates from an NFW profile outside  $R_{\text{max}}$  then this method can pick up this effect and yield a different result. So the P12 concentration estimate is expected to be more likely to agree with a measurement from a full fit of the NFW profile.

Because the density profile is slightly changed by  $f(R)$  gravity for unscreened haloes, the differences between the methods will prove to be important for the results (see Sec. 3.3.1).

### 3.2.2.2 *Rescaled mass*

A key aim of this chapter is to be able to predict how the halo concentration is affected in the screened and unscreened regimes of  $f(R)$  gravity. The mass of the transition between these two regimes can be predicted using the  $p_2$  parameter defined in Chapter 2. Therefore it is useful to rescale the mass by the transformation  $\log_{10}(M_{500}M_{\odot}^{-1}h) \rightarrow \log_{10}(M_{500}M_{\odot}^{-1}h) - p_2 \equiv \log_{10}(M_{500}/10^{p_2})$ , such that negative values correspond to the unscreened regime and positive values correspond to the screened regime. It is essential to use a halo mass overdensity of 500 here, as discussed in Sec. 3.2.1. Note that  $p_2$  can be measured using Eq. (2.20), together with Eq. (1.15) which is used to evaluate the background scalar field.

We motivate the rescaling approach in Fig. 3.1, which shows the stacked density profiles of the  $f(R)$  haloes for four bins of  $\log_{10}(M_{500}/10^{p_2})$ , shown across four panels. The GR data for the same redshifts is also shown in each panel for a comparison. The density profile has been scaled by  $r^2$  such that it peaks at distance  $r = R_s$  from the halo centre. This allows the concentration to be easily read off from the peak of the data, given that this is equal to  $R_{200}/R_s$ .

The  $f(R)$  data in the panel at the top-left corresponds to haloes that are partially screened. This is because the values of  $\log_{10}(M_{500}/10^{p_2})$  used are just slightly above zero, and because the transition from screened to unscreened is smooth and gradual these haloes are therefore unscreened in the outer, less-dense regions but still screened in the inner, denser regions. This means that the particles falling towards the intermediate regions from the outer regions of the halo feel a stronger gravitational pull. However, the particle motions in the innermost regions of the halo are not affected by the fifth force, since this region is still screened. Therefore in  $f(R)$  gravity the density at the outer regions is lower and the density at the intermediate regions is greater than in GR, but the density at the innermost regions is unaffected. This can actually lead to a larger scale radius  $R_s$  in  $f(R)$  gravity than in GR, resulting in a lower value of the concentration. However the deviation between the  $f(R)$  and GR density profiles in this regime of  $\log_{10}(M_{500}/10^{p_2})$  is still quite small.

The top-right and bottom-left panels show regimes in which the entire halo has become unscreened. It is likely that these haloes have only recently entered the unscreened regime, since the values of  $\log_{10}(M_{500}/10^{p_2})$  used are negative but still quite close to zero. Because the entire halo is now unscreened, the particles within both the inner and outer regions of the halo feel a stronger pull of gravity and thus fall towards the halo centre with a greater acceleration. Therefore the density is greater at the inner regions and lower at the outer regions than in GR. This results in a scale radius  $R_s$  that is smaller in  $f(R)$  gravity, and so the concentration is greater than in GR. This regime of  $\log_{10}(M_{500}/10^{p_2})$  has the greatest deviation between the  $f(R)$  and GR density profiles.

Finally, the bottom-right panel shows data that is deep within the unscreened regime. This is because the values of  $\log_{10}(M_{500}/10^{p_2})$  used are negative and much lower than the

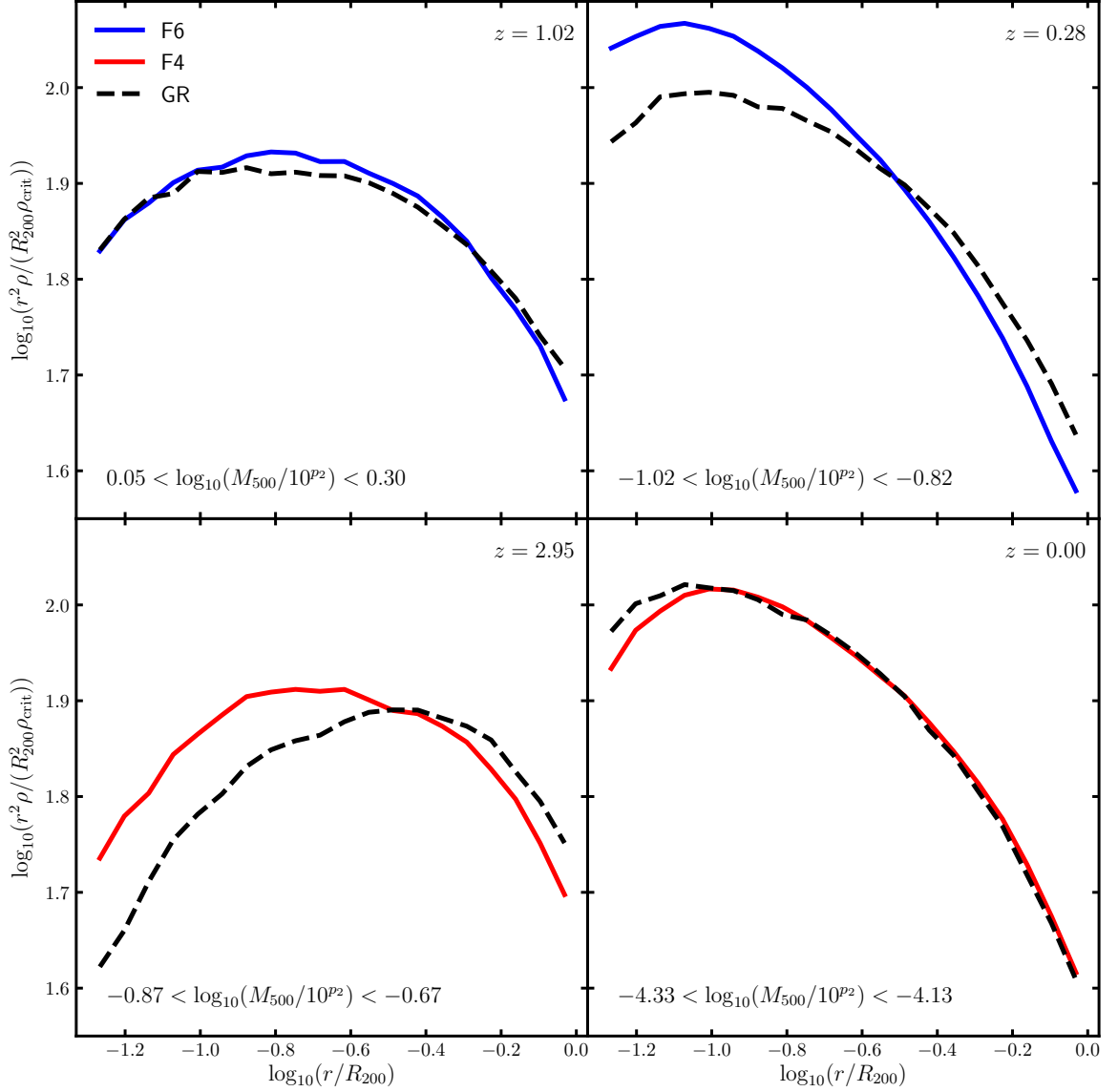


Figure 3.1: Stacked density profile, scaled by  $r^2$ , measured using the median density profile of the haloes found within a particular mass bin. The data is from the AREPO simulation (see Table 3.1), run for Hu-Sawicki  $f(R)$  gravity with  $n = 1$  (solid line) and GR (dashed line). Each panel corresponds to a particular value of the present-day scalar field  $f_{R0}$ , where F6 and F4 correspond to  $|f_{R0}| = 10^{-6}$  and  $|f_{R0}| = 10^{-4}$ , respectively, and a particular redshift  $z$ . The mass bins used cover the same range of halo masses,  $M_{500}$ , for the  $f(R)$  and GR data, and each panel corresponds to a unique range of  $\log_{10}(M_{500}/10^{p_2})$  values (annotated), where  $p_2$  is evaluated with Eq. (2.20) using the values  $|f_{R0}|$  and  $z$  of that panel.

values spanned by the top-right and bottom-left panels. Therefore the haloes in this bin are likely to have been unscreened for a significant period of time. Interestingly, the density profiles in GR and  $f(R)$  gravity are in reasonable agreement in all regions apart from the innermost region, in which the GR haloes are more dense. One possibility is that in  $f(R)$  gravity the particles that initially fall into the halo centre gain a higher velocity during this substantial period of enhanced gravitational acceleration (Shi et al., 2015), such that they are unable to settle into orbits at the innermost regions. As a result the scale radius is larger in  $f(R)$  gravity than in GR, such that the concentration is greater in GR.

From these results, it is clear that the variable  $\log_{10}(M_{500}/10^{p_2})$  can be a useful measure of the amount of screening of a halo, and so the model used to predict the enhancement of the concentration (Sec. 3.3.2) has been measured with respect to this variable. Plotting all results as a function of this variable also allows the combination of data with different values of  $f_{R0}$  and  $z$ , since these are encapsulated by  $p_2$  (see Sec. 3.3).

### 3.2.2.3 *Concentration enhancement*

The GR data of each simulation has been outputted at snapshots with the same redshifts as the  $f(R)$  data. Therefore the concentration enhancement of the haloes in a particular mass bin can be evaluated by first computing the median concentrations using the  $f(R)$  and GR data from that bin, and then taking a ratio of these quantities. The choice of binning, the measurement of the median concentration and its error, and the evaluation of the concentration enhancement and its error is discussed in this section.

In order to evaluate the concentration enhancement for the haloes of a particular snapshot, the absolute concentrations of all haloes in  $f(R)$  gravity and GR were measured using the methods discussed in Sec. 3.2.2.1. The haloes in each model were then binned by the halo mass  $M_{500}$ , with the binning chosen such that all bins are of equal width when viewed on a logarithmic axis apart from the highest-mass bin, which was allowed to be wide enough to contain at least 75 haloes in both  $f(R)$  gravity and GR. For a given snapshot, the same bins are used for both the GR and  $f(R)$  gravity datasets, and the number of bins used is the greatest possible number such that each bin contains at least 75 haloes for both



models.

The decision to use a wider highest-mass bin was taken due to the much smaller halo count at higher masses. The choice of having at least 75 haloes in each bin ensured that a balance was found between having enough bins with which to fit a reliable trend and keeping the errors low for each bin. This decision followed some tests which found that a minimum of 50 haloes per bin generates bins with a scatter that is too large, while a minimum of 100 haloes per bin results in very few bins, particularly in the AREPO and Diamond data which do not contain as many haloes as the other simulations.

In each bin, the mean logarithm of the halo mass and the median concentration was measured for both  $f(R)$  gravity and GR. The average of the two mean mass logarithms has been used to represent the mass of each bin. The error of the median concentration was evaluated using jackknife resampling (see Chapter 2), with 20 resamples.

Finally, the logarithm of the ratio of the median concentration values in  $f(R)$  gravity and GR,  $\log_{10}([c/c_{\text{GR}}]_{200})$ , was evaluated to obtain the concentration enhancement. The error of the enhancement was measured by combining the median concentration errors in quadrature to find the error of the ratio, then the error of the logarithm of this ratio. Treating the errors as independent is justified here because the particle positions in collapsed structures are uncorrelated in these two simulations even though they start from the same initial conditions. This data is shown by the symbols in Figs. 3.2 and 3.3 for an arbitrary selection of snapshots and models for each simulation. Note that the lines plotted in each panel correspond to predictions from our general model, which is discussed in Sec. 3.3.2. For each column the snapshots that are shown span the full range of available redshifts (see Table 3.1). For each bin the concentration enhancement and its error bar has been measured using the methods discussed above. The use of a wider highest-mass bin in each snapshot, as discussed above, can also be seen in each panel. Note that the same GR data is used when measuring the concentration enhancement for different  $f(R)$  gravity models of the same simulation, although the binning scheme that is used may vary slightly.

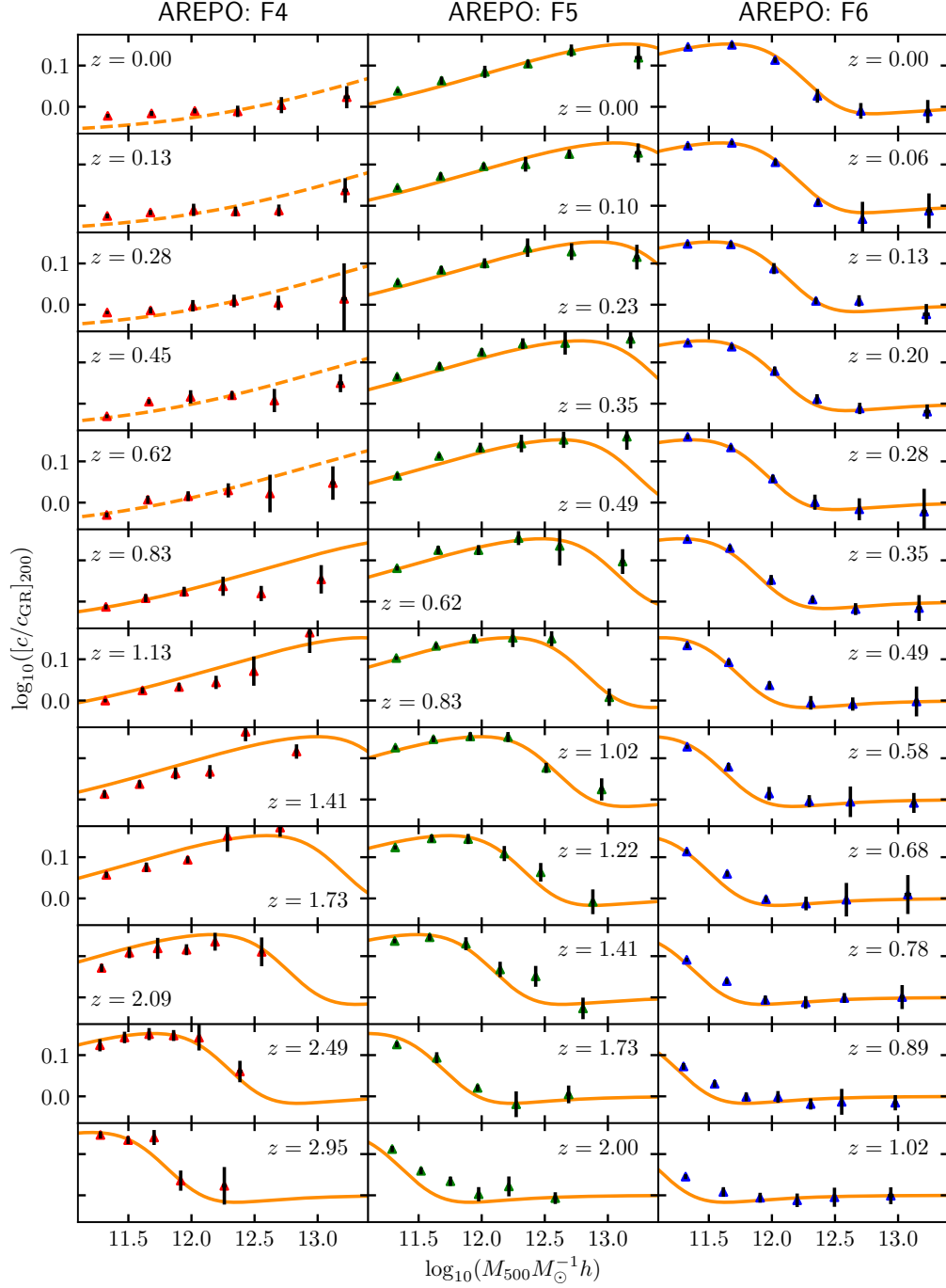


Figure 3.2: Ratio of the median concentrations of  $f(R)$  gravity and GR as a function of the halo mass, for F4 (*left column*), F5 (*middle column*) and F6 (*right column*) at various redshifts,  $z$ , as annotated. The parameter  $p_2$  is evaluated using the values  $|f_{R0}|$  and  $z$  of each panel. The data has been generated from the AREPO simulation (see Table 3.1). The one standard deviation error bars are shown. Predictions have been plotted (*solid line*) for most snapshots, and are measured using the fit of Eq. (3.9) to the data of Fig. 3.6. In panels corresponding to data excluded from the fit, the predictions are shown with dashed lines.

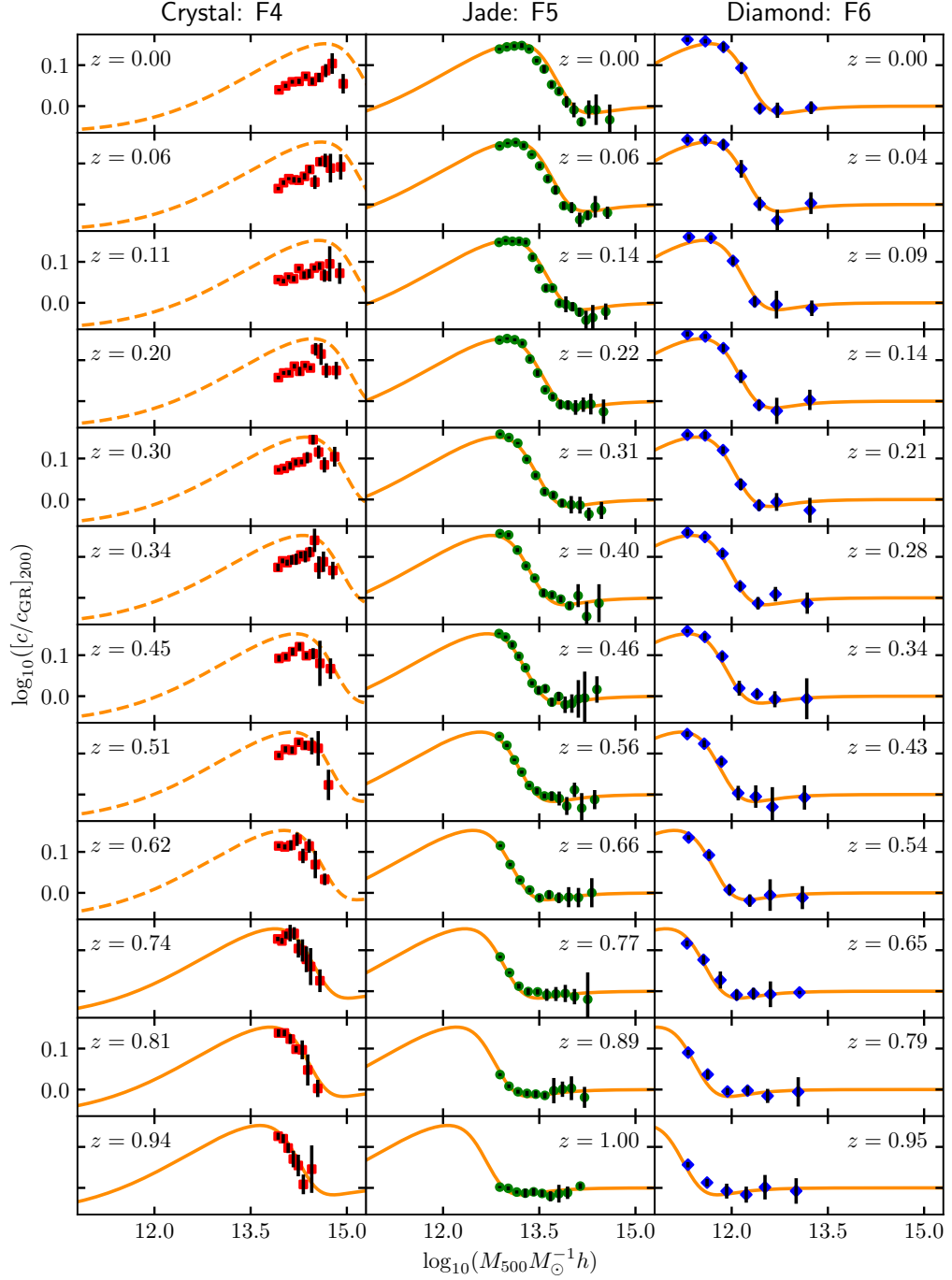


Figure 3.3: Ratio of the median concentrations of  $f(R)$  gravity and GR as a function of the halo mass, for F4 (*left column*), F5 (*middle column*) and F6 (*right column*) at various redshifts,  $z$ , as annotated. The parameter  $p_2$  is evaluated using the values  $|f_{R0}|$  and  $z$  of each panel. The data has been generated using the Crystal, Jade and Diamond ECOSMOG simulations (see Table 3.1), for F4, F5 and F6, respectively. The one standard deviation error bars are shown. Predictions have been plotted (*solid lines*) for most snapshots, and are measured using the fit of Eq. (3.9) to the data of Fig. 3.6. In panels corresponding to data excluded from the fit, the predictions are shown with dashed lines.

### 3.3 Results

The results shown in Figs. 3.2 and 3.3 are plotted against the halo mass  $M_{500}$ . However, as discussed in Sec. 3.2.2.2, performing the transformation  $\log_{10}(M_{500}M_{\odot}^{-1}h) \rightarrow \log_{10}(M_{500}/10^{p_2})$  converts the mass into a rescaled form in which negative values roughly correspond to the unscreened regime of halo mass and positive values roughly correspond to the screened regime. One advantage of this mass rescaling is that the  $f(R)$  gravity model, redshift and cosmological parameters  $\Omega_M$  and  $\Omega_\Lambda$  are all encapsulated by the  $p_2$  parameter (Eq. 2.20), so plotting against this rescaled mass can allow data from different  $f(R)$  gravity models and even from simulations run for different cosmologies to be combined and plotted together in order to extract general trends. Note that there is a unique  $p_2$  value for every combination of  $z$ ,  $f_{R0}$ ,  $\Omega_M$  and  $\Omega_\Lambda$ , so, for example, all of the data in a particular panel of Fig. 3.2 or 3.3 would be shifted by the same amount along the  $\log_{10}(M_{500}M_{\odot}^{-1}h)$  axis when applying the transformation to the mass.

We first used this plotting scheme to look into the three methods for measuring the concentration (see Sec. 3.2.2.1) to see how the choice of measurement can affect the results. This is discussed in Sec. 3.3.1. Then, focusing on the data produced from performing direct fitting of the NFW profile to individual haloes, a general model to describe the enhancement of the halo concentration in  $f(R)$  gravity was sought and is discussed in Sec. 3.3.2.

#### 3.3.1 Concentration measurement comparison

The concentration of each halo was measured using each of the three methods presented in Sec. 3.2.2.1. For each of these measures, the data was binned for every snapshot of each  $f(R)$  model and the concentration enhancement and its error was measured for each bin using the method discussed in Sec. 3.2.2.3. All of this data was plotted together against  $\log_{10}(M_{500}/10^{p_2})$  to yield Fig. 3.4, in which the three panels correspond to the three methods of measuring  $c_{200}$ .

The data from each panel follows a similar general trend. There is approximately zero enhancement of the concentration in the screened regime, then at lower mass (entering

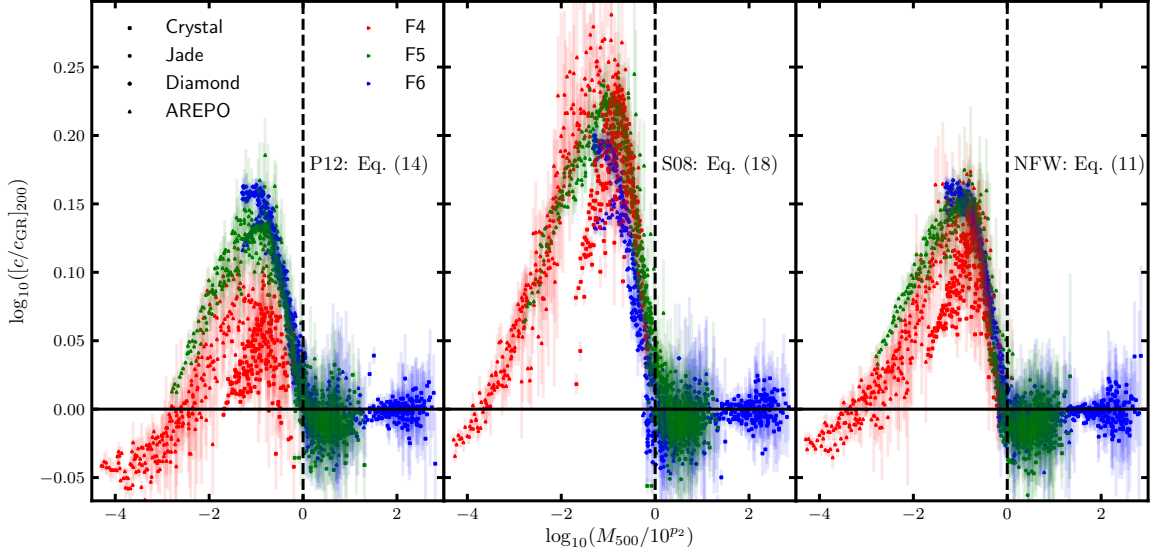


Figure 3.4: Ratio of the median concentrations of  $f(R)$  gravity and GR as a function of the rescaled mass,  $\log_{10}(M_{500}/10^{p_2})$ , for F4 (red), F5 (green) and F6 (blue). The plotted data is from the simulations summarised by Table 3.1, sifted so that only haloes with more than 1000 particles enclosed within  $R_{500}$  are included. The concentration of each halo has been calculated using the methods discussed in Sec. 3.2.2.1, namely the methods discussed by P12 (left) and S08 (middle) and by performing a direct fitting of the NFW profile to the halo density profiles (right). The one standard deviation error bars are shown.

the unscreened regime) the enhancement rises to a peak at  $\log_{10}(M_{500}/10^{p_2}) \approx -1$ , before dropping to a negative enhancement at  $\log_{10}(M_{500}/10^{p_2}) \lesssim -3$ , where the GR concentration exceeds the  $f(R)$  concentration. There is also a small dip in the concentration enhancement for  $0 < \log_{10}(M_{500}/10^{p_2}) < 1$ . The stacked profiles of Fig. 3.1 and the discussion in Sec. 3.2.2.2 can provide a physical interpretation of this behaviour.

The most accurate measurement of  $c_{200}$  is by performing a direct fitting of the NFW profile to the halo profiles, so the panel on the right in Fig. 3.4 is expected to give the most reliable result. Here, the three  $f(R)$  gravity models all show a similar behaviour and even peak at the same enhancement, which is approximately 0.15. Only the F4 data from the Crystal simulation shows any deviation from this behaviour, as it appears to have a lower concentration enhancement than the rest of the data at  $\log_{10}(M_{500}/10^{p_2}) \approx -1$ . However the good agreement of most of the data means that a general model can be fitted using a portion of the data, as discussed in Sec. 3.3.2.

The P12 data reaches the same maximum enhancement of approximately 0.15. How-

ever, the disparity between the three  $f(R)$  models is now greater, with the models each peaking at a different enhancement. F6 has the highest peak enhancement and F4 has the lowest peak. Note that the P12 data has a reduced sample which excludes all haloes with  $V_{200} < V_{\max}$  (see Sec. 3.2.2.1). The S08 data reaches a greater maximum enhancement of approximately 0.25, and shows an opposite trend in terms of the order of the models: F4 now has the highest peak enhancement and F6 has the lowest peak.

The difference in the results from the P12 and S08 methods is not surprising, yet it is significant. As shown in Fig. 3.1, data at  $\log_{10}(M_{500}/10^{p_2}) \approx -1$  has a greater density at the inner regions of the halo and a lower density at the outer regions compared with GR, due to the enhanced acceleration of the in-falling particles. The S08 method only uses data from  $R_{\max}$  to measure  $c_{200}$ , whereas the P12 method uses data from  $R_{\max}$  and  $R_{200}$ . Being at a smaller distance from the halo centre, the mass enclosed by  $R_{\max}$  is more affected by the in-fall of particles than the mass enclosed by  $R_{200}$ . So  $V_{\max}$  is more affected than  $V_{200}$ , with the result that the S08 method measures a higher  $f(R)$  concentration than the other methods. The P12 method is closer to actually fitting a density profile to the full extent of the halo, and so it yields a closer result to the full NFW fitting.

These results indicate that the choice of method for measuring the concentration can be very important in MG models. The three methods discussed in this chapter only agree perfectly for ideal NFW profiles, and therefore only the direct NFW fitting should be used for realistic cases. Note that this statement assumes that the halo density profiles in  $f(R)$  gravity can still be well described by the NFW profile, which needs to be checked explicitly (see below). However, even if NFW is no longer valid, approximate methods such as P12 and S08 should not be used instead of full NFW fitting either as they are derivatives of the latter. Indeed, using the S08 method in  $f(R)$  gravity could lead to a measurement of the halo concentration that is up to 26% greater than from performing a fit. We therefore elected to use the direct NFW fitting method for all other results in this chapter. Note that in Sec. 3.3.2 we include a check of the validity of the NFW profile in  $f(R)$  gravity.

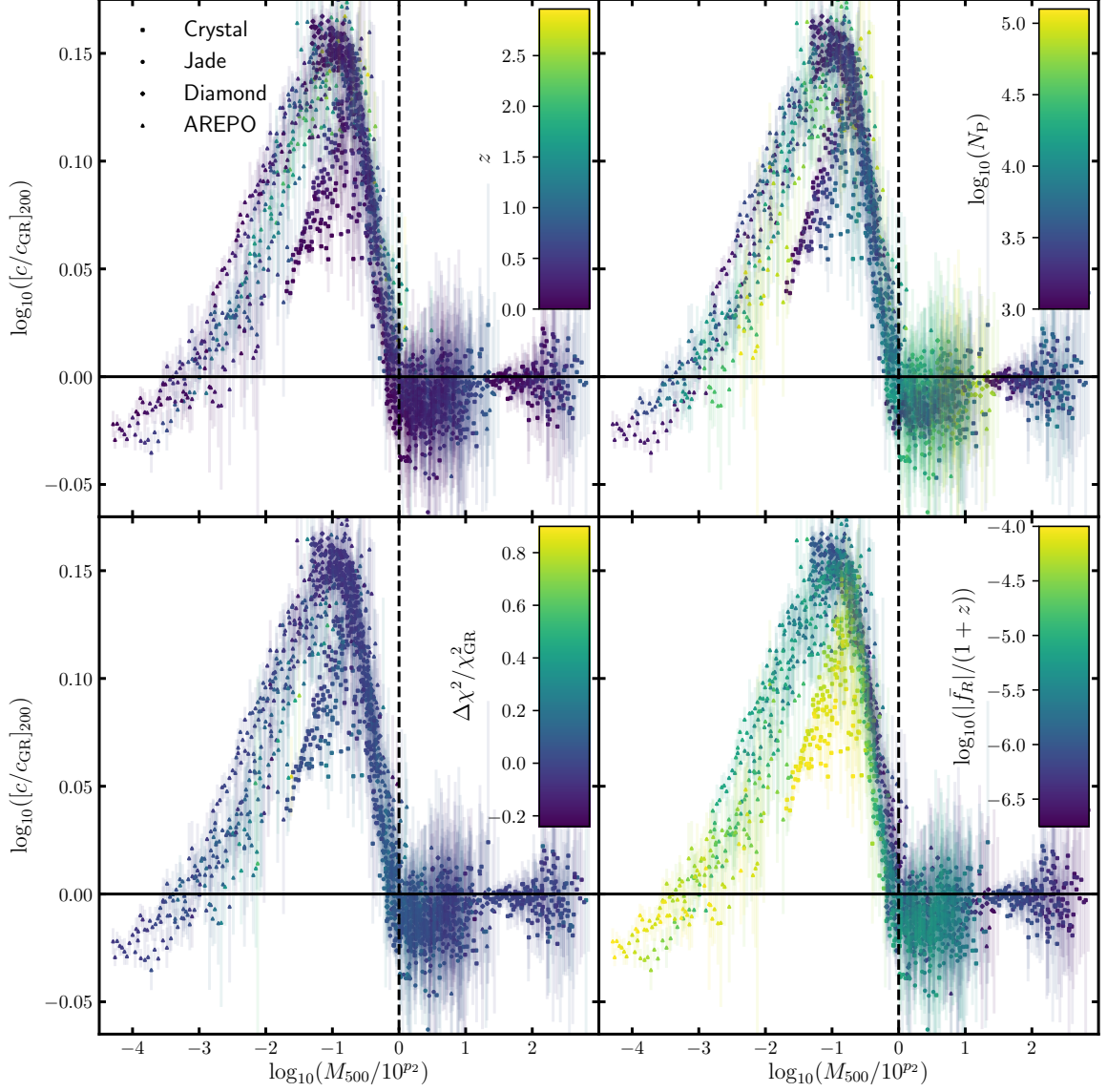


Figure 3.5: Ratio of the median concentrations of  $f(R)$  gravity and GR as a function of the rescaled mass,  $\log_{10}(M_{500}/10^{12})$ , for F4, F5 and F6. The plotted data is from the simulations summarised by Table 3.1, sifted so that only haloes with more than 1000 particles enclosed within  $R_{500}$  are included. The one standard deviation error bars are shown. The data is coloured by (clockwise from top left) the redshift, the mean number of particles within  $R_{500}$ , the logarithm of a combination of the scalar field and redshift,  $|f_R|/(1+z)$ , and the fractional difference of chi squared measures generated by NFW fits of the halo profiles.

### 3.3.2 General model for the concentration enhancement

Before applying a fitting formula to the data it was useful to check whether certain factors could be affecting the shape of the trend, therefore the data was coloured via various schemes which are shown in Fig. 3.5.

The top-left and top-right panels show the colourings by redshift and the particle number within each halo, respectively. These can both be viewed as tests of the effect of the halo resolution. For example, haloes with fewer particles can be prone to resolution effects at the innermost and outer regions, where the density can be underestimated. Haloes are less dense and more diffuse at higher redshift, which therefore leads to a greater exposure to these effects. As discussed in Sec. 3.2.2.1, in an effort to prevent these effects we restricted the fitting of the NFW profile to the radial range  $0.05R_{200}$  to  $R_{200}$ . Furthermore we only include haloes that contain at least 1000 particles within  $R_{500}$  in our sample. The coloured data of Fig. 3.5 suggests that these measures were sufficient, as it can be seen that even data at  $z \approx 3$  agrees with the main trend and every part of the trend consists of haloes with both low and high particle numbers. Therefore even for the F4 Crystal data, low resolution is unlikely to be the reason for any disparity with the main trend.

It is significant that we are able to use redshifts up to  $z = 3$  for the AREPO data. For a given  $f_{R0}$  value, haloes at high  $z$  are more screened than haloes at low  $z$ . This is because the magnitude of the background scalar field  $\bar{f}_R$  grows as a function of time, such that haloes of a given mass will eventually go from being screened to unscreened. For  $f(R)$  gravity models with a stronger scalar field (greater  $|f_{R0}|$ ), haloes at a given redshift are more unscreened and therefore have a lower  $\log_{10}(M_{500}/10^{p2})$  value. However, models that are stronger than F4 are infeasible given current constraints on  $f(R)$  gravity, and the minimum redshift that is available is  $z = 0$ . Therefore the minimum value of  $\log_{10}(M_{500}/10^{p2})$  that can be plotted is only limited by the simulation resolution, as only haloes with lower mass can exist at lower values of this rescaled mass; similarly, the maximum value of  $\log_{10}(M_{500}/10^{p2})$  is limited by the box size. For each of the  $f(R)$  gravity models tested in this analysis, the range of redshifts used provides a range of  $\log_{10}(M_{500}/10^{p2})$  that extends from the lowest value that is possible at halo mass  $M_{500} = 1.52 \times 10^{11} h^{-1} M_{\odot}$  to values in the screened regime,



at which there is approximately no enhancement of the concentration compared with GR and so the concentration is much easier to predict. A weaker model of  $f(R)$  gravity would likely exist close to or within the screened regime for  $M_{500} \geq 1.52 \times 10^{11} h^{-1} M_{\odot}$  at  $z = 0$ . Therefore, given that all three models tested in this chapter show excellent agreement for  $-0.5 \leq \log_{10}(M_{500}/10^{p_2}) \leq 0.0$ , it seems that a fit of this trend should be general for  $M_{500} \geq 1.52 \times 10^{11} h^{-1} M_{\odot}$  for arbitrary values of  $f_{R0}$  that are allowed by current constraints.

For every NFW fit, we stored the  $\chi^2$  value, which is measured by summing the squared residuals of the 20 radial bins. Storing the median  $\chi^2$  value for every mass bin, the GR and  $f(R)$  values of the latter were then combined to generate the fractional  $\chi^2$  difference. The bottom-left panel of Fig. 3.5 shows the data coloured by this measure, and can therefore be seen as a test of the validity of the NFW profile in  $f(R)$  gravity. The colouring shows that the goodness-of-fit of the NFW profile for most haloes in  $f(R)$  gravity is within 20% of the goodness-of-fit in GR. The colour-bar here shows the full range of fractional differences that were observed in the data, and we note that only a very small minority of haloes have a  $\chi^2$  that is almost 90% higher than in GR. These results are promising, and imply that systematics induced through the fitting of the NFW profile are unlikely to impact on the scatter of the halo concentration in  $f(R)$  gravity.

Finally, the data was coloured by the logarithm of  $|\bar{f}_R|/(1+z)$ , and this is shown in the bottom-right panel of Fig. 3.5. In Chapter 2 we found that complicated  $f(R)$  gravity effects, including screening, can effectively be described by this useful parameter. It is therefore useful to see how the enhancement of the halo concentration at screened and unscreened regimes can depend on this. An interesting observation is that bins with  $|\bar{f}_R|/(1+z) \lesssim 10^{-4.5}$  are in excellent agreement with a smooth trend for  $-4 \lesssim \log_{10}(M_{500}/10^{p_2}) \lesssim 3$ . The F6 data, which reaches a peak enhancement at  $z \approx 0$ , shows very good agreement with the F5 data, and both models agree well with the higher- $z$  F4 data. Therefore if a cut is made so that only data with  $|\bar{f}_R|/(1+z) \leq 10^{-4.5}$  is used in the fitting, then, at least for halo masses  $M_{500} \geq 1.52 \times 10^{11} h^{-1} M_{\odot}$ , a general model can be created that applies to arbitrary  $f_{R0}$  values provided  $|\bar{f}_R|/(1+z) \leq 10^{-4.5}$ . For models with  $|f_{R0}| > 10^{-4.5}$  the concentration enhancement does not follow the same trend, and therefore cannot be described by the

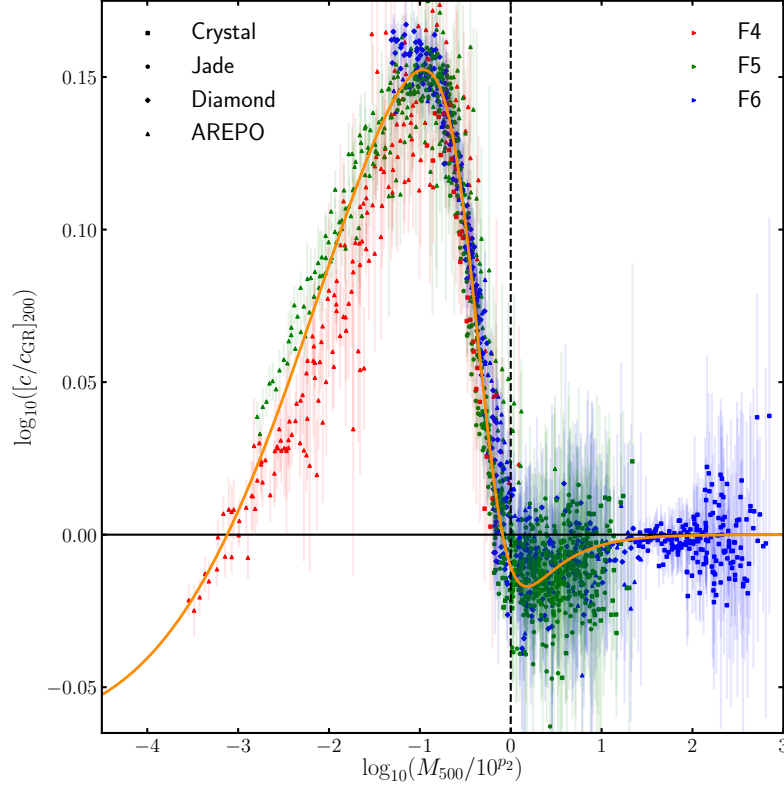


Figure 3.6: Ratio of the median concentrations of  $f(R)$  gravity and GR as a function of the rescaled mass,  $\log_{10}(M_{500}/10^{p_2})$ , for F4 (red), F5 (green) and F6 (blue). Only data with  $|\bar{f}_R|/(1+z) \leq 10^{-4.5}$  has been included and fitted with Eq. (3.9). This fit is shown by the trend-line, and the optimised parameter values are listed in Table 3.2. The plotted data is from the simulations summarised by Table 3.1, sifted so that all haloes have at least 1000 particles enclosed within  $R_{500}$ . The one standard deviation error bars are shown.

universal fitting formula below. However, we note that models with  $|f_{R0}| > 10^{-4.5}$  have already been strongly disfavoured by observations.

Cleaning the data to remove bins with  $|\bar{f}_R|/(1+z) > 10^{-4.5}$  and re-plotting yields Fig. 3.6. This shows a clear trend. As  $\log_{10}(M_{500}/10^{p_2})$  is reduced from value 3, the concentration enhancement, which is initially zero, drops slightly to a negative enhancement. Continuing into the unscreened regime the enhancement then rises to a distinct peak with value  $\approx 0.15$  at  $\log_{10}(M_{500}/10^{p_2}) \approx -1$ , before dropping down to negative values again for  $\log_{10}(M_{500}/10^{p_2}) \lesssim -3$ . From the discussion in Sec. 3.2.2.2 and the results in Fig. 3.1, the above behaviour is physical and should therefore be fully included in the fitted model.

In selecting a suitable fitting formula, both the screened and unscreened regimes of the data were considered. The data in the unscreened part of Fig. 3.6 shows good agreement

with a skewed Gaussian curve, whose steepness is different on the two sides of the peak. This requires five parameters: a scaling  $\lambda$  of the height of the curve, a shift  $\gamma$  along the  $\log_{10}([c/c_{\text{GR}}]_{200})$  axis, a width  $\omega_s$ , a skewness parameter  $\alpha$  and a parameter  $\xi_s$  to describe the location with respect to the  $\log_{10}(M_{500}/10^{p_2})$  axis. As discussed in Sec. 3.2.2.2 and from examining the top-left panel of Fig. 3.1, there is a physical motivation that the concentration should dip slightly at halo masses just greater than  $10^{p_2} h^{-1} M_{\odot}$ . Therefore the model would have to include a minimum in this regime. This was achieved by multiplying the skewed Gaussian with a tanh curve, which takes value 1 at low values of  $\log_{10}(M_{500}/10^{p_2})$  and drops to value 0 at high values. This induces a further two parameters: a location  $\xi_t$  and a width  $\omega_t$  with respect to the  $\log_{10}(M_{500}/10^{p_2})$  axis. The tanh curve also ensures that the model tends to zero at higher  $\log_{10}(M_{500}/10^{p_2})$ . This model would gradually level out at  $\log_{10}(M_{500}/10^{p_2}) < -4$ , however this is not necessarily how the concentration enhancement would behave in this regime. The only way to understand this would be to run higher-resolution simulations so that lower-mass haloes can be investigated.

By taking the above considerations into account, we arrive at a 7-parameter fitting formula, which is given by,

$$y(x) = \frac{1}{2} \left( \frac{\lambda}{\omega_s} \phi(x') \left[ 1 + \operatorname{erf} \left( \frac{\alpha x'}{\sqrt{2}} \right) \right] + \gamma \right) (1 - \tanh(\omega_t [x + \xi_t])), \quad (3.9)$$

where  $y = \log_{10}([c/c_{\text{GR}}]_{200})$ ,  $x' = (x - \xi_s)/\omega_s$  and  $x = \log_{10}(M_{500}/10^{p_2})$ . The left-hand bracket of Eq. (3.9) represents the skewed Gaussian curve, where  $\phi(x)$  is the normal distribution:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right). \quad (3.10)$$

This also includes a multiplication with the error function  $\operatorname{erf}(x')$  in order to generate a skewed curve. The error function is given by,

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (3.11)$$

The fit of Eq. (3.9) to the data of Fig. 3.6 was carried out by minimising the sum of the squared normalised residuals of the data points, where the residuals have been normalised by the one standard deviation error bars shown. The sum is evaluated in a way that treats different parts of the  $\log_{10}(M_{500}/10^{p_2})$  range equally. This has been achieved by splitting

$\lambda$	$\xi_s$	$\omega_s$	$\alpha$	$\gamma$	$\omega_t$	$\xi_t$
$0.55 \pm 0.18$	$-0.27 \pm 0.09$	$1.7 \pm 0.4$	$-6.5 \pm 2.4$	$-0.07 \pm 0.04$	$1.3 \pm 1.0$	$0.1 \pm 0.3$

Table 3.2: Optimal parameter values and errors from the fit of Eq. (3.9) to the data of Fig. 3.6. The fit is carried out by first splitting the range of  $\log_{10}(M_{500}/10^{p_2})$  into 13 equal-width bins. The squared normalised residuals of the data points are then weighted by the reciprocal of the number of data points in the current bin. The sum of these is minimised by varying the parameters.

this range into 13 equal-width bins. The squared normalised residuals are then weighted by the reciprocal of the number of data points in the current bin, prior to minimising the sum through varying the parameters. The optimal parameters are listed in Table 3.2 and the corresponding fit is shown in Fig. 3.6.

We have also considered a weighted least squares fit which neglects the weighting of the squared normalised residuals described above. This results in a model that produces nearly identical predictions to the model shown in Fig. 3.6. However, neglecting the weighting of the squared normalised residuals disfavors parts of the  $\log_{10}(M_{500}/10^{p_2})$  range that contain fewer data points, including the AREPO F4 data at  $\log_{10}(M_{500}/10^{p_2}) \lesssim -3$ . Therefore, we only include results from the fitting described above.

In order to test our model, its predictions were generated for the data shown in Figs. 3.2 and 3.3. The predictions are shown by the plotted lines. Solid lines are used in snapshots which were used to generate the fit in Fig. 3.6 and dashed lines are used in snapshots excluded from the fit (snapshots with  $|\bar{f}_R|/(1+z) > 10^{-4.5}$ ). Agreement is generally excellent between the data and the predictions in both figures. Agreement is reasonable even for the low- $z$  F4 snapshots of AREPO that were not used to generate the model, as can be seen in Fig. 3.2. Some small disparity exists in the higher- $z$  F5 and F6 snapshots, where the data does not appear to agree with the predicted minimum in the screened regime. Again, there are probably some physical effects that result in subtly different trends at different redshifts. However, given the complexity of the behaviour of the halo concentration in chameleon  $f(R)$  gravity and the simplicity of our modelling, the amount of agreement shown in these figures is indeed very good.

### 3.4 Summary, Discussion and Conclusions

In this chapter, a model has been developed for the enhancement of the halo concentration in HS  $f(R)$  gravity with  $n = 1$  using a suite of simulations that are summarised by Table 3.1. The model is shown in Fig. 3.6, and is given by Eq. (3.9) with the parameter values listed in Table 3.2. It has been defined in terms of a useful rescaling of the halo mass,  $M_{500}/10^{p_2}$ , with  $p_2$  calculated using Eq. (2.20), such that data from three different  $f(R)$  gravity models can satisfy a universal description. These models have  $\log_{10}(|f_{R0}|) = (-4, -5, -6)$ , and the fitting was carried out using data from all simulation snapshots with  $\log_{10}(|\bar{f}_R|/(1+z)) \leq -4.5$ . This universal description is shown to have very good agreement with simulations for  $M_{500}/10^{p_2}$  covering nearly 7 orders of magnitude, and covering five decades of the halo mass.

Our model has been tested by comparing its predictions of the enhancement of the concentration with an arbitrarily chosen set of snapshots from our simulations, as shown by the lines plotted in Figs. 3.2 and 3.3. These predictions show excellent agreement with the data for all snapshots, apart from the Crystal snapshots with  $\log_{10}(|\bar{f}_R|/(1+z)) > -4.5$ . This is not surprising given that this data was not used in the fit of the model. Having a general model that works for  $\log_{10}(|\bar{f}_R|/(1+z)) \leq -4.5$  will prove very useful, particularly given that an analytical theoretical modelling was not available.

The data of Fig. 3.6 shows that in the unscreened regime the enhancement of the concentration reaches a distinct peak as a function of the halo mass, but drops to negative values at lower mass, where the  $f(R)$  concentration is less than the GR concentration. As shown by Fig. 3.1, such negative enhancement occurs because the innermost regions of the haloes are less dense in  $f(R)$  gravity than in GR. This could be caused by the velocity gained by particles in haloes, which makes it difficult for them to settle into orbits at the central regions of the halo. Meanwhile in the screened regime of Fig. 3.6 there is a small dip in the concentration. Fig. 3.1 suggests that this is caused by the halo being only partially screened, so that outer particles are moved further towards the centre of the halo while the inner regions remain screened. The density profile is therefore unaffected at the innermost

regions but is greater at intermediate radii. Therefore the scale radius becomes greater, and fitting an NFW profile would then result in an estimate for the concentration that is lower in  $f(R)$  gravity than in GR. All of these effects are incorporated by the fitted model of Eq. (3.9).

Some further investigations were carried out which can be useful for further studies of the concentration in  $f(R)$  gravity, and in other similar MG theories. Firstly, in addition to applying a direct NFW profile fitting to each of the haloes to measure the concentration, two simplified approaches were also used, namely the methods that are used by Prada et al. (2012) and Springel et al. (2008). The resulting enhancement of the concentration from using these two methods (shown in Fig. 3.4) shows a difference from direct NFW fitting. This is due to the effects of  $f(R)$  gravity on the internal density profile, which means that the choice of regions of the halo to use in measuring the concentration becomes important. The method used by Springel et al. (2008) only requires the mass enclosed by the orbital radius with the maximum circular velocity. Being found at the inner regions of a halo, which become more dense as the halo becomes unscreened, this results in the concentration being overestimated by up to 26%. From this, we conclude that only the direct NFW fitting should be used in  $f(R)$  studies. Secondly, we looked at the validity of the NFW profile fitting in  $f(R)$  gravity and found that, as shown by the bottom-left panel of Fig. 3.5, for most haloes the  $\chi^2$  measure for the fit is within 20% of the GR measure, and for some haloes the fit is even better. Therefore the systematic effects caused by fitting the NFW profile in  $f(R)$  gravity are unlikely to have a significant effect on the scatter of the concentration measure.

The results of this chapter show that the  $p_2$  parameter defined in Chapter 2 can indeed be very useful in the description and modelling of complicated effects in  $f(R)$  gravity. In addition to its relatively simple one-parameter definition, it may also allow the combining of data generated by simulations run for different cosmological parameters, as  $p_2$  encapsulates the values of  $\Omega_M$  and  $\Omega_\Lambda$ . Indeed, the data for the concentration enhancement from AREPO and Diamond F6 shows excellent agreement (see Fig. 3.6), even though these two simulations were run for different cosmological parameters and using very different codes. It will be

interesting to see where else  $p_2$  can be used in  $f(R)$  studies. Of particular interest would be to see how it can simplify the modelling of the HMF. The enhancement of the HMF in  $f(R)$  gravity peaks at a particular halo mass which depends on the strength of the scalar field. A stronger scalar field allows higher-mass haloes to be unscreened, and therefore results in an enhancement of the HMF at a higher mass. At the very least, the mass of the peak enhancement of the HMF can be expected to be strongly correlated to  $p_2$ . The enhancement of the matter power spectrum could also be investigated via a similar treatment.

For the results of this chapter, we used data from four different simulations, allowing a wide range of resolutions to be used. However, one potential drawback is that these are run for dark matter only. It will therefore be important to test these results using cluster data from full-physics hydrodynamical simulations run for  $f(R)$  gravity.

# Observable-mass scaling relations in $f(R)$ gravity

## 4.1 Introduction

In this chapter, we analyse the effects of  $f(R)$  gravity on cluster observable-mass scaling relations (green dotted box of Fig. 2.1). These are modified by the effects of the fifth force on the gravitational potentials of haloes, which are intrinsically linked to the dynamical mass and gas temperature. According to [He and Li \(2016\)](#), who employed a suite of non-radiative simulations to investigate the effect of  $f(R)$  gravity on a number of mass proxies, it is possible to map between the scaling relations in  $f(R)$  gravity and GR using only the relation between the dynamical and true (or lensing) masses of a halo, which is accurately captured by our tanh fitting formula (Eq. (2.17)). Here, we test these predictions using non-radiative hydrodynamic simulations with much higher resolutions, and build upon this by checking how the addition of full-physics effects such as cooling, star formation and feedback impact the accuracy. To this end, we make use of the first simulations that simultaneously incorporate both full-physics and  $f(R)$  gravity ([Arnold et al., 2019b](#)). In addition, we propose and test a set of alternative mappings from the scaling relations in GR to their  $f(R)$  counterparts, which again require only our tanh fitting formula. We note that, for this chapter, our simulations only cover galaxy groups and low-mass clusters ( $M \lesssim 10^{14.5} M_\odot$ ).



However, in Chapter 8 we will use larger simulations (run using a retuned model for the baryonic physics) to extend these tests to  $\sim 10^{15}M_{\odot}$ .

This chapter is organised as follows: in Sec. 4.2, we summarise the scaling relation mappings proposed by He and Li (2016) and our novel alternative mappings; in Sec. 4.3, we describe our simulations and calculations of the halo masses and observable proxies; then, in Sec. 4.4, we present our results for the scaling relations of four mass proxies; finally, in Sec. 4.5, we summarise the results of this chapter and their significance for our framework.

## 4.2 Background

We will study the effects of HS  $f(R)$  gravity on the cluster scaling relations for four mass proxies: the gas temperature  $T_{\text{gas}}$ , the X-ray luminosity  $L_X$ , the integrated SZ flux, given by the Compton  $Y$ -parameter  $Y_{\text{SZ}}$ , and its X-ray analogue  $Y_X$ . These observables have a one-to-one mapping with the mass because of the link between the gravitational potential of a cluster and its temperature. During cluster formation, baryonic matter is accreted onto the dark matter halo from its surroundings. The gravitational potential energy of the gas is converted into kinetic energy as it falls in. During accretion, the in-falling gas undergoes shock heating, resulting in the conversion of its kinetic energy into thermal energy. The resulting self-similar model for cluster mass scaling relations predicts that the gravitational potential alone can determine the thermodynamical properties of a cluster (Kaiser, 1986).

The X-ray luminosity within radius  $r$  from the cluster centre is given by,

$$L_X(< r) = \int_0^r dr' 4\pi r'^2 \rho_{\text{gas}}^2(r') T_{\text{gas}}^{1/2}(r'), \quad (4.1)$$

where  $\rho_{\text{gas}}(r)$  is the gas density profile. The  $Y_{\text{SZ}}$  parameter is related to the integrated electron pressure of the cluster gas, and is given by,

$$Y_{\text{SZ}}(< r) = \frac{\sigma_{\text{T}}}{m_{\text{e}} c^2} \int_0^r dr' 4\pi r'^2 n_{\text{e}}(r') T_{\text{gas}}(r'), \quad (4.2)$$

where  $n_{\text{e}}$  is the electron number density. Meanwhile, the  $Y_X$  parameter (Kravtsov et al., 2006) is equivalent to the product of the gas mass and the mass-weighted gas temperature,

$\bar{T}_{\text{gas}}(< r)$ :

$$Y_X(< r) = \bar{T}_{\text{gas}}(< r) \int_0^r dr' 4\pi r'^2 \rho_{\text{gas}}(r'). \quad (4.3)$$

It has been shown in previous studies (e.g., [Fabjan et al., 2011](#)) that  $Y_X$  and  $Y_{\text{SZ}}$  have comparatively low scatter as mass proxies and are relatively insensitive to dynamical processes including cluster mergers. It has also been found that their scaling relations with the mass show good agreement with the self-similar model predictions even after the inclusion of full-physics effects such as feedbacks, which can heat up and blow gas out from the central regions (e.g., [Fabjan et al., 2011](#); [Truong et al., 2018](#); [Cui et al., 2018](#)).

Because these cluster mass proxies are closely related to the gravitational potential (and hence the dynamical mass) of clusters, and because in  $f(R)$  gravity the dynamical mass can be cleanly modelled (see Chapter 2), it is natural to expect that cluster observable-mass scaling relations in  $f(R)$  gravity can be modelled given their counterparts in GR. To study the effects of  $f(R)$  gravity on the scaling relations for these proxies, we adopt two methods which are described in the sections that follow.

#### 4.2.1 Effective density approach

In Chapter 2, we introduced the  $f(R)$  gravity effective density field  $\rho_{\text{eff}}$  ([He et al., 2014](#)), which can be expressed in terms of the true density field  $\rho$  using Eq. (2.10). The true density field corresponds to the intrinsic mass of simulation particles. As we discussed, the mass of haloes computed using the effective density field is equivalent to the dynamical mass.

Using non-radiative simulations run for the F5 model and GR, [He and Li \(2016\)](#) generated halo catalogues using the effective density field. The radius,  $R_{500}^{\text{eff}}$ , of these haloes enclosed an average *effective* density of 500 times the (*true*) critical density of the Universe. Both the total true and dynamical mass were computed within this radius, and the cluster observables were computed using all enclosed gas cells.

Analysing these data, [He and Li \(2016\)](#) found that haloes in GR and  $f(R)$  gravity with the same dynamical mass,  $M^{\text{GR}} = M_{\text{dyn}}^{f(R)} = M_{\text{dyn}}$ , also have the same gas temperature:

$$T_{\text{gas}}^{f(R)}(M_{\text{dyn}}^{f(R)}) = T_{\text{gas}}^{\text{GR}}(M^{\text{GR}} = M_{\text{dyn}}^{f(R)}). \quad (4.4)$$

The physical origin of this result is the intrinsic relationship between the gravitational potential and the gas temperature (see above). Two haloes with the same dynamical mass  $M_{\text{dyn}}$  (which we recall has also been computed within the same radius  $R_{500}^{\text{eff}}$ ), would also have the same gravitational potential,  $\phi = (GM_{\text{dyn}})/R_{500}^{\text{eff}}$ , and are therefore expected to have the same temperature. The authors also found that, outside the core region, the gas density profiles of haloes in GR are enhanced by a factor  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  with respect to haloes in  $f(R)$  gravity which have the same dynamical mass. This is because the gas density profile follows the true density profile more closely than the effective density profile, which itself is a result of the fact that clusters form from very large regions in the Lagrangian space, so that the ratio between the baryonic and total masses within clusters resembles the cosmic mean,  $\Omega_b/\Omega_M$  (White et al., 1993). The extra forces in MG theories and feedbacks from galaxy formation can add further complications to this through their effects on the gas density profiles, especially in the inner regions; however, as we will show in the following paragraph, the good agreement between the GR and rescaled  $f(R)$  gas density profiles still holds in the outer halo regions.

We have replicated the procedure adopted by He and Li (2016) using our full-physics and non-radiative simulations (for full details, see Sec. 4.3). In Fig. 4.1, the stacked temperature and gas density profiles of haloes from mass bins  $10^{13.7}M_{\odot} < M_{\text{dyn}}(< R_{500}^{\text{eff}}) < 10^{14.0}M_{\odot}$  and  $10^{13.0}M_{\odot} < M_{\text{dyn}}(< R_{500}^{\text{eff}}) < 10^{13.3}M_{\odot}$  are shown in the first and second columns from the left, respectively. The radial range is shown up to the mean logarithm of  $R_{500}^{\text{eff}}$  (which is almost exactly the same for GR, F6 and F5). For the non-radiative temperature profiles, shown in the third row, it is clear that the F6 and F5 predictions agree very well with GR in the outer regions. There is also encouraging agreement for the full-physics data, although there is a small disparity between F5 and GR in the outer regions for the higher-mass bin. For the  $f(R)$  gravity gas density profiles, the results both with and without the  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  rescaling are shown. As was found by He and Li (2016), the rescaled  $f(R)$  gravity profiles (shown by the dashed curves) agree very well with GR in the outer regions. Again, there is also promising agreement for the full-physics profiles.

These results suggest that for haloes in  $f(R)$  gravity and GR which have the same

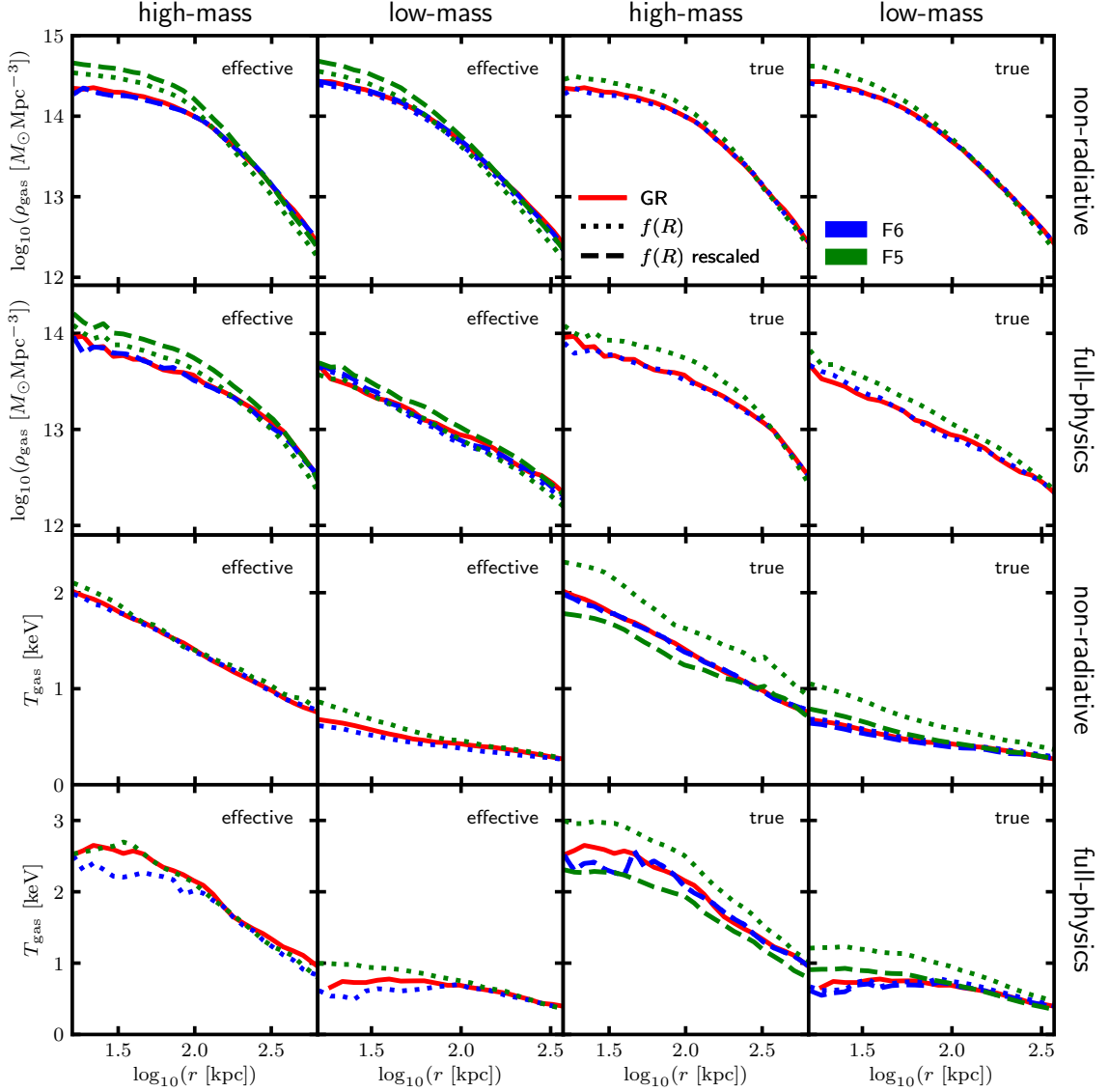


Figure 4.1: Median gas density profiles (*top two rows*) and median temperature profiles (*bottom two rows*) of FOF groups from two mass bins:  $13.7 < \log_{10}(M/M_{\odot}) < 14.0$  (*high-mass*) and  $13.0 < \log_{10}(M/M_{\odot}) < 13.3$  (*low-mass*). The group data from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1) has been used. In addition to GR (*red solid lines*), the profiles for F6 (*blue lines*) and F5 (*green lines*) are shown. Rescaled  $f(R)$  gravity profiles (*dashed lines*) are shown, along with the unaltered profiles (*dotted curves*). The rescalings correspond to the effective density (*left two columns*) and true density (*right two columns*) approaches discussed in Sec. 4.2. For the effective approach, the maximum radius shown is  $R_{500}^{\text{eff}}$  (see Sec. 4.2.1) and the halo mass  $M$  is the total dynamical mass within this radius. For the true approach, the maximum radius shown is  $R_{500}^{\text{true}}$  (see Sec. 4.2.2) and the halo mass  $M$  is the total true mass within this radius.

dynamical mass,  $M^{\text{GR}} = M_{\text{dyn}}^{f(R)}$ , the following relation is expected to apply:

$$\begin{aligned} & \int_0^r dr' 4\pi r'^2 \left( \rho_{\text{gas}}^{f(R)} \right)^a \left( T_{\text{gas}}^{f(R)} \right)^b \\ & \approx \left( \frac{M_{\text{true}}^{f(R)}}{M_{\text{dyn}}^{f(R)}} \right)^a \int_0^r dr' 4\pi r'^2 \left( \rho_{\text{gas}}^{\text{GR}} \right)^a \left( T_{\text{gas}}^{\text{GR}} \right)^b, \end{aligned} \quad (4.5)$$

where  $a$  and  $b$  represent indices of power, and we note that  $\rho_{\text{gas}}$  represents the intrinsic (not effective) gas density. Using Eqs. (4.1)-(4.3) for the mass proxies, this relation can be applied to derive the following mappings between the respective mass scaling relations in GR and  $f(R)$  gravity:

$$\frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} Y_{\text{SZ}}^{f(R)} \left( M_{\text{dyn}}^{f(R)} \right) \approx Y_{\text{SZ}}^{\text{GR}} \left( M^{\text{GR}} = M_{\text{dyn}}^{f(R)} \right), \quad (4.6)$$

$$\frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} Y_{\text{X}}^{f(R)} \left( M_{\text{dyn}}^{f(R)} \right) \approx Y_{\text{X}}^{\text{GR}} \left( M^{\text{GR}} = M_{\text{dyn}}^{f(R)} \right), \quad (4.7)$$

$$\left( \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} \right)^2 L_{\text{X}}^{f(R)} \left( M_{\text{dyn}}^{f(R)} \right) \approx L_{\text{X}}^{\text{GR}} \left( M^{\text{GR}} = M_{\text{dyn}}^{f(R)} \right). \quad (4.8)$$

Note that to obtain these relations, the two integrations in Eq. (4.5) have used the same upper limit,  $r = R_{500}^{\text{eff}}$ , for GR and  $f(R)$  gravity, as mentioned above.

He and Li (2016) demonstrated an accuracy  $\approx 3\%$  for the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  relations and  $\approx 13\%$  for  $L_{\text{X}}$ . These quantities are all cumulative: they are computed by summing over the entire volume of a halo up to some maximum radius (in this case  $R_{500}^{\text{eff}}$ ). Cumulative quantities typically depend more on the outer regions, which contain most of the volume and mass, than on the inner regions. Therefore, even though the  $f(R)$  gravity profiles (with appropriate rescaling) do not agree with GR for the inner regions (see Fig. 4.1), this is expected to have a negligible contribution overall to these mass proxies and explains why He and Li (2016) found such a high accuracy for these mappings.

In this chapter, we will expand on these tests by using full-physics hydrodynamical simulations to check how the addition of effects such as cooling and feedback can alter the accuracy of the mappings defined by Eqs. (4.6)-(4.8) and the temperature equivalence given by Eq. (4.4). Our tests with the non-radiative runs can also be used as a check for

consistency with [He and Li \(2016\)](#), who used a different simulation code and  $f(R)$  gravity solver.

### 4.2.2 True density approach

Mappings can also be done for haloes identified with the true density field. For these haloes, the radius,  $R_{500}^{\text{true}}$ , would enclose an average *true* density of 500 times the critical density of the Universe. For haloes in  $f(R)$  gravity and GR with the same true mass,  $M^{\text{GR}} = M_{\text{true}}^{f(R)}$ , the total gravitational potential at  $R_{500}^{\text{true}}$  in  $f(R)$  gravity would be a factor of  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  higher than in GR (where both the dynamical and true mass are measured within  $R_{500}^{\text{true}}$ ). According to the self-similar model predictions, the gas temperature in  $f(R)$  gravity is expected to be higher by the same factor:

$$T_{\text{gas}}^{f(R)}(M_{\text{true}}^{f(R)}) = \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} T_{\text{gas}}^{\text{GR}}(M^{\text{GR}} = M_{\text{true}}^{f(R)}). \quad (4.9)$$

On the other hand, for haloes in  $f(R)$  gravity and GR with the same true mass, the gas density profiles are expected to agree in the outer regions.

To check these assumptions, let us once again examine [Fig. 4.1](#), this time looking at the third and fourth columns from the left, which show the stacked gas density and temperature profiles for groups in mass bins  $10^{13.7} M_{\odot} < M_{\text{true}} (< R_{500}^{\text{true}}) < 10^{14.0} M_{\odot}$  and  $10^{13.0} M_{\odot} < M_{\text{true}} (< R_{500}^{\text{true}}) < 10^{13.3} M_{\odot}$ , respectively. The radial range is shown up to the mean logarithm of  $R_{500}^{\text{true}}$  for all profiles. Referring to the bottom two rows, which show the non-radiative and full-physics temperature profiles, it appears that the  $f(R)$  gravity profiles with the  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  rescaling (shown by the dashed curves) show reasonable agreement with GR in the outer regions. Again, the only exception is for the high-mass bin of the full-physics data, where there is a small disparity between F5 and GR. Looking at the top two rows, the  $f(R)$  gravity gas density profiles agree very well with GR in the outer-most regions for both mass bins.

These results for the temperature and gas density profiles yield the following predic-

tions for haloes in GR and  $f(R)$  gravity with  $M^{\text{GR}} = M_{\text{true}}^{f(R)}$ :

$$\begin{aligned} & \int_0^r dr' 4\pi r'^2 \left( \rho_{\text{gas}}^{f(R)} \right)^a \left( T_{\text{gas}}^{f(R)} \right)^b \\ & \approx \left( \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} \right)^b \int_0^r dr' 4\pi r'^2 \left( \rho_{\text{gas}}^{\text{GR}} \right)^a \left( T_{\text{gas}}^{\text{GR}} \right)^b, \end{aligned} \quad (4.10)$$

where this time the two integrations both have upper limit  $r = R_{500}^{\text{true}}$ . This prediction yields the following new mappings between the mass scaling relations in  $f(R)$  gravity and GR:

$$Y_{\text{SZ}}^{f(R)} \left( M_{\text{true}}^{f(R)} \right) \approx \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} Y_{\text{SZ}}^{\text{GR}} \left( M^{\text{GR}} = M_{\text{true}}^{f(R)} \right), \quad (4.11)$$

$$Y_{\text{X}}^{f(R)} \left( M_{\text{true}}^{f(R)} \right) \approx \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} Y_{\text{X}}^{\text{GR}} \left( M^{\text{GR}} = M_{\text{true}}^{f(R)} \right), \quad (4.12)$$

$$L_{\text{X}}^{f(R)} \left( M_{\text{true}}^{f(R)} \right) \approx \left( \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} \right)^{1/2} L_{\text{X}}^{\text{GR}} \left( M^{\text{GR}} = M_{\text{true}}^{f(R)} \right). \quad (4.13)$$

For the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  mappings, the  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  factor comes from the dependence on the gas temperature to power one in Eqs. (4.2) and (4.3). On the other hand, the X-ray luminosity, given by Eq. (4.1), depends on the gas temperature to power half, which means that the corresponding  $f(R)$  gravity and GR scaling relations are expected to differ by factor  $\left( M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)} \right)^{1/2}$  only. In Sec. 4.4, we show the results of our tests of these alternative predictions using both our non-radiative and full-physics simulations.

In this section, we have referred to two different definitions of the halo radius:  $R_{500}^{\text{eff}}$  and  $R_{500}^{\text{true}}$ . The radius  $R_{500}^{\text{eff}}$  is defined in terms of the effective density field. For an unscreened halo in  $f(R)$  gravity, the effective density is up to 4/3 times greater than the true density. As such,  $R_{500}^{\text{eff}}$  is typically a higher radius than  $R_{500}^{\text{true}}$  for haloes in  $f(R)$  gravity. On the other hand, the two radii are exactly the same in GR, where the effective density field is equivalent to the true one.

## 4.3 Simulations and methods

In Sec. 4.3.1, we describe the non-radiative and full-physics simulations that are used in this chapter. Then, in Sec. 4.3.2, we describe how we have measured the cluster mass and

four observable mass proxies from these simulations.

### 4.3.1 Simulations

The results discussed in this chapter were generated using a subset of the SHYBONE simulations (Arnold et al., 2019b; Hernández-Aguayo et al., 2020). These full-physics hydrodynamical simulations, which have been run using AREPO (Springel, 2010), employ the IllustrisTNG galaxy formation model (Weinberger et al., 2017; Pillepich et al., 2018c) and include runs for GR, HS  $f(R)$  gravity and nDGP (we will only use the GR and  $f(R)$  runs in this chapter).

For every full-physics run used in this chapter, we also utilise a non-radiative counterpart which does not include cooling, star formation or stellar and black hole feedback processes. Both the full-physics and non-radiative simulations span a comoving box of length  $62h^{-1}\text{Mpc}$ . These runs each start with  $512^3$  dark matter particles and the same number of initial gas cells, and begin at redshift  $z = 127$ . All results in this chapter are computed at  $z = 0$ . The cosmological parameters have values  $(h, \Omega_M, \Omega_b, \Omega_\Lambda, n_s, \sigma_8) = (0.6774, 0.3089, 0.0486, 0.6911, 0.9667, 0.8159)$ , where  $n_s$  is the power-law index of the primordial density power spectrum. The mass resolution is set by the DM particle mass  $m_{\text{DM}} = 1.28 \times 10^8 h^{-1} M_\odot$  and an average gas cell mass of  $m_{\text{gas}} \approx 2.5 \times 10^7 h^{-1} M_\odot$ . In addition to GR, the runs include the F6 and F5 HS models, all starting from identical initial conditions at  $z = 127$ .

In the calculation of the gas temperature, we have assumed that the primordial hydrogen mass fraction has a value  $X_H = 0.76$  and set the adiabatic index to  $\gamma = 5/3$  (for a monatomic gas). For the non-radiative simulations we assume that the gas is made up of fully ionised hydrogen and helium.

### 4.3.2 Group catalogues

The group catalogues were constructed using SUBFIND (Springel et al., 2001), which locates the gravitational potential minimum of the FOF groups using the true density field. For each group, both radii  $R_{500}^{\text{true}}$  and  $R_{500}^{\text{eff}}$  were computed around the gravitational potential



minimum, enclosing, respectively, average *true* and *effective* densities of 500 times the critical density of the Universe. For each radius definition, the total enclosed dynamical and true masses were measured, in addition to the group observables. Quantities measured within  $R_{500}^{\text{eff}}$  have been used to test the scaling relation mappings from the effective density approach described in Sec. 4.2.1, while the quantities measured within  $R_{500}^{\text{true}}$  have been used to test the predictions of the true density approach discussed in Sec. 4.2.2.

In the computation of the halo temperature, we have excluded the core regions in which the complex thermal and dynamical processes during cluster formation and evolution can lead to a significant degree of dispersion between the halo temperature profiles. We have set the core region to the radial range  $r < 0.15R$ , where  $R$  can be either  $R_{500}^{\text{eff}}$  or  $R_{500}^{\text{true}}$ . This range is consistent with previous studies of cluster scaling relations (e.g., Fabjan et al., 2011; Le Brun et al., 2017; Truong et al., 2018).

The halo gas temperature has been computed using a mass-weighted average:

$$\bar{T}_{\text{gas}} = \frac{\sum_i m_{\text{gas},i} T_i}{\sum_i m_{\text{gas},i}}, \quad (4.14)$$

where  $m_{\text{gas},i}$  and  $T_i$  are, respectively, the mass and temperature of gas cell  $i$ . The summations have been performed over all gas cells whose positions fall within the radial range  $0.15R < r < R$ . The integrated SZ flux is given by:

$$Y_{\text{SZ}} = \frac{\sigma_{\text{T}}}{m_{\text{e}} c^2} \sum_i N_{\text{e},i} T_i, \quad (4.15)$$

where  $N_{\text{e},i}$  is the number of electrons in gas cell  $i$  and the sum includes the same cells as for  $\bar{T}_{\text{gas}}$ . The X-ray analogue of the integrated SZ flux is equal to the product of the total gas mass  $M_{\text{gas}}$ , of all gas cells within  $R_{500}$ , and  $\bar{T}_{\text{gas}}$ :

$$Y_{\text{X}} = M_{\text{gas}} \times \bar{T}_{\text{gas}}. \quad (4.16)$$

Finally, the X-ray luminosity is calculated using:

$$L_{\text{X}} = \sum_i m_{\text{gas},i} \rho_{\text{gas},i} T_i^{1/2}, \quad (4.17)$$

where  $\rho_{\text{gas},i}$  is the gas density of gas cell  $i$  and the summation is performed over the same gas cells as for the  $\bar{T}_{\text{gas}}$  calculation.

## 4.4 Results

In Sec. 4.4.1, we discuss our results for the cluster scaling relations in HS  $f(R)$  gravity. Then, in Sec. 4.4.2, we test the validity of our analytical tanh formula for the dynamical mass enhancement, given by Eq. (2.17), in the presence of full physics. There we will also present an example in which we map between the GR and  $f(R)$  mass scaling relations based on this approximate fitting formula, rather than the actual values of  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  from the simulations. Finally, in Sec. 4.4.3, we propose a new test of gravity using the  $Y_X$ - $\bar{T}_{\text{gas}}$  relation, which does not require inferences of the cluster mass.

### 4.4.1 Scaling relations

Using our simulation data, we have tested the scaling relation mappings described in Sec. 4.2. Due to the small box size,  $62h^{-1}\text{Mpc}$  (comoving), of our simulations, we can only examine haloes with mass  $M_{500} \lesssim 10^{14.5}M_{\odot}$ . We show all objects with  $M_{500} \geq 10^{13}M_{\odot}$  (groups and clusters) in Figs. 4.2–4.5, which typically includes  $\sim 100$  haloes for a given model. Note that it is difficult to rigorously test our scaling relation mappings for the cluster regime ( $M_{500} \gtrsim 10^{14}M_{\odot}$ ), where there are only 5–10 haloes in the present simulations. However, for F5, groups are typically unscreened and low-mass clusters are partially screened, while for F6 low-mass groups are partially screened and higher-mass objects are completely screened; haloes with  $M_{500} \gtrsim 10^{14.5}M_{\odot}$  will be mostly screened for F5 and entirely screened for F6 (see, e.g., Fig. 4.6 below and Fig. 2.3). Therefore, while we do not have a significant number of such large cluster-sized objects, we do expect the scaling relations calibrated for GR to be valid for them.

In addition to showing data points for individual haloes, all plots include curves showing a moving average. This is computed using a moving window of fixed size equal to 10 haloes, for which the mean logarithm of the mass and the median proxy are displayed. We note that the highest-mass haloes have been included in the moving average, even though the highest mean mass is only  $\sim 10^{14}M_{\odot}$ . We also show sub-plots with the smoothed relative difference between the  $f(R)$  and GR curves, as well as the halo scatter in GR. The latter

is computed by fitting a linear model to the GR data and computing the root-mean-square residuals within mass bins.

For the panels labelled ‘effective’ in Figs. 4.2–4.5, all measurements of the mass and observables have been taken within  $R_{500}^{\text{eff}}$  (see Sec. 4.2.1), and the relations are plotted against the dynamical mass. This allows the effective density mappings given by Eqs. (4.4) and (4.6)–(4.8), originally proposed by He and Li (2016), to be tested. On the other hand, an outer radius  $R_{500}^{\text{true}}$  (see Sec. 4.2.2) is imposed for all measurements for the data displayed in the panels labelled ‘true’. These are plotted against the true mass, and can be used to test the true density mappings given by Eqs. (4.9) and (4.11)–(4.13).

#### 4.4.1.1 *Temperature scaling relations*

The results for the gas temperature scaling relations are shown in Fig. 4.2. The non-radiative data is displayed in the left two columns and the full-physics data is shown in the right two columns. For all models and hydrodynamical schemes, the data follows a power-law behaviour, as expected from the self-similar model. The correlation is particularly tight for the non-radiative data, with an overall scatter of 7%. The non-radiative runs contain gas and dark matter particles, but do not feature baryonic processes (apart from basic hydrodynamics) such as radiative cooling, stellar and black hole feedback and star formation. It is therefore expected that the thermodynamical properties can be largely determined from the gravitational potential, which is observed in the results. On the other hand, there is  $\sim 10\%$  overall scatter in the full-physics data, and the gas temperature is typically higher. This can be explained by the inclusion of feedback mechanisms which act as an additional source of heating of the surrounding gas and cause some departures from self-similarity. These mechanisms have a stronger effect on lower-mass haloes, resulting in a particularly high (10–20%) scatter for these objects.

For the effective density approach, Eq. (4.4) is expected to hold: the temperature is predicted to be equal for haloes in GR and  $f(R)$  gravity with the same dynamical mass. In Fig. 4.2, the non-radiative and full-physics results from our effective catalogue are shown in the first and third columns from the left, respectively. For both F6 and F5, there is excellent

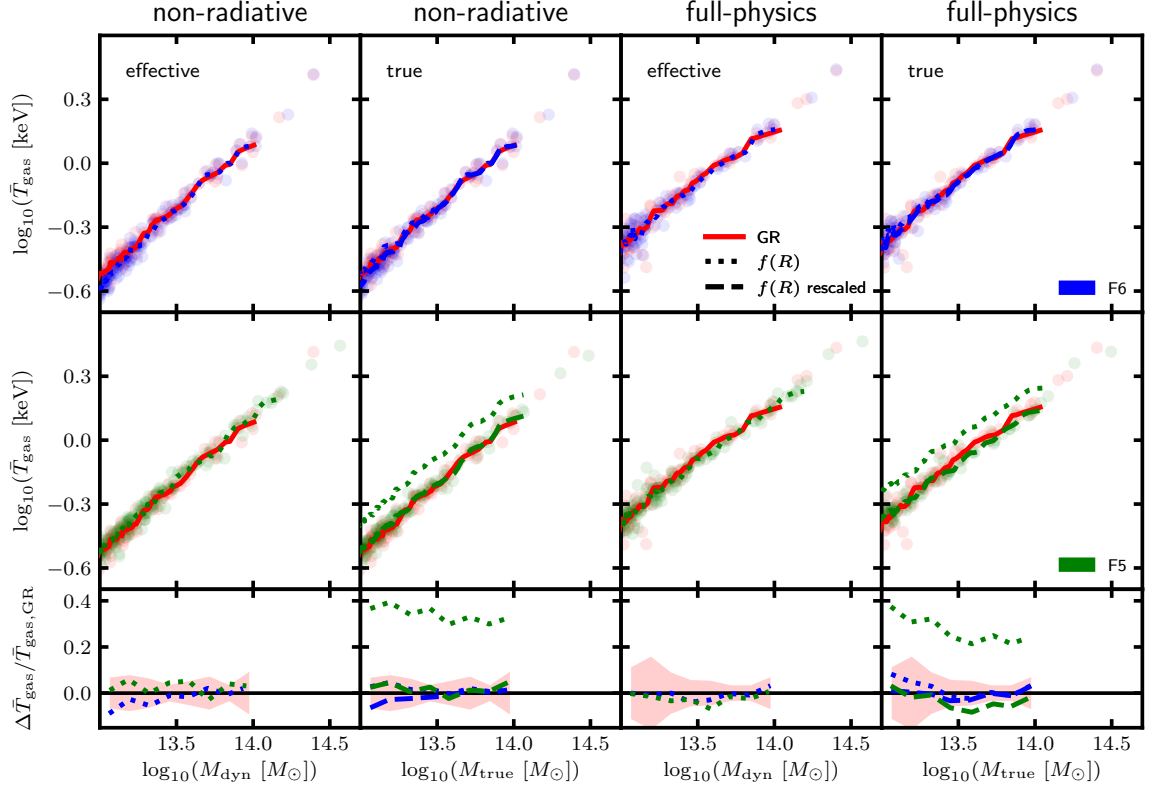


Figure 4.2: Gas temperature plotted as a function of mass for FOF haloes from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1). The curves correspond to the median mass-weighted temperature and the mean logarithm of the mass computed within a moving window of fixed size equal to 10 haloes. Data has been included for GR (red solid lines) together with the F6 (blue lines) and F5 (green lines)  $f(R)$  gravity models. Rescalings to the  $f(R)$  gravity temperature have been carried out as described in Secs. 4.2.1 and 4.2.2. For the ‘true’ density approach, the rescaled data (dashed lines) is shown along with the unaltered data (dotted lines). For this data, the mass corresponds to the total true mass within the radius  $R_{500}^{\text{true}}$ , and the temperature has also been computed within this radius. For the ‘effective’ density approach, no rescaling is necessary, the mass corresponds to the total dynamical mass within  $R_{500}^{\text{eff}}$ , and the temperature has also been computed within this radius. Data points are displayed, with each point corresponding to a GR halo (red points) or to a halo in F6 (blue points) or F5 (green points), including the rescaling for the ‘true’ density data. *Bottom row*: the smoothed relative difference between the  $f(R)$  gravity and GR curves in the above plots; the red shaded regions indicate the size of the halo scatter in GR.

agreement with the GR data, with typical agreement  $\lesssim 5\%$ . This agreement for the non-radiative data backs up the findings from [He and Li \(2016\)](#), while the full-physics results do not show clear evidence for a departure from Eq. (4.4) caused by feedback processes and cooling. These results agree with the self-similar model predictions: two haloes in  $f(R)$  gravity and GR which have the same dynamical mass  $M_{\text{dyn}}$  (and therefore the same radius  $R_{500}^{\text{eff}}$ ) also have the same gravitational potential,  $\phi = GM_{\text{dyn}}/R_{500}^{\text{eff}}$ .

In order to test the new mappings predicted by the true density approach, the temperature of each halo in  $f(R)$  gravity has been divided by the mass ratio  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$ . In Fig. 4.2, both the data for individual haloes and corresponding moving averages are shown with this rescaling applied (dashed lines), along with the moving averages for the unaltered data (dotted lines). It is expected, from Eq. (4.9), that the data with the rescaling should agree with GR. For the non-radiative results in the second column, this indeed appears to be the case, with an excellent agreement that is generally within just a few percent. For the full-physics data there is still reasonable  $\lesssim 10\%$  agreement, but the F5 temperature appears to be lower than the GR temperature for  $\log_{10} [M_{\text{true}} (< R_{500}^{\text{true}}) / M_{\odot}] \gtrsim 13.5$ .

This small deviation is consistent with the full-physics temperature profiles shown in Fig. 4.1. The plots in the bottom right of that figure show the temperature profiles with the  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  rescaling applied. The profiles are shown for the halo mass bins  $10^{13.7} M_{\odot} < M_{\text{true}} (< R_{500}^{\text{true}}) < 10^{14.0} M_{\odot}$  and  $10^{13.0} M_{\odot} < M_{\text{true}} (< R_{500}^{\text{true}}) < 10^{13.3} M_{\odot}$ . For the high-mass full-physics profiles, the rescaled F5 profile is clearly lower than the profile in GR across most of the radial range. This can explain the lower rescaled F5 temperature observed at the high-mass end of the full-physics data. On the other hand, for the lower-mass bin with full-physics and for the non-radiative data the agreement between the rescaled  $f(R)$  gravity and GR temperature profiles is very good, particularly at the outer radii which have greater overall contribution to the mass-weighted temperature. Similar agreement is shown between the  $f(R)$  gravity and GR profiles for the plots in the bottom-left of Fig. 4.1. Again, for the high-mass full-physics data there is some deviation between F5 and GR, particularly in the outermost regions. But this is not as noticeable as for the profiles with the true density rescalings.

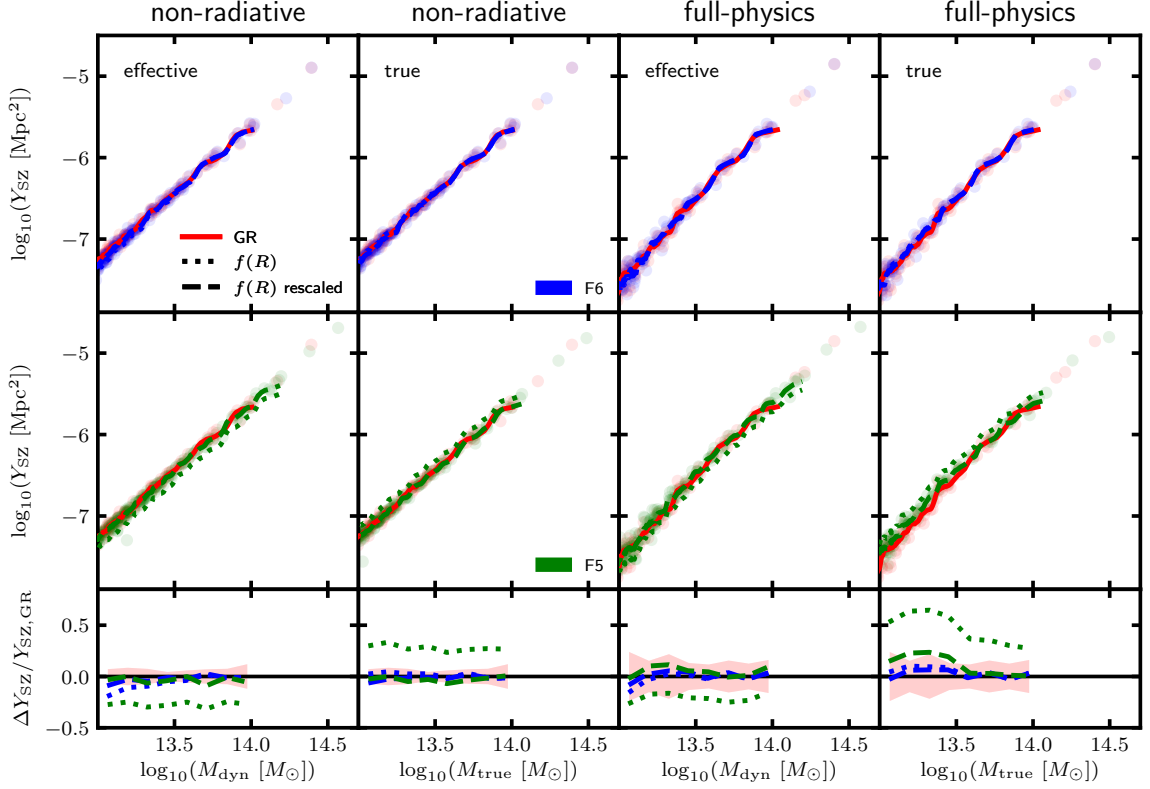


Figure 4.3: Compton  $Y$ -parameter of the SZ effect plotted as a function of mass for FOF haloes from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1). Apart from the  $Y_{\text{SZ}}$  parameter being shown rather than the mass-weighted temperature, the layout of this figure is otherwise identical to Fig. 4.2. The rescalings of the  $f(R)$   $Y_{\text{SZ}}$  data are carried out as described in Sec. 4.2.

The small difference between F5 and GR is likely to be caused by a difference in the levels of feedback – which itself is determined by the interrelations between modified gravity (including screening or the lack of it) and baryonic physics – in these higher-mass haloes for the two models. Encouragingly, this appears to have only a small effect on the effective density data, where there is good agreement between F5 and GR for high-mass groups. However, the effect is greater for the true density data. To understand the implications that this could have on tests of gravity using the cluster regime, we will need simulations with a larger box size. This will be addressed in Chapter 8, where we will use a re-calibrated full-physics model to probe masses up to  $M_{500} \sim 10^{15} M_{\odot}$ .

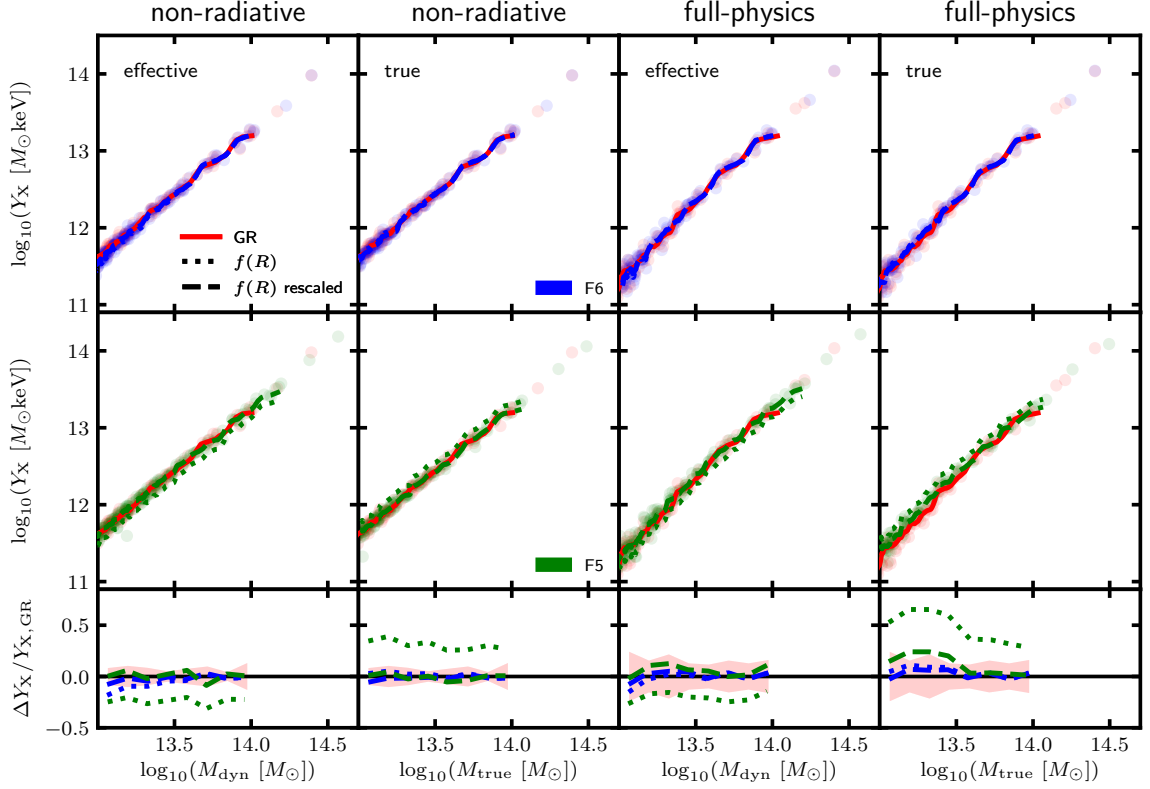


Figure 4.4: X-ray analogue of the Compton  $Y$ -parameter plotted as a function of mass for FOF haloes from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1). Apart from the  $Y_{\text{X}}$  parameter being shown rather than the mass-weighted temperature, the layout of this figure is otherwise identical to Fig. 4.2. The rescalings of the  $f(R)$   $Y_{\text{X}}$  data are carried out as described in Sec. 4.2.

#### 4.4.1.2 $Y_{\text{SZ}}$ and $Y_{\text{X}}$ scaling relations

Our results for the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  scaling relations are shown in Figs. 4.3 and 4.4. The  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  parameters are, by definition, tightly correlated. Their results therefore follow similar patterns, and both show very tight correlations with the halo mass, with a scatter of  $\sim 8\%$  and  $\sim 19\%$  for the non-radiative and full-physics data, respectively. There are also no clear outliers in the full-physics data, unlike for the temperature and the X-ray luminosity data (see below). This is because of the competing effects of feedback processes on the gas density and gas temperature (Fabjan et al., 2011). Comparing the non-radiative and full-physics profiles in Fig. 4.1, it can be seen that the additional processes in the full-physics runs cause haloes to have a lower gas density, particularly at the inner regions, and a higher gas temperature. This is caused by stellar and black hole feedbacks, which generate high-

energy winds that heat up the surrounding gas and blow it out from the central regions. Such competing effects are approximately cancelled out in the product of the gas density with the gas temperature, as in Eqs. (4.2) and (4.3).

In order to test the mappings predicted by the effective density approach, given by Eqs. (4.6) and (4.7), the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  values measured for  $f(R)$  gravity have been multiplied by the mass ratio  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$ . For the non-radiative plots in Figs. 4.3 and 4.4, there is excellent agreement between this rescaled data and GR. There is also a strong agreement for the higher-mass full-physics data. For the mass range  $13.2 < \log_{10}(M_{\text{dyn}}(< R_{500}^{\text{eff}})M_{\odot}^{-1}) < 13.5$ , however, there is some disparity of  $\lesssim 20\%$  between the rescaled F5 data and GR.

A similar level of accuracy is observed for the mappings given by Eqs. (4.11) and (4.12), which are predicted by the true density approach. To test these,  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  are divided by  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  to generate rescaled data for F6 and F5. Again, this data shows excellent agreement, within a few percent, with GR for the non-radiative simulations and the high-mass end of the full-physics data. But for the lower-mass full-physics data there is a significant disagreement between F5 and GR of up to  $\sim 30\%$ , which is higher than for the effective density rescalings.

The disparities found in the low-mass full-physics data can be explained using Fig. 4.1. Looking at the full-physics profiles for the true density mass bins, it is observed that for a large portion of the inner halo regions the gas density is higher in F5 than in GR. These profiles converge at  $r \approx 10^{2.5}\text{kpc}$  for both mass bins. The high-mass haloes have higher overall radius  $R_{500}^{\text{true}}$ , which means that the profiles are converged for a large portion of the outer radii. This means the disparities at lower radii have a negligible overall contribution to the integrals for  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$ . But for the lower-mass haloes, the profiles are converged only for a small portion of the overall radius range, causing  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  to be greater in F5 than in GR. A similar reasoning can be used to explain the disparities for the results with the effective density rescaling, although the difference in agreement at the outer regions for each mass bin is not quite as substantial here, which explains why the full-physics data for the effective density approach shows less overall deviation between F5 and GR in Figs. 4.3 and 4.4.



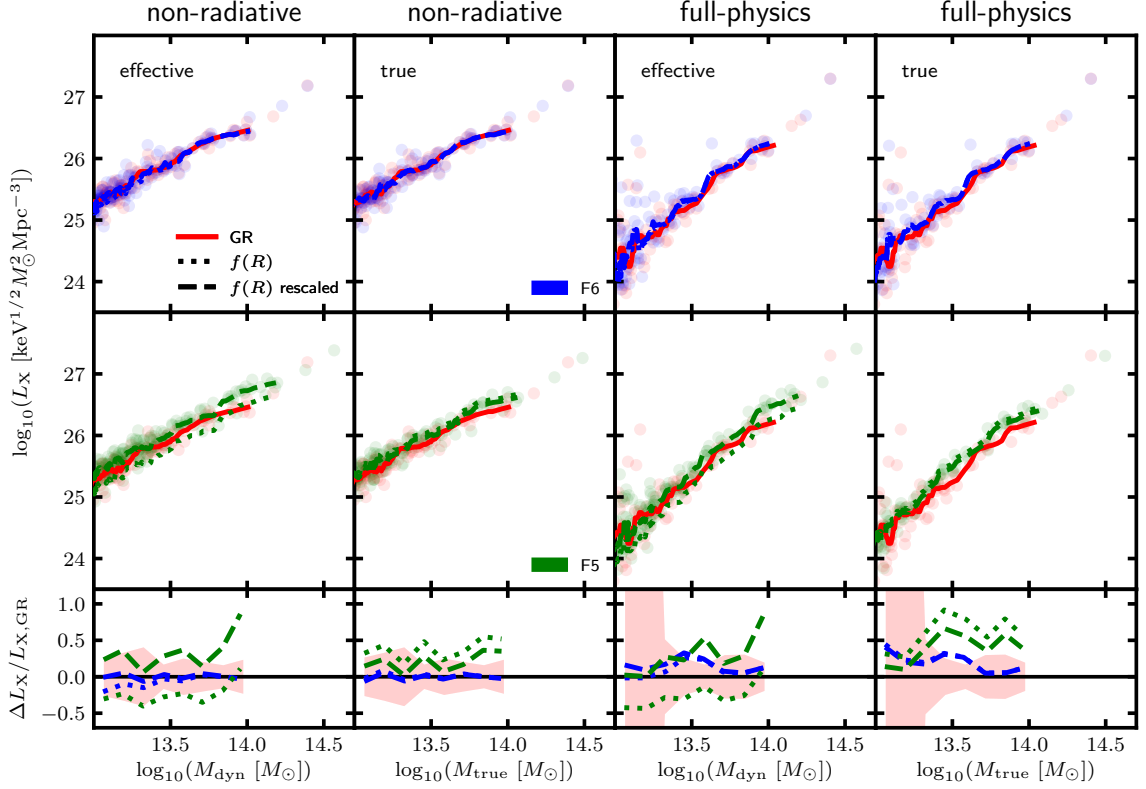


Figure 4.5: X-ray luminosity plotted as a function of the mass for FOF haloes from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1). Apart from the X-ray luminosity being shown rather than the mass-weighted temperature, the layout of this figure is otherwise identical to Fig. 4.2. The rescalings of the  $f(R)$   $L_X$  data are carried out as described in Sec. 4.2.

The difference between the full-physics  $f(R)$  and GR scaling relations at low mass is likely to be explained by baryonic processes such as feedback which are absent in the non-radiative simulations. However, in studies of clusters, these lower-mass groups are of less interest. The strong agreement at the higher masses is therefore very encouraging for our framework to constrain  $f(R)$  gravity using the high-mass end of the halo mass function.

#### 4.4.1.3 X-ray luminosity scaling relations

The results for the X-ray luminosity scaling relations are shown in Fig. 4.5. Compared with the temperature,  $Y_{SZ}$  and  $Y_X$  data, the X-ray luminosity is much more scattered, with particularly large dispersion in lower-mass haloes and  $\sim 25\%$  scatter at higher masses. One explanation for this is that the X-ray luminosity, defined in Eq. (4.1), depends on

the gas density to power two. This means that the inner regions of the group, which have a higher gas density than the outer regions, have a greater overall contribution to the  $L_X$  integral than for the other observables discussed in this chapter. The inner halo regions are expected to be more impacted by unpredictable dynamical processes during cluster formation, including halo mergers. In particular, they are more prone to gas heating and blowing-out of gas caused by feedback mechanisms. While the competing effects of these processes on the gas density and gas temperature roughly cancel for the  $Y_{SZ}$  and  $Y_X$  observables, this is not the case for  $L_X$ , which depends on the gas density to power two and the gas temperature to power half. This results in a number of significant outliers, as can be seen in the full-physics data of Fig. 4.5.

For the mapping defined using the effective density field, given by Eq. (4.8), it has been expected that the GR X-ray luminosity should be equal to the  $f(R)$  gravity value multiplied by the factor  $\left(M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}\right)^2$ . From Fig. 4.5, the rescaled data in F5 appears to be higher than in GR by  $\sim 30\%$  on average for both the non-radiative and the full-physics simulations. A similar level of deviation is also observed for the true density results, where Eq. (4.13) predicts that the GR and  $f(R)$  gravity X-ray luminosity should be equal after the values in  $f(R)$  gravity are divided by the factor  $\left(M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}\right)^{1/2}$ . Again, the rescaled F5 X-ray luminosity is significantly greater than in GR on average.

As for the  $Y_{SZ}$  and  $Y_X$  mappings, the disparity found here can be explained by looking at the gas density profiles in Fig. 4.1. For both the non-radiative and full-physics data, the gas density in the inner halo regions is greater for F5 (with appropriate rescaling applied) than for GR. Because the inner regions have a greater contribution to the X-ray luminosity than for other proxies, as described above, this causes these differences in the inner regions to become significant overall, even for the non-radiative data for which the F5 and GR profiles are converged above a lower radius. This results in the general offset for the full range of masses as shown in Fig. 4.5. As described above, the X-ray luminosity is also more strongly influenced by feedback processes, which can further increase the offset between F5 and GR if the feedback behaves differently in these two models.

Our observation that the mappings have a poorer performance for the X-ray luminosity

than for the other proxies is consistent with [He and Li \(2016\)](#), who observed a disparity of  $\sim 13\%$ . Unless further corrections are applied to account for the unpredictable effects of feedback in the mappings, therefore, the X-ray luminosity is unlikely to be a reliable proxy for mass determination in accurate cluster tests of  $f(R)$  gravity.

#### 4.4.1.4 *Further comments*

As described in Sec. 4.3.2, we have excluded the core region of  $r < 0.15R_{500}$  when calculating the thermal properties of our simulated FOF groups. We have also experimented excluding core regions of size  $r < 0.1R_{500}$  and  $r < 0.2R_{500}$ : for the scaling relations  $Y_{\text{SZ}}-M$ ,  $Y_X-M$  and  $\bar{T}_{\text{gas}}-M$ , the relative differences between  $f(R)$  gravity and GR are barely affected; on the other hand, the effect is larger for the  $L_X-M$  scaling relation, because  $L_X$  is more sensitive to the inner halo regions than the other proxies. However, the  $L_X-M$  relation is not ideal for reliable tests of gravity anyway, as noted above, and is mainly included in this thesis for completeness and for comparison with the other relations. Even if the entire core is included in the calculations, we have found that the effect on the  $Y_{\text{SZ}}-M$  and  $Y_X-M$  results is still very small, providing further confirmation that these relations can be used for reliable tests of gravity.

The scatter of the full-physics GR scaling relations shown in Figs. 4.2–4.5 is typically higher than the scatter quoted in recent studies that have also used simulations which include star formation, cooling and stellar and black hole feedback. For example, we observe a root-mean-square dispersion of  $\sim 19\%$  for the  $Y_{\text{SZ}}-M$  and  $Y_X-M$  relations, while [Le Brun et al. \(2017\)](#) and [Truong et al. \(2018\)](#) reported  $\sim 10\%$  and  $\sim 15\%$ , respectively. This is likely to be caused by our restricted halo population (modified gravity simulations are much more computationally expensive than their standard gravity counterparts which limits the affordable box-size and resolution), which contains a large number of low-mass groups that are more susceptible to feedback. Our results suggest that the  $\bar{T}_{\text{gas}}-M$  relation has the lowest scatter and the  $L_X-M$  relation has the highest scatter, and this is consistent with the above works.

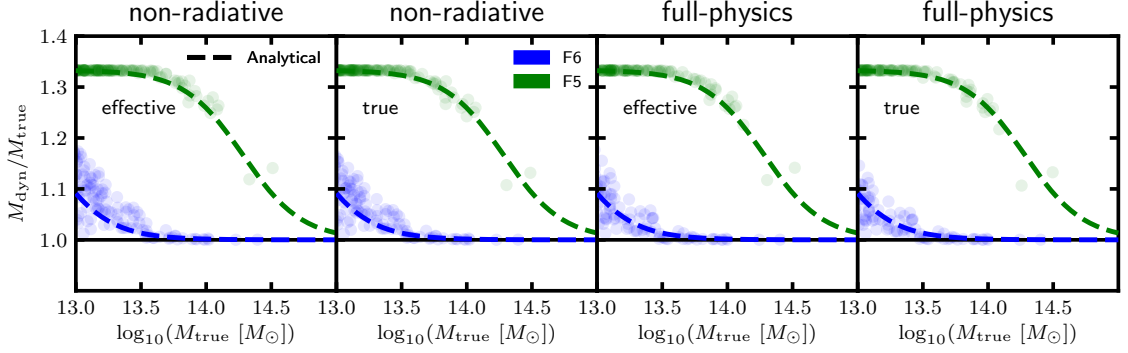


Figure 4.6: Ratio of the dynamical mass to the true mass of haloes plotted as a function of the true mass. The data points correspond to FOF haloes from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1 for details). Data is included for F6 (blue) and F5 (green). For the data labelled ‘effective’, the dynamical and true mass have been measured within the radius  $R_{500}^{\text{eff}}$  (defined in Sec. 4.2.1), while the radius  $R_{500}^{\text{true}}$  (defined in Sec. 4.2.2) has been used for the data labelled ‘true’. Analytical predictions for the mass enhancement have been computed using Eq. (2.17) and are shown (dashed curves) for each model.

#### 4.4.2 Halo mass ratio calculation

For the results discussed in Sec. 4.4.1, the rescalings to the observables in  $f(R)$  gravity have been computed using direct measurements of the true mass and the dynamical mass from the simulations. However, for studies of clusters using real observations, measurements of both the true mass and dynamical mass are unlikely to be available. In this case, our analytical model for the ratio of the dynamical mass to the true mass, given by Eq. (2.17), can be used.

This model was calibrated using a suite of DMO simulations. To check how it performs for data that includes full physics, we have plotted the model predictions on top of actual measurements of the dynamical mass enhancement for the FOF groups in the SHYBONE simulations. This is shown in Fig. 4.6. Data for both the effective density and true density catalogues have been included, for which the dynamical and true halo masses have been measured within  $R_{500}^{\text{eff}}$  and  $R_{500}^{\text{true}}$ , respectively. We have made use of all available data with  $M_{\text{true}} > 10^{13} M_{\odot}$ , including haloes with  $M_{\text{true}} \sim 10^{14.5} M_{\odot}$ . These results indicate that there is very good agreement between the analytical predictions and the actual data for both F6 and F5, regardless of the hydrodynamical scheme that is employed. Interestingly,

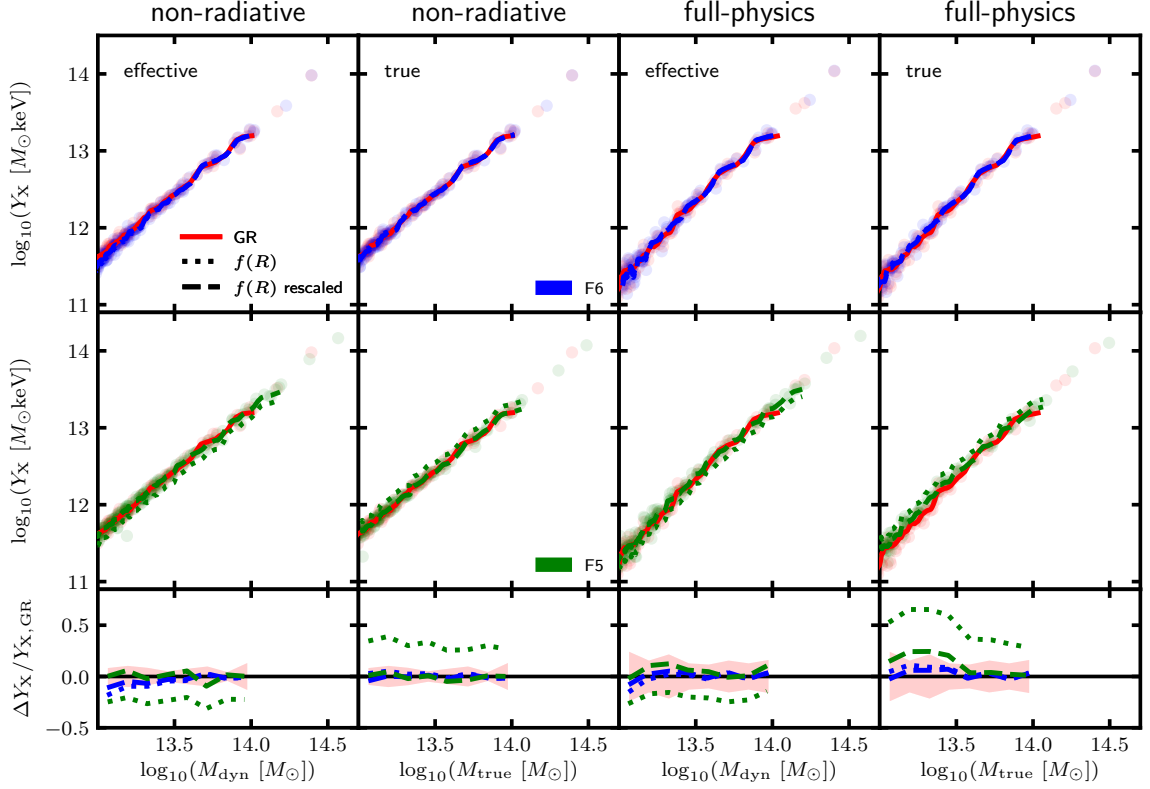


Figure 4.7: X-ray analogue of the Compton  $Y_X$ -parameter plotted as a function of the mass for FOF haloes from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1). The layout and format of this figure are identical to those of Fig. 4.4. The results shown here are also mostly the same as for Fig. 4.4, however the rescalings of  $Y_X$  in the  $f(R)$  gravity data have been generated using the analytical tanh formula given by Eq. (2.17).

even though Eq. (2.17) was originally calibrated using measurements of the dynamical and true mass within  $R_{500}^{\text{true}}$ , it still performs very well for data measured within  $R_{500}^{\text{eff}}$ , which is typically a higher radius.

We have also tested the mappings of Eqs. (4.7) and (4.12) for the  $Y_X$  parameter, with Eq. (2.17) used to compute the required rescalings to the  $f(R)$  gravity data. This is shown in Fig. 4.7. From comparing this plot with Fig. 4.4, it can be seen that there is almost no difference in the rescaled data in both figures. This confirms that Eq. (2.17) can be applied to derive the mappings between GR and  $f(R)$  scaling relations for, at least, the mass range  $10^{13} M_{\odot} < M_{500} < 10^{14} M_{\odot}$ . Given the very good agreement up to  $10^{14.5} M_{\odot}$  shown in Fig. 4.6, it is expected that our formula can be applied in the cluster regime as well. However, this should be tested more rigorously using full-physics simulations with a

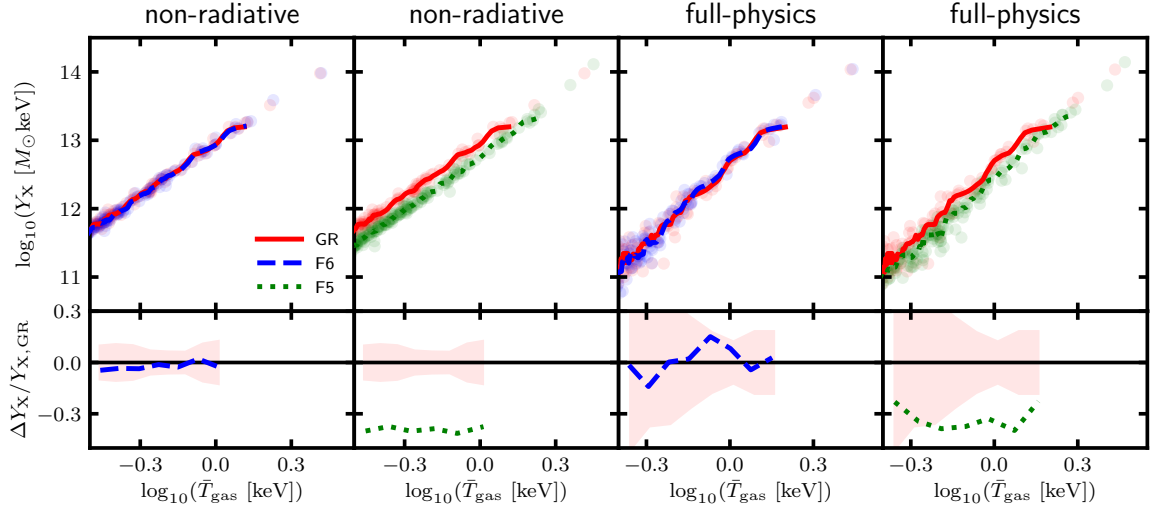


Figure 4.8: The X-ray analogue of the Compton  $Y$ -parameter plotted as a function of the mass-weighted temperature for FOF haloes from the non-radiative and full-physics SHYBONE simulations (see Sec. 4.3.1). The curves correspond to the median  $Y_X$  versus the mean logarithm of the temperature computed within a moving window of fixed size equal to 10 haloes. Data has been included for GR (red solid lines) together with the F6 (blue lines) and F5 (green lines)  $f(R)$  gravity models.  $Y_X$  and the temperature have been computed within the radius  $R_{500}^{\text{true}}$  (see Sec. 4.2.2). Data points are displayed, with each point corresponding to a GR halo (red points), or to a halo in F6 (blue points) or F5 (green points). Bottom row: the smoothed relative difference between the  $f(R)$  gravity and GR curves in the above plots; the red shaded regions indicate the size of the halo scatter in GR.

larger box size.

#### 4.4.3 $Y_X$ -temperature scaling relation

So far, we have only considered scaling relations that can be used to infer the cluster mass, which is a vital ingredient for tests of gravity that use the cluster abundance (Fig. 1.2). However, tests of gravity can also be conducted using the relations themselves. For example, Hammami and Mota (2017); Del Popolo et al. (2019) used the temperature-mass scaling relation to probe screened MG models. The cluster mass can be determined using other observations, such as weak lensing, making such scaling relations observable, and our full-physics results (Figs. 4.2–4.4) confirm that the  $\bar{T}_{\text{gas}}-M$  and  $Y_X-M$  (and  $Y_{\text{SZ}}-M$ ) relations can be used as reliable probes on group and cluster scales, with differences between GR and F5 in the range 20-50% (when no rescaling is applied).

However, scaling relations which do not involve the mass can also be modelled and used.

In Fig. 4.8, we show the scaling relation between the  $Y_X$  parameter and the gas temperature, where both observables have been computed within the radius  $R_{500}^{\text{true}}$ . A significant relative difference of 30-40% is observed between the GR and F5 models for both the non-radiative and full-physics data, indicating that this relation can offer a powerful test of gravity using group- and cluster-sized objects. A key advantage of using the  $Y_X-\bar{T}_{\text{gas}}$  (or  $Y_{\text{SZ}}-\bar{T}_{\text{gas}}$ ) scaling relation is that it does not involve measuring the cluster mass, and hence no need for mass calibrations or synergies with other observations such as weak lensing. It also has a relatively low scatter compared to, for example, the  $L_X-\bar{T}_{\text{gas}}$  relation that was considered by Arnold et al. (2014).

## 4.5 Summary, Discussion and Conclusions

In this chapter, we have made use of the first full-physics simulations that have been run for both GR and  $f(R)$  gravity (along with non-radiative counterparts), to study the effects of the fifth force of  $f(R)$  gravity on the scaling relations between the cluster mass and four observable proxies: the gas temperature (Fig. 4.2), the  $Y_{\text{SZ}}$  and  $Y_X$  parameters (Figs. 4.3 and 4.4) and the X-ray luminosity (Fig. 4.5). To understand these effects in greater detail, we have also examined the effects of both  $f(R)$  gravity and full-physics on the gas density and temperature profiles (see Fig. 4.1). In doing so, we have been able to test two methods for mapping between scaling relations in  $f(R)$  gravity and GR.

The first method was proposed by He and Li (2016). This proposes a set of mappings, given by Eqs. (4.4) and (4.6)-(4.8), that can be applied to haloes whose mass and radius are measured using the effective density field (see Sec. 4.2.1). A second, new, approach is proposed in Sec. 4.2.2, and predicts another set of mappings, given by Eqs. (4.9) and (4.11)-(4.13), that can be applied to haloes whose mass and radius are measured using the true density field. Both sets of mappings involve simple rescalings that depend only on the ratio of the dynamical mass to the true mass in  $f(R)$  gravity. As shown by Figs. 4.6 and 4.7, even with the inclusion of full-physics processes this ratio can be computed with high accuracy using our analytical tanh formula, which is given by Eq. (2.17).

For the mass-weighted gas temperature and the  $Y_{\text{SZ}}$  and  $Y_X$  observables, we found that

the F6 and F5 scaling relations, with appropriate rescaling applied (using either method discussed above), match the GR relations to within a few percent for the full mass-range tested for the non-radiative simulations. With the inclusion of full-physics effects such as feedbacks, star formation and cooling, the rescaled  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  scaling relations continue to show excellent agreement with GR for mass  $M_{500} \gtrsim 10^{13.5} M_{\odot}$ , which includes group- and cluster-sized objects. These proxies also show relatively low scatter as a function of the cluster mass, compared with other observables.  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  are therefore likely to be suitable for accurate determination of the cluster mass in tests of  $f(R)$  gravity. The mappings for the gas temperature show a very high accuracy for lower-mass objects, but show a small  $\lesssim 5\%$  offset between F5 and GR for higher-mass objects.

The mappings do not work as well for the X-ray luminosity  $L_{\text{X}}$ , for which the F5 relations after rescaling are typically enhanced by  $\sim 30\%$  compared with GR. This is caused by the unique dependency of  $L_{\text{X}}$  on the gas density to power two, and the gas temperature to power half, which means that the inner halo regions have a greater contribution than for the other proxies and the competing effects of feedback on the temperature and gas density profiles are less likely to cancel out. This issue, in addition to the fact that  $L_{\text{X}}$  has a highly scattered correlation with the cluster mass, means that this proxy is unlikely to be suitable for cluster mass determination in tests of  $f(R)$  gravity.

We also considered the  $Y_{\text{X}}-\bar{T}_{\text{gas}}$  scaling relation (Fig. 4.8), and found that this is suppressed by 30-40% in the F5 model relative to GR. This offers a potential new and useful test of gravity with group- and cluster-sized objects which avoids the systematic uncertainties incurred from mass calibration.

We note that the box size  $62h^{-1}\text{Mpc}$  of the simulations used in this chapter is more suited to studying galaxy-sized objects than group- or cluster-sized objects. Indeed, there are only  $\sim 100$  objects with  $M_{500} > 10^{13} M_{\odot}$  and  $\sim 5$ -10 objects with  $M_{500} > 10^{14} M_{\odot}$  in the simulations, making it impossible to test the mappings discussed in this chapter for the most massive galaxy clusters to be observed. In Chapter 8, we will present a re-calibrated full-physics model which can be used to run larger simulations that can be used to reliably probe halo masses up to  $M_{500} \sim 10^{15} M_{\odot}$ .



Our results also provide insights into the viability of extending cluster tests of gravity to the group-mass regime. An advantage of using lower-mass objects is that these objects can be unscreened (or partially screened) even for weaker  $f(R)$  models, offering the potential for tighter constraints using data from ongoing and upcoming SZ and X-ray surveys (e.g., [Merloni et al., 2012](#); [Ade et al., 2016](#)) which are now entering this regime. On the other hand, as we have seen above, the scatter induced by feedback mechanisms becomes more significant in group-sized haloes, which means that additional work will need to be conducted to characterise this effect and to understand its impact on model tests.

Finally, we note that our parameter  $p_2$ , which is used to compute the ratio of the dynamical mass to the true mass, depends only on the quantity  $|\bar{f}_R|/(1+z)$ , and not on the model parameters  $n$  and  $f_{R0}$  of HS  $f(R)$  gravity. This dependence was derived by using the thin-shell model (Chapter 2), which does not depend on the details of the  $f(R)$  model. We therefore expect our scaling relation mappings to perform similarly for any combination of the HS  $f(R)$  parameters, and potentially other chameleon-type or thin-shell-screened models. However, due to the high computational cost of running full-physics simulations of  $f(R)$  gravity and other models, we do not seek to confirm this conjecture in this thesis.

# A self-consistent pipeline for unbiased constraints of $f(R)$ gravity

## 5.1 Introduction

In Chapters 2-4, we used a combination of DMO and full-physics simulations to model the effects of the  $f(R)$  gravity fifth force on the cluster dynamical mass, the halo concentration and the observable-mass scaling relations. These models are core components of our proposed framework for  $f(R)$  constraints using cluster number counts (see Fig. 2.1): our model for the enhancement of the concentration can be used for conversions between cluster mass definitions, which is required if, for example, the theoretical predictions and observations use different spherical overdensities; and our model for the dynamical mass enhancement can be used to predict the  $f(R)$  scaling relation given a GR counterpart relation, and this can be used to relate the observable mass function ( $dn/dY$ ) to the theoretical mass function ( $dn/dM$ ). In this chapter, we will incorporate the remaining components of this framework, including the theoretical model for the HMF in  $f(R)$  gravity and the MCMC sampling used to constrain  $f_{R0}$ , and we will test the full constraint pipeline using mock cluster catalogues generated for both GR and  $f(R)$  fiducial cosmologies.

The chapter is arranged as follows: in Sec. 5.2, we provide an overview of the effects of the  $f(R)$  fifth force on the cluster properties; in Sec. 5.3, we describe our MCMC constraint

pipeline, including the calculation of the log-likelihood and the generation of the mocks; in Sec. 5.4, we present constraints using the GR and  $f(R)$  mocks; then, in Sec. 5.5, we highlight potential sources of bias in our pipeline; finally, we summarise our main findings in Sec. 5.6.

## 5.2 Background

In this section, we summarise the main effects of  $f(R)$  gravity on the properties of galaxy clusters. In Sec. 5.2.1, we recap our model for the dynamical mass enhancement and present a new model for its scatter. Then, in Secs. 5.2.2 and 5.2.3, we recap our model for the concentration enhancement and our mapping between observable-mass scaling relations in  $f(R)$  gravity and GR. Finally, in Sec. 5.2.4, we outline the modelling by Cataneo et al. (2016) for the  $f(R)$  enhancement of the HMF.

### 5.2.1 Dynamical mass enhancement and its scatters

In Chapter 2, we presented our general formula for the ratio  $\mathcal{R}$  of the dynamical mass to the true mass:

$$\mathcal{R} = \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} = \frac{7}{6} - \frac{1}{6} \tanh \left( p_1 \left[ \log_{10} \left( M_{\text{true}}^{f(R)} M_{\odot}^{-1} h \right) - p_2 \right] \right). \quad (5.1)$$

We showed that  $p_1$  is approximately constant, with best-fit value  $2.21 \pm 0.01$ , while the parameter  $p_2$  closely follows the following physically motivated linear relation:

$$p_2 = (1.503 \pm 0.006) \log_{10} \left( \frac{|\bar{f}_R(z)|}{1+z} \right) + (21.64 \pm 0.03). \quad (5.2)$$

For this chapter, we have again used the simulation data from Chapter 2 (Table 2.1) to model the root-mean-square scatter of the dynamical mass enhancement,  $\sigma_{\mathcal{R}}$ . Our model is shown by the solid line in Fig. 5.1 (we provide a detailed description of this model and our fitting procedure in Appendix B.1). The rescaled mass,  $\log_{10}(M_{500} M_{\odot}^{-1} h) - p_2 \equiv \log_{10}(M_{500}/10^{p_2})$ , is expected to take positive values for haloes that are screened and negative values for haloes that are unscreened. The scatter peaks for haloes that are partially screened, with  $\log_{10}(M_{500} M_{\odot}^{-1} h) \sim p_2$ , whereas it falls to roughly zero for lower

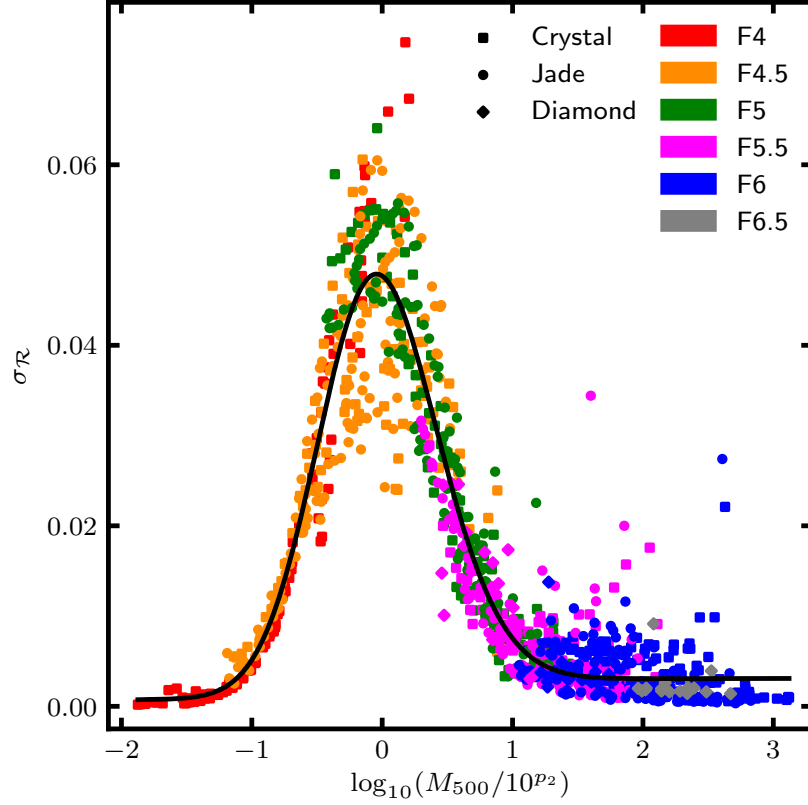


Figure 5.1: Root-mean-square scatter in the dynamical mass enhancement as a function of the rescaled halo mass  $\log_{10}(M_{500}/10^{p_2})$ , where  $p_2$  is given by Eq. (5.2). The data points, which correspond to mass bins spanning  $10^{11}h^{-1}M_{\odot} \lesssim M_{500} \lesssim 10^{15}h^{-1}M_{\odot}$ , have been generated using the DMO simulations Crystal (*squares*), Jade (*circles*) and Diamond (*diamonds*), which are described in Chapter 2. The data spans redshifts  $0 \leq z \leq 1$  and includes present-day scalar field amplitudes  $|f_{R0}| = 10^{-6.5}$  (*grey*),  $10^{-6}$  (*blue*),  $10^{-5.5}$  (*magenta*),  $10^{-5}$  (*green*),  $10^{-4.5}$  (*orange*) and  $10^{-4}$  (*red*). The solid line represents our best-fit model, which is given by Eq. (B.1).

and higher masses. Physically, this makes sense: at sufficiently high masses where *all* haloes are screened and have  $\mathcal{R} \approx 1$ , it follows that the scatter  $\sigma_{\mathcal{R}}$  is very small, and a similar argument can be applied for haloes deep in the unscreened regime. Between these two regimes, the physics is more complicated, giving rise to greater dispersion in the chameleon screening; for example, haloes which do not have a high enough mass to be self-screened can still be environmentally screened by nearby massive haloes.

### 5.2.2 Halo concentration

In Chapter 3, we used the rescaled logarithmic mass  $x = \log_{10}(M_{500}/10^{p_2})$  to calibrate the following formula for the enhancement of the halo concentration  $c_{200}$ :

$$\log_{10} \left| \frac{c}{c_{\text{GR}}} \right|_{200} = \frac{1}{2} \left( \frac{\lambda}{\omega_s} \phi(x') \left[ 1 + \operatorname{erf} \left( \frac{\alpha x'}{\sqrt{2}} \right) \right] + \gamma \right) \times (1 - \tanh(\omega_t [x + \xi_t])), \quad (5.3)$$

where  $x' = (x - \xi_s)/\omega_s$ ,  $\phi(x')$  is the normal distribution and  $\operatorname{erf}(\alpha x'/\sqrt{2})$  is the error function. The parameters have best-fit values  $\lambda = 0.55$ ,  $\omega_s = 1.7$ ,  $\xi_s = -0.27$ ,  $\alpha = -6.5$ ,  $\gamma = -0.07$ ,  $\omega_t = 1.3$  and  $\xi_t = 0.1$ .

### 5.2.3 Observable-mass scaling relations

In Chapter 4, we used the SHYBONE simulations to verify a set of proposed mappings between the  $f(R)$  scaling relations and their GR power-law counterparts. Here, we will only recap the ‘true density’ rescaling for the  $Y_{\text{SZ}}(M)$  relation, which is used in our constraint pipeline. For haloes in  $f(R)$  gravity and GR that have the same true mass,  $M_{\text{true}}^{f(R)} = M^{\text{GR}}$ , we showed:

$$Y_{\text{SZ}}^{f(R)}(M_{\text{true}}^{f(R)}) \approx \frac{M_{\text{dyn}}^{f(R)}}{M_{\text{true}}^{f(R)}} Y_{\text{SZ}}^{\text{GR}}(M^{\text{GR}} = M_{\text{true}}^{f(R)}). \quad (5.4)$$

In this case, the total gravitational potential of the  $f(R)$  haloes is enhanced by a factor of  $M_{\text{dyn}}^{f(R)}/M_{\text{true}}^{f(R)}$  compared to the GR haloes. The temperature is then enhanced by the same amount, giving rise to this factor in Eq. (5.4). We showed that this mapping holds for halo masses  $M_{500} \gtrsim 10^{13.5} M_{\odot}$ .

### 5.2.4 Halo mass function

In this section, we will outline the [Cataneo et al. \(2016\)](#) model for the  $f(R)$  enhancement of the HMF, which we have adopted for our constraint pipeline. This is computed using the [Sheth and Tormen \(1999\)](#) prescription of the HMF:

$$n_{\text{ST}} \equiv \frac{dn}{d \ln M} = \frac{\bar{\rho}_M}{M} \frac{d \ln \nu}{d \ln M} \nu f(\nu), \quad (5.5)$$

where the multiplicity function  $\nu f(\nu)$  is given by:

$$\nu f(\nu) = A \sqrt{\frac{2}{\pi} a \nu^2} \left[ 1 + (a \nu^2)^{-p} \exp\left(-\frac{a \nu^2}{2}\right) \right]. \quad (5.6)$$

For the parameters  $A$ ,  $a$  and  $p$ , [Cataneo et al. \(2016\)](#) used the fits by [Despali et al. \(2016\)](#), which extend the [Sheth and Tormen \(1999\)](#) HMF to be a function of generic halo overdensity  $\Delta$ . For the latter, [Cataneo et al. \(2016\)](#) used value  $300\Omega_M(z)$  (i.e., here the halo mass  $M$  is  $M_{300m}$ ). The peak height  $\nu$  is given by:

$$\nu = \frac{\delta_c}{\sigma(M, z)}, \quad (5.7)$$

where  $\delta_c$  is the linearly extrapolated threshold density for spherical collapse and  $\sigma(M, z)$  is the linear root-mean-square fluctuation of the matter density within spheres of mass  $M$  containing an average density of  $\bar{\rho}_M(z)$ . The latter can be computed using the  $\Lambda$ CDM linear power spectrum (for both GR and  $f(R)$  gravity) with the publicly available code CAMB ([Lewis et al., 2000](#)).

The  $f(R)$  effects are incorporated through  $\delta_c$ : in GR, this is given by:

$$\delta_c^{\text{GR}}(z) \approx \frac{3}{20} (12\pi)^{\frac{2}{3}} [1 + 0.0123 \log_{10} \Omega_M(z)], \quad (5.8)$$

while in  $f(R)$  gravity it can be expressed as:

$$\delta_c^{\text{eff}}(M, z) \equiv \epsilon(M, z) \times \delta_c^{f(R)}(M, z). \quad (5.9)$$

The function  $\delta_c^{f(R)}(M, z)$  is the prediction of the linearly extrapolated threshold density for spherical collapse in  $f(R)$  gravity. This treats haloes and their surrounding environment as co-centred spherically symmetric top-hat overdensities (note the environment can be underdensities) which are co-evolved from an initial time to the time of collapse. This procedure, which is based on the method developed by [Li and Efstathiou \(2012\)](#); [Lombriser et al. \(2013\)](#), takes into account both the mass-dependent self-screening and the environmental screening of the fifth force. However, while giving qualitatively correct predictions, the method is unable to very accurately capture the complex nonlinear dynamics of structure formation in  $f(R)$  gravity. This limitation is accounted for using the correction factor  $\epsilon(M, z)$ , which [Cataneo et al. \(2016\)](#) modelled and fitted using DMO simulations. Their

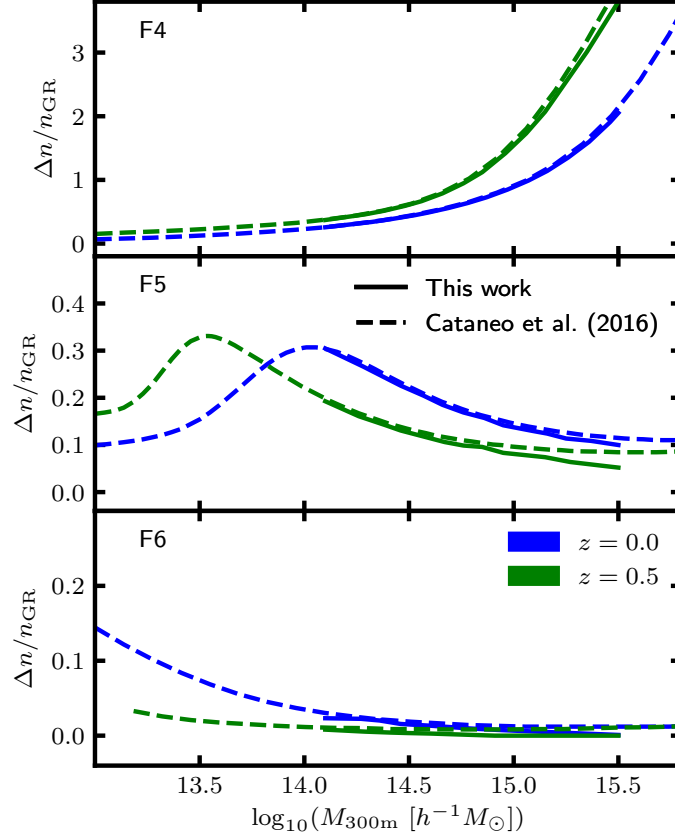


Figure 5.2: Halo mass function enhancement in  $f(R)$  gravity with respect to GR as a function of the halo mass. The solid lines show the predictions from our 5D interpolation of  $\delta_c^{f(R)}$  (see Sec. 5.2.4) and the dashed lines show the results from Cataneo et al. (2016). The predictions have been generated using the WMAP9 cosmological parameters and  $f(R)$  models F4 (top row), F5 (middle row) and F6 (bottom row), at redshifts 0 (blue lines) and 0.5 (green lines).

best-fit model can accurately reproduce the  $f(R)$  enhancement of the HMF for redshifts  $0.0 \leq z \leq 0.5$  and field strengths  $10^{-6} \leq |f_{R0}| \leq 10^{-4}$ .

For this thesis, we have evaluated  $\delta_c^{f(R)}$  on a grid of  $M$ ,  $z$ ,  $\Omega_M$ ,  $\sigma_8$  and  $f_{R0}$ , and obtained the relation  $\delta_c^{f(R)}(M, z, \Omega_M, \sigma_8, f_{R0})$  using 5D interpolation. For a given set of cosmological and  $f(R)$  parameters, we can use this to predict  $\delta_c^{f(R)}(M, z)$ , which can then be used to predict  $\delta_c^{\text{eff}}(M, z)$  using the model for  $\epsilon(M, z)$  taken from Cataneo et al. (2016). The  $f(R)$  enhancement of the HMF is given by the ratio between  $n_{\text{ST}}|_{f(R)}$  and  $n_{\text{ST}}|_{\text{GR}}$ , which are evaluated using  $\delta_c = \delta_c^{\text{eff}}$  and  $\delta_c = \delta_c^{\text{GR}}$ , respectively.

For illustrative purposes, we show, in Fig. 5.2, our predictions of the HMF enhancement as a function of the halo mass for F6, F5 and F4 at redshifts 0.0 and 0.5. We also show the

predictions from [Cataneo et al. \(2016\)](#) as a comparison. Both sets of predictions assume the 9-year WMAP cosmological parameter estimates ([Hinshaw et al., 2013](#)). There are some small differences between the two sets of predictions, which are likely caused by subtle differences in the calculations of  $\delta_c^{f(R)}$ . The largest difference is observed at  $M_{300\text{m}} \gtrsim 10^{15} h^{-1} M_\odot$  for F5 at  $z = 0.5$ . We note that the enhancement is expected to drop to zero at high masses where haloes become completely screened, therefore the behaviour of the solid lines here appears to be physically reasonable. We also note that we set the enhancement to zero wherever our calculations predict a negative (unphysical) enhancement. This is the case for  $M_{300\text{m}} \gtrsim 10^{15} h^{-1} M_\odot$  for F6 at  $z = 0.5$ .

### 5.3 Methods

In this section, we describe the main components of our constraint pipeline, including the mass function predictions (Sec. 5.3.1), the observable-mass scaling relation (Sec. 5.3.2), the mock cluster catalogues (Sec. 5.3.3) and the MCMC sampling (Sec. 5.3.4).

#### 5.3.1 Theoretical mass function

In order to make constraints using cluster number counts, it is necessary to have a parameter-dependent theoretical model for the HMF. For our pipeline, we start with a GR HMF and apply the  $f(R)$  enhancement using:

$$n^{f(R)} = n^{\text{GR}} \times \frac{n_{\text{ST}}|_{f(R)}}{n_{\text{ST}}|_{\text{GR}}}, \quad (5.10)$$

where the ratio is computed using the [Sheth and Tormen \(1999\)](#) prescription, as described in Sec. 5.2.4, and we have chosen the [Tinker et al. \(2008\)](#) calibration for  $n^{\text{GR}}$ .

Before Eq. (5.10) can be applied, the halo mass definition must be considered. As mentioned in Sec. 5.2.4, the model for the ratio in Eq. (5.10) was calibrated by [Cataneo et al. \(2016\)](#) using overdensity  $\Delta = 300\Omega_{\text{M}}(z)$ ; however, with the framework in Fig. 2.1, we hope to use data from SZ and X-ray surveys, which often measure cluster properties with overdensity 500. Therefore, it is necessary to convert the HMF between these two definitions.



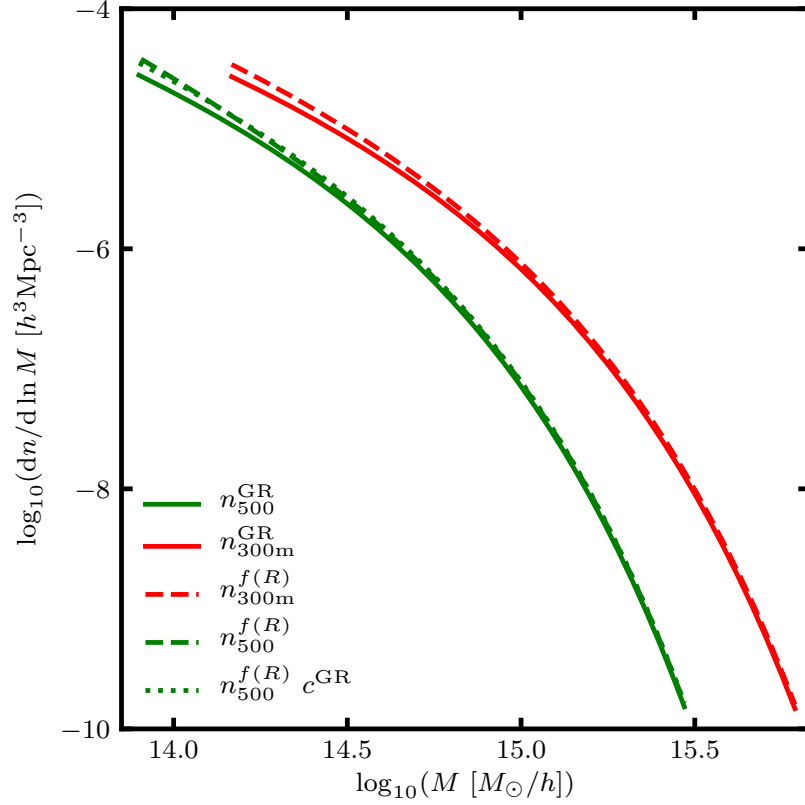


Figure 5.3: Halo mass function in GR (*solid lines*) and F5 (*dashed lines*), with the mass defined using spherical overdensities 500 (*green lines*) and  $300\Omega_M(z)$  (*dashed lines*). The mass conversions and  $f(R)$  enhancement have been applied as described in Sec. 5.3.1; the green dotted line shows the F5 HMF prediction that results from neglecting the  $f(R)$  enhancement of the halo concentration in the mass conversion  $300\Omega_M(z) \rightarrow 500$ .

In Fig. 5.3, we show each step of the mass conversion procedure for the F5 model at  $z = 0$ . We start with the Tinker et al. (2008) HMF with overdensity 500 ( $n_{500}^{\text{GR}}$ ), which we compute using the python package HMF (Murray et al., 2013), and convert this to overdensity  $300\Omega_M(z)$  ( $n_{300m}^{\text{GR}}$ ) using the Duffy et al. (2008) concentration-mass-redshift relation. We then apply the  $f(R)$  enhancement using Eq. (5.10) to get  $n_{300m}^{f(R)}$ . Finally, to convert this back to overdensity 500 ( $n_{500}^{f(R)}$ ), we use the  $f(R)$  concentration-mass-redshift relation, which is computed by applying the concentration enhancement, given by Eq. (5.3), to the Duffy et al. (2008) relation. We also show, with the dotted green line, the prediction with the concentration enhancement neglected; the effect here is quite small, since cluster-size haloes are mostly screened in F5. For further details of the formulae used to convert the halo mass and the HMF from one mass definition to another, we refer the reader to

Appendix B.2.

The final result  $n_{500}^{f(R)}(M_{500})$  provides the theoretical prediction of the cluster abundance in  $f(R)$  gravity. This is computed following the above steps for each set of parameter values sampled by our MCMC pipeline. We note that our mass conversions are evaluated assuming an NFW profile, which has also been used in previous cluster tests of  $f(R)$  gravity (e.g., Cataneo et al., 2015). However, this may not provide an accurate description for haloes that are not dynamically relaxed and it does not account for the effects of baryons on the total mass profile. We plan to investigate the latter effect using clusters identified from the realistic full hydrodynamical simulations in  $f(R)$  gravity described in Chapter 8. However, we remark here that the main use of the concentration-mass relation in our pipeline is to perform mass conversions as described above, and so it would not be strictly needed if a theoretical HMF for the required mass definition  $M_{\Delta}$  ( $M_{500}$  in this case) is already in place.

### 5.3.2 Observable-mass scaling relation

As discussed in Sec. 5.2.3, the  $f(R)$  scaling relation can be computed by simply rescaling a GR relation using our model for the dynamical mass enhancement. For the GR relation, we adopt the power-law mapping between  $Y_{\text{SZ}}$  and the halo mass calibrated by the Planck Collaboration (Ade et al., 2016):

$$E^{-\beta}(z) \left[ \frac{D_A^2(z) \bar{Y}_{500}}{10^{-4} \text{Mpc}^2} \right] = Y_{\star} \left[ \frac{h}{0.7} \right]^{-2+\alpha} \left[ \frac{(1-b)M_{500}}{6 \times 10^{14} M_{\odot}} \right]^{\alpha}, \quad (5.11)$$

where  $E(z) = H(z)/H_0$  and  $D_A(z)$  is the angular diameter distance. This includes parameters  $\beta$  for the  $z$ -evolution,  $Y_{\star}$  for the normalisation and  $\alpha$  for the power-law slope with respect to the mass. It also includes a bias parameter  $(1-b)$  which accounts for differences between the X-ray determined masses used in the calibration, which are subject to hydrostatic equilibrium bias, and the true mass. Planck have also provided the following formula for the intrinsic lognormal scatter of the relation:

$$P(\log Y_{500}) = \frac{1}{\sqrt{2\pi}\sigma_{\log Y}} \exp \left[ -\frac{\log^2(Y_{500}/\bar{Y}_{500})}{2\sigma_{\log Y}^2} \right], \quad (5.12)$$

where  $\sigma_{\log Y}$  is a fixed spread.

We assume a fixed value of 0.8 for the hydrostatic equilibrium bias parameter, which is consistent with the range 0.7 to 1.0 adopted by Planck, and we treat  $Y \equiv D_A^2(z)Y_{\text{SZ}}$  as the cluster SZ observable, rather than  $Y_{\text{SZ}}$ . This leaves four scaling relation parameters which are allowed to vary in our MCMC sampling. We adopt the following Gaussian priors from Planck:  $\log Y_* = -0.19 \pm 0.02$ ,  $\alpha = 1.79 \pm 0.08$ ,  $\beta = 0.66 \pm 0.50$  and  $\sigma_{\log Y} = 0.075 \pm 0.010$ .

To obtain the  $f(R)$  scaling relation  $Y_{\text{SZ}}^{f(R)}(M_{500})$  from the above  $Y_{\text{SZ}}^{\text{GR}}(M_{500})$  relation, we rescale the right-hand side of Eq. (5.11) by the mass ratio  $\mathcal{R}$ , which is predicted using Eq. (5.1) with scatter given by Eq. (B.1). This rescaling is based on Eq. (5.4), which means that the mass  $M_{500}$  in the expressions  $Y_{\text{SZ}}^{f(R)}(M_{500})$  and  $Y_{\text{SZ}}^{\text{GR}}(M_{500})$  above is the true mass; we note that, although the Planck masses were originally determined using X-ray measurements, the value  $(1 - b) = 0.8$  assumed for the mass bias is consistent with weak lensing measurements (e.g., [Hoekstra et al., 2015](#)).

Finally, we note that the scaling relation adopted in this work is intended to be representative of general scaling relations between the mass and SZ and X-ray observables, not just the Planck  $Y_{\text{SZ}}(M_{500})$  relation. This justifies our decision to encapsulate  $D_A^2(z)$  in the cluster observable and to fix the hydrostatic equilibrium bias; indeed, scaling relations for other observables — for example, the SZ significance and the  $Y_X$  parameter — do not include the function  $D_A^2(z)$  or a bias parameter (e.g., [de Haan et al., 2016](#); [Bocquet et al., 2019](#)). Regardless of the observable, the main purpose of this chapter is to check that our constraint pipeline can give reasonable constraints of  $f_{R0}$  using a realistic scaling relation which includes both intrinsic scatter and the  $f(R)$  enhancement. It would be very straightforward to adapt this pipeline for other cluster observables, or for more than one cluster observable.

### 5.3.3 Mock catalogues

We test our framework (Fig. 2.1) using mock cluster catalogues in place of observational data. We have generated mocks for both the GR and F5 models, using fiducial cosmological parameter values based on the Planck 2018 CMB constraints ([Aghanim et al., 2020](#)):  $(\Omega_M, \sigma_8, h, \Omega_b, n_s) = (0.3153, 0.8111, 0.6736, 0.04931, 0.9649)$ . For the scaling relation para-

meters, we assume the central values of the Gaussian priors listed in Sec. 5.3.2.

To generate the mocks, we first compute the predicted count per unit mass per unit redshift:

$$\frac{dN}{dz d \ln M} = \frac{dn}{d \ln M} \times \frac{dV_c(z)}{dz}, \quad (5.13)$$

where  $V_c(z)$  is the comoving volume enclosed by the survey area between redshifts 0 and  $z$  and the first term is the theoretical HMF  $n_{500}^{f(R)}$ , which is computed as described in Sec. 5.3.1 for the fiducial cosmology. For this work, we assume a survey area of  $5000 \text{ deg}^2$  and a maximum redshift of  $z = 0.5$ , which is the upper redshift used to calibrate the  $f(R)$  enhancement of the HMF (Sec. 5.2.4). In the future, it will be important to develop models of the  $f(R)$  HMF that work for a wider redshift range, which will be applicable to real cluster survey data.

The predicted number of clusters is:

$$N_{\text{tot}} = \int_{0.0}^{0.5} dz \int_{-\infty}^{\infty} d \ln M \frac{dN}{dz d \ln M}. \quad (5.14)$$

For each mock, we randomly draw the masses and redshifts of  $N_{\text{tot}}$  clusters using  $dN/dz d \ln M$ , which is effectively a probability density. For each cluster  $i$ , we then draw a mass ratio  $\mathcal{R}_i$  using a normal distribution with mean given by Eq. (5.1) and standard deviation given by Eq. (B.1). The intrinsic observable  $Y'_i (= D_A^2 Y_{500,i})$  of each cluster is then drawn using the lognormal distribution given by Eq. (5.12), where  $\bar{Y}_{500}$  is computed using Eq. (5.11) and rescaled by  $\mathcal{R}_i$ .

We assume a fixed  $1\sigma$  measurement uncertainty of 10%. The measured observable  $Y_i$  is therefore drawn from a normal distribution with mean  $Y'_i$  and standard deviation  $0.1Y'_i$ . We note that this choice of a fixed fractional uncertainty is intended to keep our calculations simple and general (for example, a more complicated model may be specific to a particular observational survey). We have also considered 5% and 20% uncertainties and have found that the inferred parameter constraints do not significantly differ, suggesting that this uncertainty is not the dominant source of error in the constraint pipeline (e.g., compared to the intrinsic scatters in the cluster scaling relation or the  $f(R)$  dynamical mass enhancement).

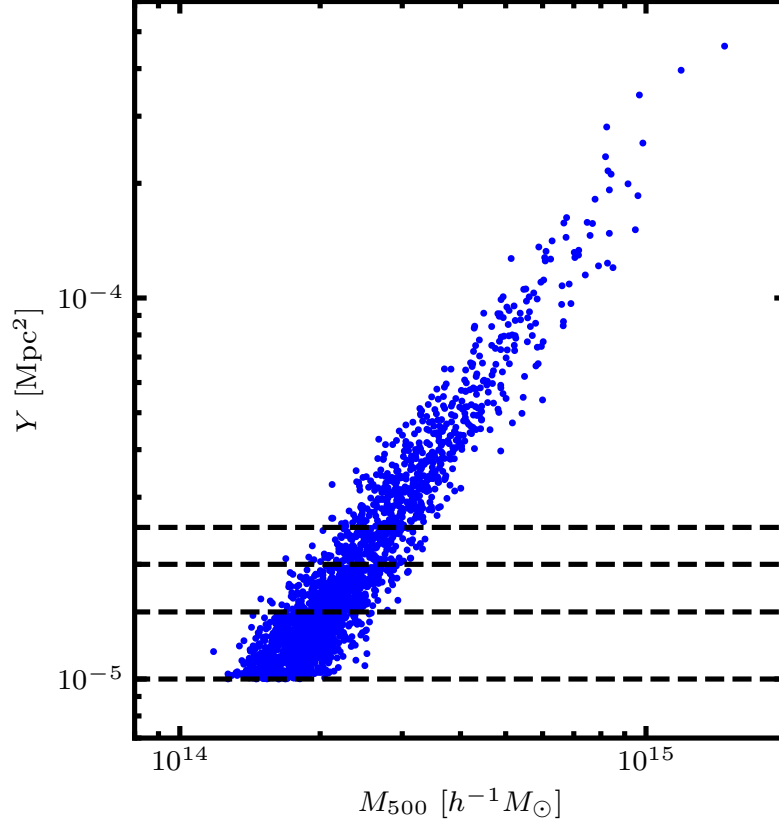


Figure 5.4: SZ  $Y$ -parameter as a function of the halo mass for clusters from a GR mock catalogue with observational flux limit  $Y_{\text{cut}} = 10^{-5} \text{Mpc}^2$ . The dashed lines indicate the cuts  $10^{-5} \text{Mpc}^2$ ,  $1.5 \times 10^{-5} \text{Mpc}^2$ ,  $2 \times 10^{-5} \text{Mpc}^2$  and  $2.5 \times 10^{-5} \text{Mpc}^2$ .

Finally, we remove all clusters for which  $Y_i$  is below some observational flux limit  $Y_{\text{cut}}$ . For the main results of this chapter, we use  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ ; however, we will also discuss the effects of using cuts  $10^{-5} \text{Mpc}^2$ ,  $2 \times 10^{-5} \text{Mpc}^2$  and  $2.5 \times 10^{-5} \text{Mpc}^2$ . For each mock, we store only the cluster redshift  $z_i$  (which is assumed to have no error) and the measured observable  $Y_i$ .

An F5 mock with  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$  contains  $\sim 1350$  clusters. Generating a GR mock is more straightforward, since there is no need to include the  $f(R)$  enhancements of the HMF or the scaling relation. In this case, there are  $\sim 1150$  clusters for  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ . For illustrative purposes, in Fig. 5.4 we show the measured  $Y$ -parameters of the clusters as a function of the mass  $M_{500}$  for a GR mock with  $Y_{\text{cut}} = 10^{-5} \text{Mpc}^2$ . Horizontal dashed lines are included to indicate the four flux thresholds considered in this work, to give an idea of the mass range of clusters found above each.

### 5.3.4 MCMC sampling

For our parameter constraints, we use the unbinned Poisson likelihood (e.g., [Artis et al., 2021](#)):

$$\ln \mathcal{L} = - \int dz dY \frac{dN}{dz dY}(z, Y) + \sum_i \ln \frac{dN}{dz dY}(z_i, Y_i), \quad (5.15)$$

where the first term represents the predicted cluster count and the second term is a summation performed over all mock clusters. The expression  $dN/dz dY$  represents the theoretical prediction of the count per unit  $z$  per unit  $Y$ .

Since our theoretical HMF is defined in terms of the mass  $M$  rather than the observable  $Y$ , it is more convenient to re-express the first term with an integral over  $\ln M$  (e.g., [de Haan et al., 2016](#)):

$$\begin{aligned} & - \int_{0.0}^{0.5} dz \int_{Y_{\text{cut}}}^{\infty} dY \frac{dN}{dz dY}(z, Y) \\ & = - \int_{0.0}^{0.5} dz \int_{-\infty}^{\infty} d \ln M P(Y > Y_{\text{cut}} | M, z) \frac{dN}{dz d \ln M}(M, z), \end{aligned} \quad (5.16)$$

where  $dN/dz d \ln M$  can be computed using Eq. (5.13), and the redshift integral is evaluated between  $z = 0$  and the maximum redshift  $z = 0.5$  of the mock.  $P(Y > Y_{\text{cut}} | M, z)$  represents the probability that, for a given mass and redshift, the measured  $Y$ -parameter exceeds the flux threshold. This depends on both the measurement uncertainty and the intrinsic log-normal scatter of  $Y$ :

$$P(Y > Y_{\text{cut}} | M, z) = \int_{-\infty}^{\infty} d \ln Y' P(Y > Y_{\text{cut}} | Y') P(Y' | M, z), \quad (5.17)$$

where  $P(Y > Y_{\text{cut}} | Y')$  is the probability that the measured value  $Y$  exceeds  $Y_{\text{cut}}$ , given an intrinsic value  $Y'$ , and  $P(Y' | M, z)$  is the probability density of a cluster having intrinsic value  $Y'$  given that it has mass  $M$  and redshift  $z$ . As discussed in Sec. 5.3.3, the mocks use a fixed measurement uncertainty of 10%, which means that the former can be estimated using a normal distribution with mean  $Y'$  and standard deviation  $0.1Y'$ . The probability density  $P(Y' | M, z)$  is more complicated, since this depends both on the intrinsic scatter of the  $Y(M)$  scaling relation and the scatter of the mass ratio  $\mathcal{R}$ :

$$P(Y' | M, z) = \int_1^{4/3} d\mathcal{R} P(Y' | \bar{Y}(M, z, \mathcal{R})) P(\mathcal{R} | M, z), \quad (5.18)$$

where  $P(\mathcal{R}|M, z)$  is the probability density of a cluster having mass ratio  $\mathcal{R}$  given that it has mass  $M$  and redshift  $z$ . This is computed using a normal distribution with mean given by Eq. (5.1) and standard deviation given by Eq. (B.1). The other probability density,  $P(Y'|\bar{Y}(M, z, \mathcal{R}))$ , is computed using Eq. (5.12), with  $\bar{Y}$  calculated using Eq. (5.11) and rescaled by a factor of  $\mathcal{R}$ . Together, Eqs. (5.16)-(5.18) form a 4D integral, which we compute using a fixed grid in  $(\ln M, z, \ln Y, \mathcal{R})$ .

For the second term in Eq. (5.15), we can again re-express into a form that depends on  $dN/(dz d \ln M)$  using:

$$\begin{aligned} \frac{dN}{dz dY}(z_i, Y_i) &= \int d \ln Y' \int d \ln M' \\ &\times P(Y_i|Y') P(Y'|M', z_i) \frac{dN}{dz d \ln M'}(M', z_i), \end{aligned} \quad (5.19)$$

where the probability density functions  $P(Y_j|Y')$  and  $P(Y'|M', z_i)$  represent the measurement uncertainty and intrinsic scatter, respectively. The latter is computed using Eq. (5.18), meaning that Eq. (5.19) is really a 3D integral. We compute this for each mock cluster using a fixed grid in  $(\ln Y', \ln M', \mathcal{R})$ , then evaluate the sum in Eq. (5.15).

We have used the python package EMCEE (Foreman-Mackey et al., 2013) for the MCMC sampling. For all of the results discussed in this chapter, we have used 28 walkers each travelling 2700 steps (we discard the first 600 steps to ensure that the chains are well converged). At each step, the log-likelihood is computed for the sampled parameters as described above. In addition to the  $f_{R0}$  parameter, the cosmological parameters  $\Omega_M$  and  $\sigma_8$  and the four scaling relation parameters  $Y_*$ ,  $\alpha$ ,  $\beta$  and  $\sigma_{\log Y}$  are sampled. For the cosmological parameters, we adopt uniform (flat) priors  $\log_{10} |f_{R0}| \in [-7, -4]$  and  $\sigma_8 \in [0.60, 0.95]$ , and for  $\Omega_M$  we use either a flat prior  $\Omega_M \in [0.15, 0.50]$  or a Gaussian prior  $\Omega_M = 0.3153 \pm 0.0073$  which is based on the Planck 2018 CMB constraints (Aghanim et al., 2020). For the scaling relation parameters, we adopt the Gaussian priors listed in Sec. 5.3.2.

The flat prior  $[-7, -4]$  for  $\log_{10} |f_{R0}|$  extends beyond the range  $[-6, -4]$  used to calibrate the HMF enhancement model (Cataneo et al., 2016). For sampled values in the range  $-7 \leq \log_{10} |f_{R0}| \leq -6$ , we first calculate the HMF enhancement for  $\log_{10} |f_{R0}| = -6$ , then linearly interpolate between  $|f_{R0}| = 0$  (GR) and  $|f_{R0}| = 10^{-6}$  to estimate the enhancement. For example, this means that the estimated enhancement for  $|f_{R0}| = 10^{-7}$  would be 10%

of the enhancement for  $|f_{R0}| = 10^{-6}$ . We note that, because clusters are expected to be completely screened for this range of  $\log_{10} |f_{R0}|$  values, it is not necessary to use a physically accurate method here, so long as the predicted enhancement lies between GR and F6. We use a similar approach to estimate the dynamical mass enhancement for this range of  $\log_{10} |f_{R0}|$ , where, again, the enhancement is very close to zero anyway.

## 5.4 Results

In this section, we discuss the main results of this chapter. In Sec. 5.4.1, we use a GR mock to check that our pipeline can give reasonable constraints of the  $\Lambda$ CDM and scaling relation parameters. Then, in Sec. 5.4.2, we use our full pipeline to constrain the  $f_{R0}$  parameter of  $f(R)$  gravity, using a combination of GR and F5 mocks.

### 5.4.1 GR pipeline

In order to verify that our pipeline can give reasonable  $\Lambda$ CDM constraints and successfully account for the intrinsic scatter of the  $Y(M_{500})$  relation and measurement uncertainty in the mock, we first test our ‘GR pipeline’, where the  $f(R)$  corrections to the HMF and scaling relation are excluded. We show the constraints, which have been inferred using a GR mock with  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ , in Fig. 5.5. The blue contours are obtained using the flat prior  $\Omega_{\text{M}} \in [0.15, 0.50]$ , while the red contours are obtained using the Gaussian prior  $\Omega_{\text{M}} = 0.3153 \pm 0.0073$  from Planck 2018.

For the flat  $\Omega_{\text{M}}$  prior, the contours are in good agreement with the fiducial parameter values, which are indicated by the green lines. In the top panel of each column, we show the marginalised distributions of each parameter, with the mean and standard deviation quoted in Table 5.1. In Table 5.1, we also show the combination of parameters that gave the highest log-likelihood during the sampling ( $\mathcal{L}_{\text{max}}$ ); these can be thought of as the ‘most likely’ set of values. The distributions of the scaling relation parameters closely match the Gaussian priors. Meanwhile, the constraints  $0.34 \pm 0.04$  for  $\Omega_{\text{M}}$  and  $0.79 \pm 0.04$  for  $\sigma_8$  – while still within  $1\sigma$  agreement – are slightly offset from the fiducial values, and the same goes for the highest-likelihood values 0.34 and 0.79. As shown by the constraints in red, using a tighter



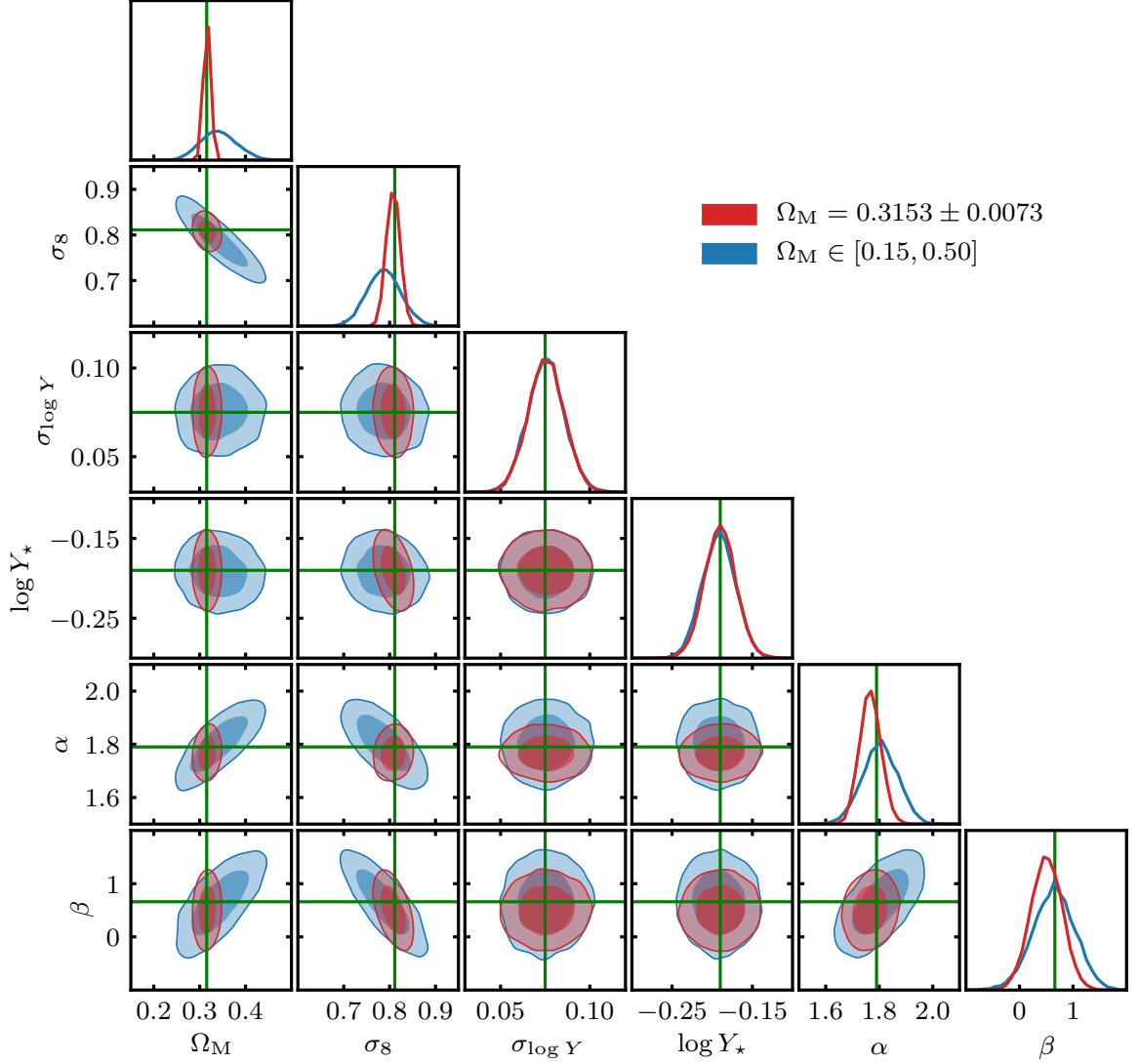


Figure 5.5: Parameter constraints using our GR pipeline, which does not include  $f(R)$  enhancements of the HMF and the scaling relation (see Sec. 5.3), using a GR mock with observational flux threshold  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ . The two sets of constraints are generated using a flat prior  $[0.15, 0.50]$  (blue) and a Gaussian prior  $0.3153 \pm 0.0073$  (red) in  $\Omega_M$ . The dark and light regions of the contours represent 68% and 95% confidences, respectively. The distributions of the sampled parameter values are shown in the top panels of each column, with the mean and standard deviation of each parameter quoted in Table 5.1. The fiducial cosmological parameter values of the GR mock are indicated by the green lines.

Parameter	Fiducial value	Prior	Flat $\Omega_M$ prior 68% range	Gaussian $\Omega_M$ prior 68% range	$\Omega_M$ prior $\mathcal{L}_{\max}$
$\Omega_M$	0.3153	—	$0.34 \pm 0.04$	$0.316 \pm 0.007$	0.3175
$\sigma_8$	0.8111	$[0.60, 0.95]$	$0.79 \pm 0.04$	$0.808 \pm 0.015$	0.8037
$\sigma_{\log Y}$	0.075	$0.075 \pm 0.010$	$0.076 \pm 0.010$	$0.076 \pm 0.010$	0.073
$\log Y_*$	-0.19	$-0.19 \pm 0.02$	$-0.19 \pm 0.02$	$-0.19 \pm 0.02$	-0.19
$\alpha$	1.79	$1.79 \pm 0.08$	$1.80 \pm 0.07$	$1.77 \pm 0.04$	1.76
$\beta$	0.66	$0.66 \pm 0.50$	$0.6 \pm 0.4$	$0.5 \pm 0.3$	0.58

Table 5.1: Parameter constraints using our GR pipeline. The mean and standard deviation are quoted (68% range) along with the parameter combinations giving the highest log-likelihood ( $\mathcal{L}_{\max}$ ). The constraints correspond to the distributions shown in Fig. 5.5.

Gaussian prior in  $\Omega_M$  results in narrower contours and constraints  $\Omega_M = 0.316 \pm 0.007$  and  $\sigma_8 = 0.808 \pm 0.015$  which match the fiducial values more closely.

The initial offset of the  $\Omega_M$  and  $\sigma_8$  constraints from the fiducial values is caused by a well-known degeneracy between these two parameters: increasing either of these will boost the predicted amplitude of the HMF. Therefore, the effects of increasing (decreasing)  $\Omega_M$  and decreasing (increasing)  $\sigma_8$  on the HMF can roughly cancel out. This causes the elongated shape of the blue  $\Omega_M$ - $\sigma_8$  contour.

We also observe degeneracies between  $\Omega_M$ ,  $\sigma_8$ ,  $\alpha$  and  $\beta$ . One explanation for this is that  $\alpha$  and  $\beta$  can also affect the predicted HMF. For example, increasing  $\alpha$  (i.e., increasing the slope of the  $Y(M)$  scaling relation) will cause the predicted  $Y$ -parameter to be reduced for clusters with  $0.8M_{500} < 6 \times 10^{14} M_\odot$  (since  $(1-b)^{-1} 6 \times 10^{14} M_\odot$  is the pivot mass of the power-law function in Eq. (5.11)), which includes the majority of clusters in our mocks (see Fig. 5.4). This means that fewer clusters will be predicted to have  $Y > Y_{\text{cut}}$ , and therefore the inferred cluster count will be lower, which can be countered by a larger  $\Omega_M$ . The effects of changing  $\alpha$ ,  $\beta$ ,  $\Omega_M$  and  $\sigma_8$  may balance out overall, giving rise to the observed degeneracies in the blue contours of Fig. 5.5.

By adopting the tighter  $\Omega_M$  prior, these degeneracies appear to be mostly eliminated. This shows the importance of accurate independent measurements of  $\Omega_M$  in the use of galaxy cluster number counts to constrain cosmological models and parameters.

### 5.4.2 $f(R)$ pipeline

We now test the full  $f(R)$  gravity constraint pipeline, which includes the  $f(R)$  effects on the HMF and scaling relation, as described in Secs. 5.3.1 and 5.3.2. In Fig. 5.6, we show the constraints inferred using GR and F5 mocks with  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ . For these results, we use the Gaussian prior of  $\Omega_M$  in order to prevent the  $\Omega_M$ - $\sigma_8$  degeneracy observed in Fig. 5.5. As we will show in Sec. 5.5, using a flat prior for  $\Omega_M$  can otherwise lead to biased constraints of  $\log_{10} |f_{R0}|$ .

For the constraints obtained from the GR mock, which are indicated by the red contours in Fig. 5.6, the  $\log_{10} |f_{R0}|$  posterior distribution is roughly uniform for the range

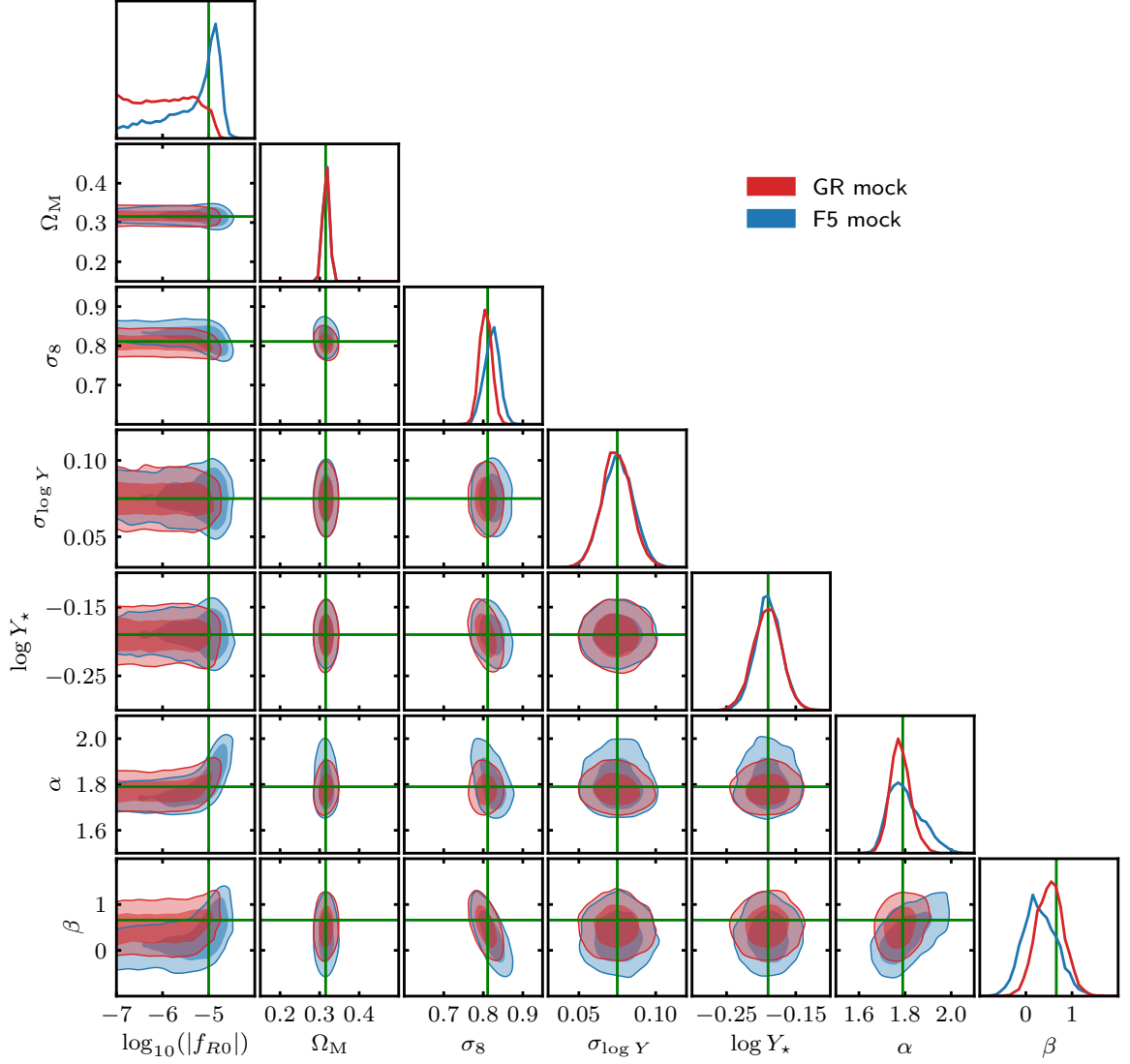


Figure 5.6: Parameter constraints using our full  $f(R)$  pipeline, as detailed in Sec. 5.3, using GR (*red*) and F5 (*blue*) mocks with observational flux threshold  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ . The dark and light regions of the contours represent 68% and 95% confidences, respectively. The distributions of the sampled parameter values are shown in the top panels of each column, with the mean and standard deviation quoted in Table 5.2. The fiducial cosmological parameter values of the mocks are indicated by the green lines, including the value  $\log_{10} |f_{R0}| = -5$  for the F5 mock.

Parameter	Fiducial value	Prior	GR constraints		F5 constraints	
			68% range	$\mathcal{L}_{\max}$	68% range	$\mathcal{L}_{\max}$
$\log_{10} f_{R0} $	—	$[-7, -4]$	$\leq -5.56$	-6.75	$-5.1^{+0.3}_{-1.0}$	-4.92
$\Omega_M$	0.3153	$0.3153 \pm 0.0073$	$0.316 \pm 0.007$	0.317	$0.316 \pm 0.008$	0.313
$\sigma_8$	0.8111	$[0.60, 0.95]$	$0.806 \pm 0.015$	0.806	$0.821 \pm 0.019$	0.815
$\sigma_{\log Y}$	0.075	$0.075 \pm 0.010$	$0.075 \pm 0.010$	0.072	$0.075 \pm 0.010$	0.079
$\log Y_*$	-0.19	$-0.19 \pm 0.02$	$-0.19 \pm 0.02$	-0.19	$-0.190 \pm 0.019$	-0.18
$\alpha$	1.79	$1.79 \pm 0.08$	$1.78 \pm 0.04$	1.77	$1.80 \pm 0.07$	1.82
$\beta$	0.66	$0.66 \pm 0.50$	$0.5 \pm 0.3$	0.51	$0.3 \pm 0.4$	0.44

Table 5.2: Parameter constraints using our full  $f(R)$  pipeline. The 68% range columns show the mean and standard deviation for all parameters other than  $\log_{10}|f_{R0}|$ ; for the latter, the 68% upper bound is shown for the GR mock constraints and the median and 68-percentile is shown the F5 mock constraints. The parameter combinations giving the highest log-likelihood ( $\mathcal{L}_{\max}$ ) are also shown. The constraints correspond to the distributions shown in Fig. 5.6.

$-7 \leq \log_{10} |f_{R0}| \lesssim -5$  and drops to zero for  $\log_{10} |f_{R0}| \gtrsim -5$ . This rules out  $f(R)$  models stronger than F5, whereas weaker models are difficult to distinguish from GR for this sample of clusters. We show our constraints of the parameter values in Table 5.2. Since the  $\log_{10} |f_{R0}|$  posterior does not follow a normal distribution, we quote an upper bound rather than the mean and standard deviation. In this case, 68% of the sampled points have  $\log_{10} |f_{R0}| \leq -5.56$ . We note that this threshold may depend on the width of the  $\log_{10} |f_{R0}|$  prior: for a wider prior (i.e., extending the lower bound of the prior to some value smaller than  $-7$  while fixing the upper bound of the prior) and a uniform  $\log_{10} |f_{R0}|$  posterior, it is reasonable to expect the 68% upper bound to be lower. Therefore, it is perhaps more useful to look at the combination of parameter values that give the highest log-likelihood. In this case, the most likely combination has  $\log_{10} |f_{R0}| = -6.75$ , which is quite close to the lower bound of the prior (although we note that, given the flat posterior distribution of  $\log_{10} |f_{R0}|$ , the point with  $\mathcal{L}_{\max}$  might not be much more significant than points with only slightly smaller log-likelihood values). The constraints for the other parameters are in excellent agreement with the fiducial values. Therefore, the results suggest that our pipeline can successfully constrain  $f_{R0}$  using cluster samples in a GR universe.

The constraints for the F5 mock are indicated by the blue contours in Fig. 5.6. The  $\log_{10} |f_{R0}|$  constraints appear to be in good agreement with the fiducial value  $-5$ , which lies within the 68% confidence region of the contours. This region only extends down to  $\log_{10} |f_{R0}| \approx -6.5$ , clearly favouring  $f(R)$  gravity over GR. The constraints also appear to rule out models with  $\log_{10} |f_{R0}| \gtrsim -4.5$ . The median and 68-percentile range of the sampled values is  $\log |f_{R0}| = -5.1^{+0.3}_{-1.0}$ , while the highest-likelihood parameter combination has  $\log_{10} |f_{R0}| = -4.92$ . Both of these results are very close to the fiducial value of  $-5$ . The constraints for the other parameters are again in very reasonable agreement with the fiducial values. This result suggests that our pipeline can clearly identify if the underlying universe model is F5.

Despite this promising agreement, it is interesting to note that the  $\log_{10} |f_{R0}|$  posterior distribution has a long tail over the range  $-7 < \log_{10} |f_{R0}| < -5$ . Over this range of points,  $\sigma_8$  appears to have value 0.83-0.84 on average, while  $\alpha$  and  $\beta$  have values  $\sim 1.75$  and  $\sim 0.0$

on average (see the blue contours in Fig. 5.6). As  $\log_{10} |f_{R0}|$  is lowered, the predicted amplitude of the HMF will be reduced. The increased  $\sigma_8$  can act against this, as can the lowered  $\alpha$ , which, as discussed in Sec. 5.4.1, can increase the predicted cluster count for clusters with  $0.8M_{500} < 6 \times 10^{14} M_{\odot}$ . The latter can also give a scaling relation that more closely matches the F5 result: this is because the scaling relation in F5 is enhanced at lower masses, which may be approximated by the constraint pipeline as a power-law with shallower slope. This degeneracy also comes into play for  $\log_{10} |f_{R0}| > -5$ , where  $\sigma_8$  becomes slightly lower on average and  $\alpha$  and  $\beta$  become higher. Overall, this reduces the precision of the  $\log_{10} |f_{R0}|$  constraint, and is perhaps the reason why the  $\log_{10} |f_{R0}|$  posterior peaks at a value that is slightly higher than  $-5$ . This can also explain why the  $\beta$  constraints predict a value  $0.3 \pm 0.4$  that is slightly lower than the fiducial value 0.66. By using tighter priors in  $\sigma_8$ ,  $\alpha$  or  $\beta$  it may be possible to eliminate this bias (see Sec. 5.5.3 for a detailed discussion).

We have also tested our pipeline using an F4.5 mock (with  $\log_{10} |f_{R0}| = -4.5$ ), and in Appendix B.3 we show that this model is clearly distinguished from F5.

## 5.5 Potential biases in model constraints

In Sec. 5.4, we demonstrated that our framework can give very reasonable constraints of  $\log_{10} |f_{R0}|$  for both GR and F5 mocks (Fig. 5.6). An important feature of this constraint framework (Fig. 2.1) is the inclusion of corrections for the effects of  $f(R)$  gravity on the internal cluster properties, which are expected to prevent biased constraints. In Sec. 5.5.1, we will assess potential sources of bias in the constraint pipeline, including an incomplete treatment of the scaling relation. Then, in Sec. 5.5.2, we will check the effects of the cluster sample, including selection criteria, on the constraints. Finally, we will demonstrate how the various parameter degeneracies can be prevented by using tighter parameter priors in Sec. 5.5.3.

For all of the figures in this section, we will only show constraints for parameters that are either biased or contribute to parameter degeneracies. Therefore, we exclude the  $\log Y_{\star}$  and  $\sigma_{\log Y}$  constraints, since these always match the Gaussian priors very closely (e.g., see

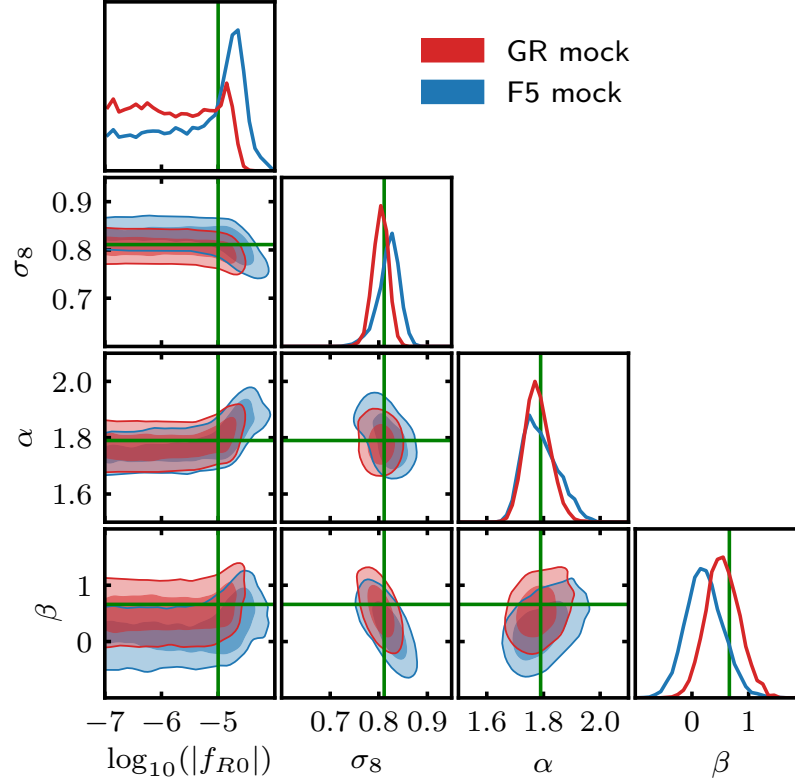


Figure 5.7: Parameter constraints generated using the same GR (*red*) and F5 (*blue*) mocks as Fig. 5.6, but with a simplified constraint pipeline in which the  $f(R)$  effects on the observable-mass scaling relation are neglected.

Figs. 5.5 and 5.6). For similar reasons, we will also exclude  $\Omega_M$  constraints that have been inferred using the Gaussian prior from Planck 2018.

### 5.5.1 Constraint pipeline

#### 5.5.1.1 Power-law scaling relation

In Fig. 5.7, we show constraints inferred using the same GR and F5 mocks as used for Fig. 5.6. However, here the  $f(R)$  effects on the SZ scaling relation (Eq. (5.11)) have been neglected, i.e., a power-law scaling relation without  $f(R)$  corrections is used in the (incomplete)  $f(R)$  pipeline.

For the GR mock constraints, shown by the red contours in Fig. 5.7, the  $\log_{10} |f_{R0}|$  posterior appears to be uniformly distributed over the range  $-7 \leq \log_{10} |f_{R0}| \lesssim -4.5$ . This extends beyond the range  $-7 \leq \log_{10} |f_{R0}| \lesssim -5$  observed using the full pipeline



in Fig. 5.6, and the range  $\log_{10} |f_{R0}| \leq -5.36$  containing 68% of the sampled points has a higher upper bound than the range  $\leq -5.56$  given in Table 5.2 for the full pipeline. Therefore, even though the GR mock is generated using a power-law scaling relation, it seems that neglecting the  $f(R)$  effects on the scaling relation in the pipeline leads to less precise and weaker constraints of  $\log_{10} |f_{R0}|$  overall.

The F5 mock constraints, which are shown by the blue contours, still give a peaked  $\log_{10} |f_{R0}|$  posterior distribution. However, there are now a greater proportion of sampled points within the range  $-7 \leq \log_{10} |f_{R0}| \lesssim -5$ . This means that the 68% confidence contours extend to  $\log_{10} |f_{R0}| = -7$ , indicating that the pipeline is unable to convincingly rule out GR. There are also a greater number of sampled points with  $\log_{10} |f_{R0}| \gtrsim -4.5$ ; indeed, the highest-likelihood parameter combination has  $\log_{10} |f_{R0}| = -4.56$ , which is much higher than the fiducial value  $-5$  and the value  $-4.92$  when using the full pipeline. The median and 68-percentile range is  $\log_{10} |f_{R0}| = -5.1^{+0.5}_{-1.2}$ , which is less precise than the constraint  $\log_{10} |f_{R0}| = -5.1^{+0.3}_{-1.0}$  with the full  $f(R)$  pipeline.

In summary, our constraints for the GR and F5 mocks indicate that assuming a power-law observable-mass scaling relation can lead to imprecise and biased constraints of  $f(R)$  gravity. This appears to be linked to parameter degeneracies, where we again observe a lowered  $\sigma_8$  and increased  $\alpha$  for  $\log_{10} |f_{R0}| \gtrsim -5$ , and an increased  $\sigma_8$  and lowered  $\alpha$  and  $\beta$  at  $\log_{10} |f_{R0}| \lesssim -5$ .

### 5.5.1.2 *Mass ratio scatter*

For our constraints using the F5 mock in Fig. 5.6, we included the scatter of the dynamical mass enhancement, given by Eq. (B.1), in both the mock and the log-likelihood calculation. We now consider the effect of neglecting this scatter from the mock and the likelihood. The new result is shown by the red contours in Fig. 5.8, along with the previous results in blue. Without this scatter, the observable-mass scaling relation is less scattered overall; as a result, the  $f(R)$  constraints are more precise, with 68-percentile range  $\log_{10} |f_{R0}| = -4.89^{+0.15}_{-0.35}$  as opposed to  $\log_{10} |f_{R0}| = -5.1^{+0.3}_{-1.0}$ . In particular, the red 68% contours do not feature the tail towards low  $\log_{10} |f_{R0}|$ . These results indicate that excluding the scatter

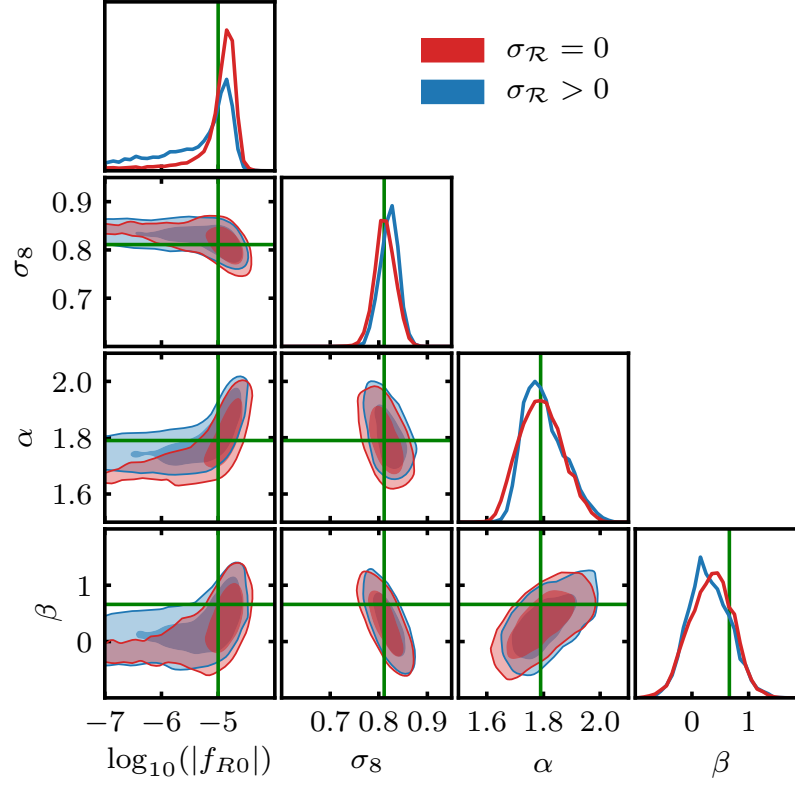


Figure 5.8: Parameter constraints generated using our constraint pipeline, where the blue constraints are the same as the F5 mock constraints in Fig. 5.6 and the red constraints are generated with the scatter of the dynamical mass enhancement set to zero in both the mock and log-likelihood.

could lead to  $f(R)$  constraints with an unrealistically high precision.

## 5.5.2 Cluster sample

### 5.5.2.1 Flux threshold

In addition to the observational cut  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$  which is used in the main results of this chapter, we have also generated mocks with cuts  $10^{-5} \text{Mpc}^2$ ,  $2 \times 10^{-5} \text{Mpc}^2$  and  $2.5 \times 10^{-5} \text{Mpc}^2$ . From Fig. 5.4, a cut of  $10^{-5} \text{Mpc}^2$  means that the lowest mass clusters, with  $M_{500} \sim 10^{14} h^{-1} M_{\odot}$ , are included in the sample. In the F5 model, the HMF is more enhanced at these lower halo masses (see Fig. 5.2), therefore it is expected that using lower-mass objects can give more precise constraints of  $\log_{10} |f_{R0}|$ .

In Fig. 5.9, we show constraints generated from F5 mocks with these three cuts. For  $Y_{\text{cut}} = 2.5 \times 10^{-5} \text{Mpc}^2$ , the sampled  $\log_{10} |f_{R0}|$  distribution is quite uniform for  $-7 <$

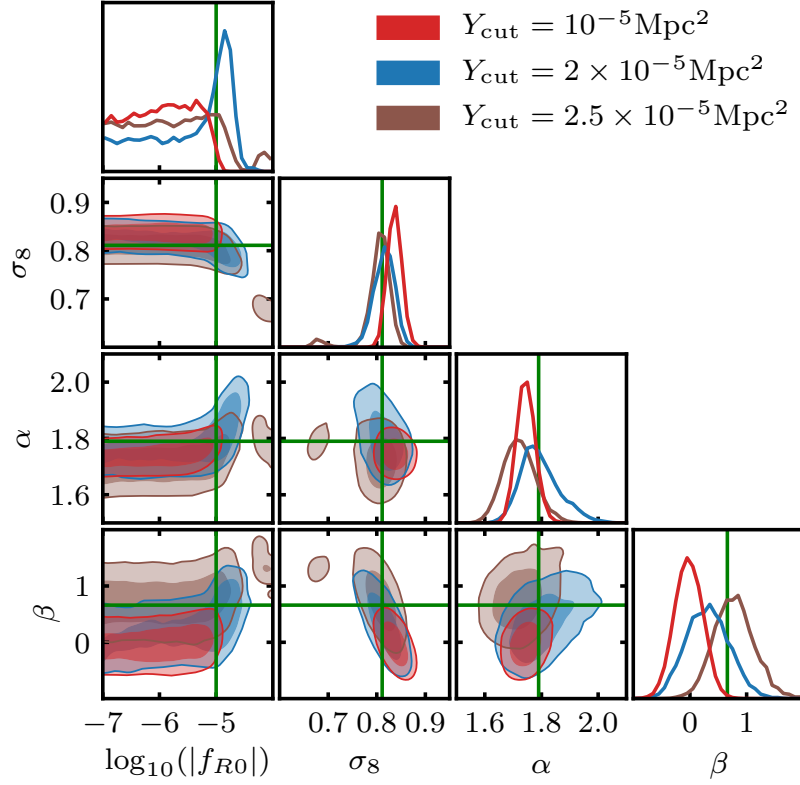


Figure 5.9: Parameter constraints generated using our constraint pipeline, using F5 mocks with observational flux thresholds of  $10^{-5}\text{Mpc}^2$  (red),  $2 \times 10^{-5}\text{Mpc}^2$  (blue) and  $2.5 \times 10^{-5}\text{Mpc}^2$  (brown).

$\log_{10} |f_{R0}| \lesssim -5$ , indicating that this high-mass cluster sample cannot be used to distinguish the F5 model from weaker models, including GR. This is not surprising, given that higher-mass clusters are better-screened in F5, which means that their number count deviates from the GR prediction less strongly (see Fig. 5.2). On the other hand, the constraints for  $Y_{\text{cut}} = 2 \times 10^{-5}\text{Mpc}^2$  clearly favour  $\log_{10} |f_{R0}|$  values close to  $-5$ . However, the 68% contours still extend to  $\log_{10} |f_{R0}| = -7$ , which is very close to GR. This is improved upon with  $Y_{\text{cut}} = 1.5 \times 10^{-5}\text{Mpc}^2$ , which is able to convincingly distinguish the F5 model from GR, as we showed in Fig. 5.6.

In Fig. 5.9, we also show the constraints from the F5 mock with  $Y_{\text{cut}} = 10^{-5}\text{Mpc}^2$ . Interestingly, despite containing lower-mass clusters than the other mocks, the sampled  $\log_{10} |f_{R0}|$  values are approximately evenly distributed over  $-7 \lesssim \log_{10} |f_{R0}| \lesssim -5$ . One possible reason is that this mock catalogue includes many more low-mass, unscreened, clusters, and the main constraining power comes from different objects than the previous

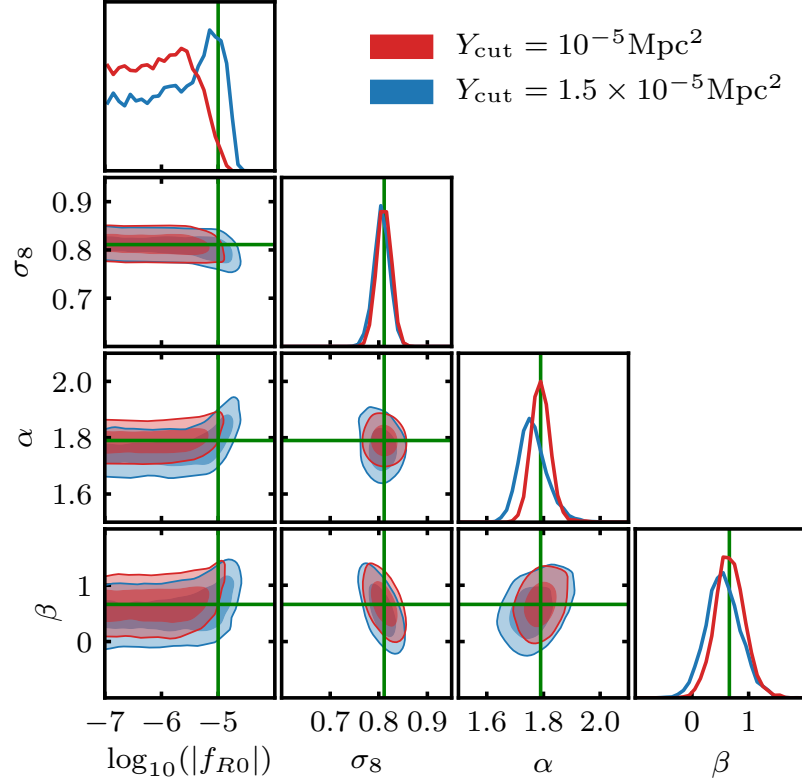


Figure 5.10: Parameter constraints generated using our constraint pipeline, using GR mocks with observational flux thresholds  $10^{-5}\text{Mpc}^2$  (red) and  $1.5 \times 10^{-5}\text{Mpc}^2$  (blue). The latter is a different realisation (generated in the same way) from the GR mock used in Fig. 5.6, and is included to show the potential effects of sample variance on the constraints.

cases. We note that for these constraints, the  $\sigma_8$ ,  $\alpha$  and  $\beta$  parameters are all biased. As we have already discussed, these parameters can be varied in such a way that the predicted theoretical HMF in GR (i.e., with low  $\log_{10} |f_{R0}|$ ) can match the F5 HMF with the fiducial cosmological parameters. Our results here show that this can cause biased constraints which appear to prefer GR over  $f(R)$  gravity even though this is an F5 mock, and this seems to be more relevant for cluster samples that extend to lower masses. As we will show in Sec. 5.5.3, these degeneracies can be prevented by using tighter parameter priors.

We note that the biased results described above only apply to an F5 fiducial cosmology. The red contours in Fig. 5.10 show the constraints inferred using a GR mock with  $Y_{\text{cut}} = 10^{-5}\text{Mpc}^2$ . These are consistent with GR, with 68% of the sampled points in the range  $\log_{10} |f_{R0}| \leq -5.71$ , which is even more precise than the  $\log_{10} \leq -5.56$  constraint from Fig. 5.6. Meanwhile, the constraints for  $\sigma_8$ ,  $\alpha$  and  $\beta$  show an excellent match with the

fiducial values. Therefore, the bias described above may not be an issue for cluster samples in a GR universe.

### 5.5.2.2 *Sample variance*

In order to check the effect of sample variance on the constraints, we have generated several GR mocks with  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ , following the method discussed in Sec. 5.3.3. In all cases, the inferred constraints of  $\log_{10} |f_{R0}|$  are consistent with GR, with the 68% constraint contours spanning  $-7 \leq \log_{10} |f_{R0}| \lesssim -5$  just like the red contours in Fig. 5.6.

However, we have occasionally observed peaks in the  $\log_{10} |f_{R0}|$  posterior distribution close to  $-5$ , which are related to the degeneracies between  $\log_{10} |f_{R0}|$ ,  $\sigma_8$ ,  $\alpha$  and  $\beta$  mentioned above. An example is shown with the blue contours in Fig. 5.10. As we have discussed, in the constraints using the F5 mock in Fig. 5.6, we can see a ‘rise’ in the  $\log_{10} |f_{R0}|$ - $\alpha$  contour at  $\log_{10} |f_{R0}| > -5$ ; there is a similar ‘rise’ in the case of the blue contours in Fig. 5.10. This is because a larger  $\alpha$ , which means a steeper scaling relation and hence underpredicted cluster number counts, could be compensated by a stronger gravity, so that to the pipeline, the GR mock would appear to be reasonably fitted with an  $f(R)$  model with slightly larger  $\alpha$ . We also see a slight ‘drop’ in the  $\log_{10} |f_{R0}|$ - $\sigma_8$  contour, where the lowered  $\sigma_8$  can again counteract the strengthened gravity. These effects can lead to more points sampled around  $\log_{10} |f_{R0}| = -5$ , and because even stronger gravity is disfavoured an artificial peak is formed at  $-5$ . While the peak in  $\log_{10} |f_{R0}|$  here is smaller than the peak observed for the F5 mock in Fig. 5.6, it is important to be wary that degeneracies can lead to a particular value of  $\log_{10} |f_{R0}|$  being favoured even for a GR fiducial cosmology. Like the other sources of bias discussed in this chapter, this issue can be eliminated by using tighter priors, as we will show in the next section.

### 5.5.3 Tighter priors

For many of the results discussed in this chapter, we have observed degeneracies between  $\log_{10} |f_{R0}|$ ,  $\Omega_M$ ,  $\sigma_8$ ,  $\alpha$  and  $\beta$ . Together, these parameters can vary such that the theoretical GR HMF is consistent with the F5 mocks, or similarly the theoretical F5 HMF can be

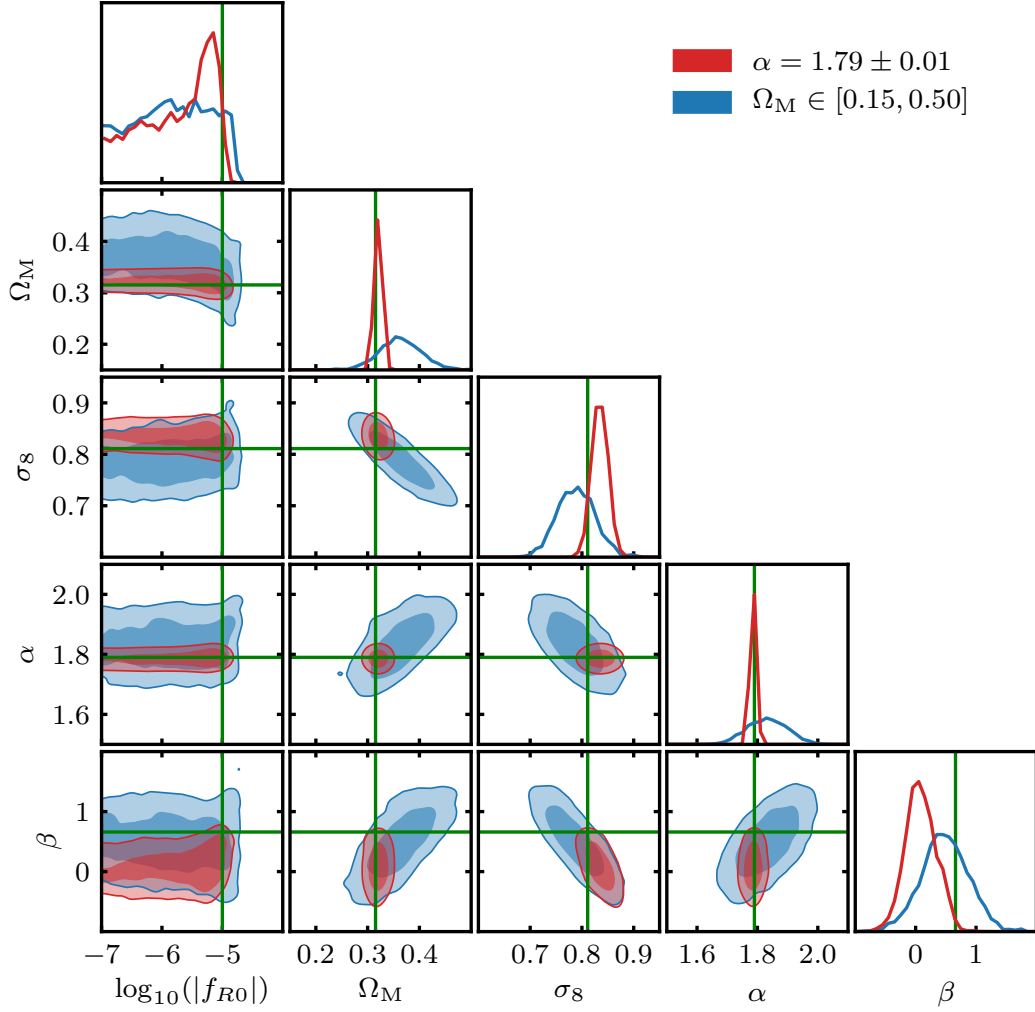


Figure 5.11: Parameter constraints generated by our constraint pipeline using: an F5 mock with flux threshold  $Y_{\text{cut}} = 10^{-5} \text{Mpc}^2$  and a tight Gaussian prior  $1.79 \pm 0.01$  for  $\alpha$  (red); and an F5 mock with flux threshold  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$  and a flat prior  $[0.15, 0.50]$  for  $\Omega_M$  (blue).

made consistent with the GR mocks. For our main results with the  $f(R)$  constraint pipeline (Fig. 5.6), we have been using a tight Gaussian  $\Omega_M$  prior from Planck 2018. In Fig. 5.11, the blue constraints have been generated using the same F5 mock as the blue constraints in Fig. 5.6; however, here a flat prior  $[0.15, 0.50]$  has been adopted for  $\Omega_M$ . This gives rise to the degeneracy between  $\Omega_M$  and  $\sigma_8$  (observed earlier in Fig. 5.5), which leads to a uniform distribution in  $\log_{10} |f_{R0}|$ . Fig. 5.12 provides an illustration of this degeneracy: here, the HMF prediction for F6 with increased  $\Omega_M$  and reduced  $\sigma_8$  closely resembles the F5 prediction, particularly at lower masses which dominate the mock cluster samples. By using

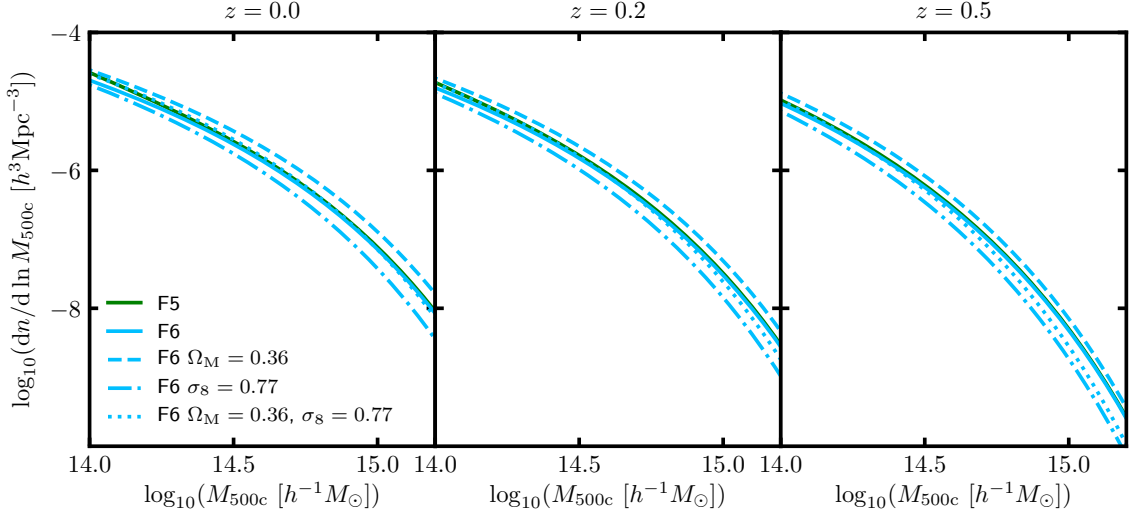


Figure 5.12: Predictions of the HMF in F6 (*light blue*) and F5 (*dark green*) at redshifts 0.0, 0.2 and 0.5. We show predictions with the fiducial parameter values  $\Omega_M = 0.3153$  and  $\sigma_8 = 0.8111$  (*solid lines*), an increased  $\Omega_M$  (*dashed line*), a reduced  $\sigma_8$  (*dash-dotted line*) and both an increased  $\Omega_M$  and reduced  $\sigma_8$  (*dotted line*). This figure illustrates not only the well-known degeneracy between  $\Omega_M$  and  $\sigma_8$  in determining the HMF, but also their degeneracy with  $f_{R0}$ : by tuning the values of these two parameters, an F6 model can closely mimic the HMF of an F5 model; note that the latter degeneracy may be broken by looking at multiple redshifts or by having more precise knowledge of  $\Omega_M$  and/or  $\sigma_8$ .

the tight  $\Omega_M = 0.3153 \pm 0.0073$  prior for our main results in Sec. 5.4.2, we have prevented this issue. The tight prior on  $\Omega_M$  can potentially be replaced by combining cluster number counts with other cosmological probes that are sensitive to  $\Omega_M$ , such as the CMB.

We have also shown that there is a degeneracy between  $\sigma_8$  and the SZ scaling relation parameters  $\alpha$  and  $\beta$ . Although we have used Gaussian priors for the latter, they can still vary enough to cause biased constraints. In Sec. 5.5.2.1, we found that this degeneracy caused the  $\log_{10}|f_{R0}|$  constraints using the F5 mock with  $Y_{\text{cut}} = 10^{-5} \text{Mpc}^2$  to resemble GR (see Fig. 5.9). In Fig. 5.11, the red contours show the  $\log_{10}|f_{R0}|$  constraints for the same mock, but this time using a tighter  $\alpha$  prior of  $1.79 \pm 0.01$ . The  $\log_{10}|f_{R0}|$  posterior distribution now peaks close to the fiducial value  $-5$ , though the constraints on  $\sigma_8$  and  $\beta$  are similarly biased as before. In this case, as in Sec. 5.5.2.1, the constrained  $\beta$  value is lower, which means less time evolution; because the time evolution is normalised at  $z = 0$ , this implies that, for a given cluster mass  $M_{500}$ , the measured observable  $Y$  at  $z > 0$  is smaller than the true value, and so fewer detectable clusters would be predicted.

This is compensated by a larger  $\sigma_8$  (actually a similar degeneracy can be observed in the GR case, see the  $\sigma_8$ - $\beta$  contour in Fig. 5.5), but one side effect is that smaller  $\log_{10} |f_{R0}|$  values are more likely to be allowed, leading to a uniform posterior distribution in Fig 5.9, which is alleviated in Fig. 5.11 with the tighter prior on  $\alpha$  but nevertheless not completely eliminated. Looking at the red contours in the left column of Fig. 5.11, we can see that at  $\log_{10} |f_{R0}| \approx -5$ ,  $\beta$  and  $\sigma_8$  both match their correct values, which suggests that if we can tighten the prior on either  $\sigma_8$  or  $\beta$ , the constraint on  $\log_{10} |f_{R0}|$  can be further improved.

Therefore, a conclusion from this discussion is that, with better knowledge of the scaling relation parameters, it is possible to reduce the effect of these degeneracies. However, we note that it may be difficult to constrain the scaling relation parameters with even greater precision. In this case, the degeneracies could be prevented by using a synergy with weak lensing data, which can estimate the cluster mass with higher precision. Even if this data is only available for a subset of the clusters, it can still be incorporated in the log-likelihood (e.g., Bocquet et al., 2019).

## 5.6 Summary, Discussion and Conclusions

In this chapter, we have combined all of our models for the effects of  $f(R)$  gravity on cluster properties into an MCMC pipeline for constraining the amplitude of the present-day background scalar field,  $|f_{R0}|$ . We have adopted the model from Cataneo et al. (2016) for the  $f(R)$  enhancement of the HMF, and used this, along with our model for the enhancement of the halo concentration, to produce a model-dependent prediction of the cluster number counts (Sec. 5.3.1). We have also used our model for the enhancement of the dynamical mass in  $f(R)$  gravity to convert a GR power-law observable-mass scaling relation, which is based on the Planck  $Y_{\text{SZ}}(M_{500})$  relation (Ade et al., 2016), into a form consistent with  $f(R)$  gravity, where the fifth force enhances the relation at sufficiently low masses (Sec. 5.3.2). These models are all incorporated in our log-likelihood (Sec. 5.3.4), which we have used to infer parameter constraints using a set of mock cluster catalogues (Sec. 5.3.3).

Using a combination of GR and F5 mocks, we have shown that our pipeline is able to give reasonable parameter constraints that are consistent with the fiducial cosmology



(Figs. 5.5 and 5.6). For the GR mock, the constraints conclusively rule out  $f(R)$  models with  $\log_{10} |f_{R0}| \gtrsim -5$  and favour values in the range  $-7 \leq \log_{10} |f_{R0}| \lesssim -5$  where  $-7$  is the lowest value considered by our MCMC sampling. Meanwhile, the constraints inferred using the F5 mock favour values close to the fiducial value of  $-5$ , with 68% range  $-5.1^{+0.3}_{-1.0}$  and a ‘most likely’ value of  $-4.92$ . We have also shown that the constraints inferred from both mocks can be imprecise and biased if the  $f(R)$  enhancement of the scaling relation is not accounted for (Fig. 5.7). Therefore, this should be properly modelled in future tests of  $f(R)$  gravity in order to prevent biased constraints. This will become particularly relevant as cluster catalogues start to enter the galaxy group regime (e.g., Pillepich et al., 2018a; Lovisari et al., 2021), where more objects can be unscreened in  $f(R)$  gravity.

Throughout this chapter, the main obstacle to precise and unbiased constraints has stemmed from degeneracies between  $f_{R0}$ ,  $\Omega_M$ ,  $\sigma_8$  and the scaling relation parameters  $\alpha$  and  $\beta$ , all of which can influence the predicted cluster count. We have shown that the degeneracies can be prevented by using a tighter Gaussian prior for  $\Omega_M$  and by having better knowledge of the scaling relation parameters (Fig. 5.11). The latter can potentially be achieved by including lensing data for a subset of the clusters. If wide or flat parameter priors are used, this may give rise to biased constraints of  $\log_{10} |f_{R0}|$ . For example, we have found that the parameter degeneracies can have a more significant effect for cluster samples that extend to lower masses (Sec. 5.5.2.1).

Our constraint pipeline can be improved in a couple of ways. First, while the HMF model of Cataneo et al. (2016) is accurate, it only covers the redshift range  $[0, 0.5]$ . An extended model that works for a larger redshift range, as well as for wider ranges of other cosmological parameters (not restricted to the  $\Omega_M$  and  $\sigma_8$  parameters as we have focused on here), would be very useful. Calibrating this model for spherical overdensity  $\Delta = 500$  would also mean that conversions between halo mass definitions would no longer be required. Second, the MCMC pipeline should be extended so that independent cluster data, such as weak lensing, can be included in the model constraint. Once these tasks are completed, this pipeline can be used to constrain  $f(R)$  gravity using observations. It is also straightforward to extend our framework to other gravity models; we have already started to do this for

the nDGP model (Chapter [6](#)).

---

# Cluster and halo properties in DGP gravity

## 6.1 Introduction

So far in this thesis, we have calibrated models for the effects of the  $f(R)$  gravity fifth force on the dynamical mass (Chapter 2), the halo concentration (Chapter 3) and the observable-mass scaling relations (Chapter 4) of clusters, and we have used these to create a robust MCMC constraint pipeline that uses cluster number counts to probe gravity (Chapter 5). However, our general framework for cluster tests of gravity (Fig. 1.2) is intended to be easily extended to other MG models. In this chapter, we will study cluster and halo properties in nDGP. As discussed in Chapter 1, this model gives rise to departures from GR above a particular ‘cross-over’ scale, resulting in a fifth force which enhances the total strength of gravity (at smaller scales, the fifth force is screened out by the Vainshtein screening mechanism). As with the fifth force in HS  $f(R)$  gravity, the fifth force in nDGP could alter cluster properties such as the temperature and density profile. If these are not taken into account in cluster constraints, then cluster mass measurements could become biased. We will address this by studying four models of nDGP, which exhibit different strengths of the fifth force, using a combination of DMO and full-physics simulations that cover a wide range of resolutions and box sizes. This allows us to study and model the effects of

the nDGP fifth force on the halo concentration and observable-mass scaling relations. By combining our DMO simulations, we also examine the halo abundance over a continuous mass range extending from Milky Way galaxy-sized to large cluster-sized haloes.

This chapter is organised as follows: in Sec. 6.2, we describe the nDGP simulations used in the analyses of this chapter and the method for calculating the halo properties; our main results are presented and discussed in Sec. 6.3; and, in Sec. 6.4, we give the main conclusions and discuss the significance of our results.

## 6.2 Simulations and methods

Since Eq. (1.22) for the scalar field is highly nonlinear, the fifth force in the nDGP model can display a wide spectrum of behaviours, depending on time, scale and the mass of the objects being considered. Therefore, numerical simulations are essential for predicting its cosmological properties and implications accurately. For earlier works that make use of nDGP simulations, see, e.g., [Chan and Scoccimarro \(2009\)](#); [Schmidt \(2009\)](#); [Khouri and Wyman \(2009\)](#); [Li et al. \(2013\)](#); [Falck et al. \(2014, 2015\)](#). We describe the DMO and full-physics simulations used in this chapter in Sec. 6.2.1. Then, in Sec. 6.2.2, we describe our methods for computing the thermal properties and concentration of our haloes.

### 6.2.1 Simulations

Our simulations were run using AREPO ([Springel, 2010](#)). One of these is the first cosmological simulation to simultaneously incorporate both full baryonic physics (implemented using the IllustrisTNG model) and nDGP\*. This simulation, which is part of the SHYBONE simulation suite (see [Arnold et al., 2019b](#); [Hernández-Aguayo et al., 2020](#)), has box size  $62h^{-1}\text{Mpc}$  and consists of  $512^3$  dark matter particles, with mass  $1.28 \times 10^8 h^{-1} M_\odot$ , and (initially) the same number of Voronoi gas cells, which have mass  $\sim 2.5 \times 10^7 h^{-1} M_\odot$  on average. We also have four DMO  $N$ -body simulations, whose specifications are provided

---

\*We note that the IllustrisTNG model was tuned using standard gravity simulations. However, the differences between the GR and nDGP predictions for the stellar and gas properties of galaxies are generally small compared to typical observational scatters (see, e.g., Fig. 8 of [Hernández-Aguayo et al., 2020](#)), making a full retuning of the TNG parameters for the nDGP model unnecessary.

Specifications and models	Simulations			
	L62	L200	L500	L1000
box size / $h^{-1}\text{Mpc}$	62	200	500	1000
particle number	$512^3$	$1024^3$	$1024^3$	$1024^3$
particle mass / $h^{-1}M_\odot$	$1.52 \times 10^8$	$6.39 \times 10^8$	$9.98 \times 10^9$	$7.98 \times 10^{10}$
nDGP models	N1,N5	N0.5,N1,N2,N5	N0.5,N1,N2,N5	N0.5,N1,N2,N5

Table 6.1: Specifications of the four DMO simulations that are used, along with the full-physics SHYBONE simulations, to study cluster and halo properties in nDGP. These are labelled L62, L200, L500 and L1000, according to their box size. The simulations have all been run for GR in addition to the nDGP models listed, where N0.5, N1, N2 and N5 correspond to  $H_0 r_c = 0.5, 1, 2, 5$ , respectively.

in Table 6.1. These have box sizes  $62h^{-1}\text{Mpc}$ ,  $200h^{-1}\text{Mpc}$ ,  $500h^{-1}\text{Mpc}$  and  $1000h^{-1}\text{Mpc}$ . Throughout this chapter, we will refer to these as L62, L200, L500 and L1000, respectively. These span a wide range of mass resolutions – from  $1.52 \times 10^8 h^{-1}M_\odot$  in L62 to  $7.98 \times 10^{10} h^{-1}M_\odot$  in L1000 – allowing us to study haloes spanning, continuously, the mass range  $\sim 10^{11} h^{-1}M_\odot$  to  $\sim 10^{15} h^{-1}M_\odot$ .

The simulations have all been run with cosmological parameters  $(h, \Omega_M, \Omega_b, \sigma_8, n_s) = (0.6774, 0.3089, 0.0486, 0.8159, 0.9667)$ . All simulations include runs with N5 and N1, in addition to GR. The L200, L500 and L1000 simulations also feature runs with N2 and N0.5, allowing us to thoroughly explore the effects of different strengths of the fifth force on halo properties. The simulations all begin at redshift  $z = 127$ . For this study, we use 12 particle snapshots from each simulation which span the redshift range  $0 \leq z \leq 3$ .

For completeness and as a first check, we show here the matter power spectra of the simulated nDGP models; because this is not the primary focus, we shall only discuss the result briefly. The upper panels of Fig. 6.1 show the matter power spectra generated using the  $z = 0$  and  $z = 1$  snapshots of L500 (similar results can be found for the L200 and L1000 boxes). The relative differences between the nDGP and GR spectra are shown in the lower panels, where we have also included the predictions from linear theory (dashed lines). On large scales ( $k \lesssim 0.1 h\text{Mpc}^{-1}$ ), the observed relative differences closely match the linear predictions; here, the fifth force enhances the power by  $\sim 25\%$  in N0.5 and by a few percent in N5 at  $z = 0$ . The enhancement is even greater at intermediate scales, where the

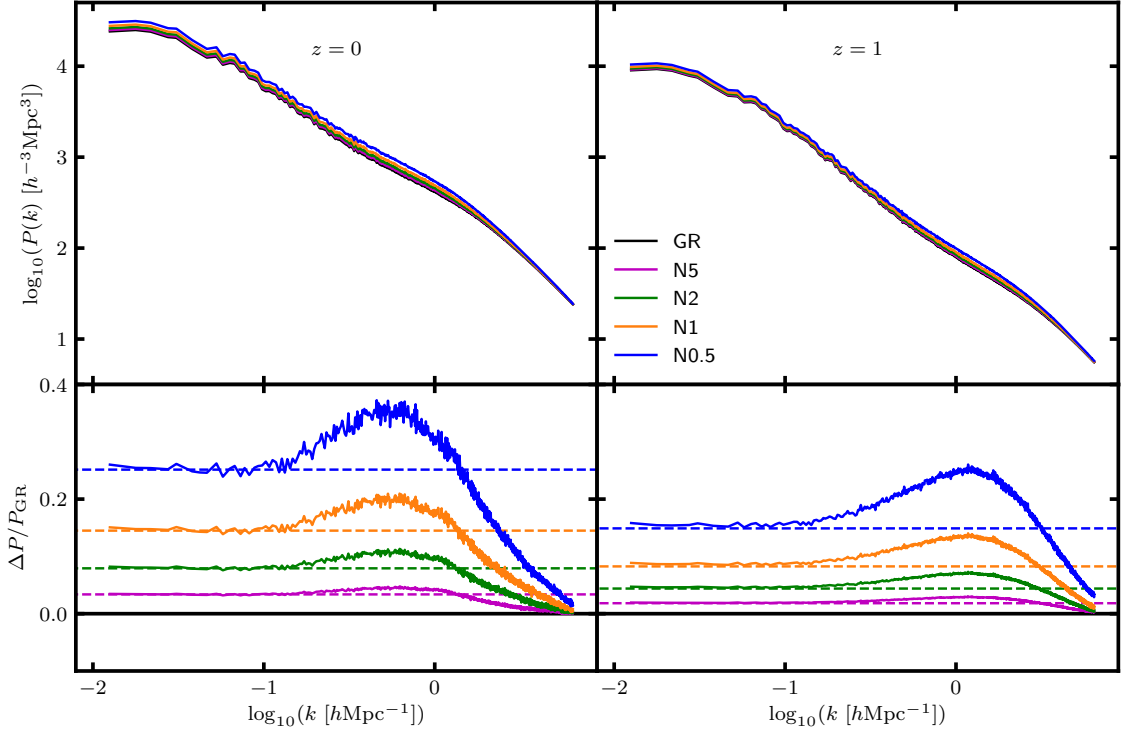


Figure 6.1: Matter power spectrum (*top row*) and its relative difference in nDGP with respect to GR (*bottom row*), as a function of the wavenumber at redshifts 0 and 1. The data has been generated using our dark-matter-only L500 simulation (see Table 6.1), which has been run for GR (*black*) and the nDGP models N5 (*magenta*), N2 (*green*), N1 (*orange*) and N0.5 (*blue*). The dashed lines in the bottom row show the linear theory predictions of the relative difference.

N0.5 power is enhanced by up to  $\sim 35\%$ . This is a consequence of mode-coupling at these scales. At smaller scales ( $k \gtrsim 1 h\text{Mpc}^{-1}$ ), which correspond to halo scales, the Vainshtein screening of the fifth force suppresses the power spectrum enhancement. These results are consistent with previous works (e.g., Schmidt, 2009; Schmidt et al., 2010; Winther et al., 2015). While the trends are similar at  $z = 0$  and  $z = 1$ , the nDGP enhancement is smaller for the latter due to the fifth force being weaker at earlier times.

### 6.2.2 Halo catalogues

At each simulation snapshot, we have generated halo catalogues using the SUBFIND code (Springel et al., 2001). For each halo, we have measured the mass-weighted gas temperature and the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  parameters using Eqs. (4.14)-(4.16), again excluding gas cells within the radial range  $r < 0.15R_{500}$ . The halo concentration is measured using full fitting of the

NFW profile to the halo density, as described in Sec. 3.2.2.1.

## 6.3 Results

In Sec. 6.3.1, we present our results for the observable-mass scalings using our full-physics simulations. Then, in Sec. 6.3.2, we study and model the concentration-mass-redshift relation in nDGP. Finally, in Sec. 6.3.3, we examine the HMF in nDGP.

### 6.3.1 Observable-mass scaling relations

In Figs. 6.2-6.4, we plot the mass-weighted gas temperature and the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  parameters against the halo mass  $M_{500}$ . In addition to showing individual data points for each halo in the mass range  $M_{500} > 10^{13} M_{\odot}$ , we also plot lines showing the median observable as a function of the mean logarithm of the mass. These averages have been computed using a moving window with a fixed size of 10 haloes. This approach, which is consistent with our study of the observable-mass scaling relations using the  $f(R)$  SHYBONE simulations (Chapter 4), is preferred over using a set of fixed-width bins, which would contain much fewer haloes at high mass than at low mass. The moving averages make use of all haloes with mass  $M_{500} > 10^{13} M_{\odot}$ , including cluster-sized haloes with  $M_{500} \gtrsim 10^{14} M_{\odot}$ . We note, however, that because there are only a few haloes with this mass (owing to the small box size of the full-physics simulations), the highest mean mass of the moving average is only  $\sim 10^{14} M_{\odot}$ . The lower panels of Figs. 6.2-6.4 show the relative differences between the observable medians in nDGP and GR. These are smoothed by computing the mean relative difference within 8 mass bins. We also show the root-mean-square halo scatter in GR for each of these bins (grey shaded regions).

The  $\bar{T}_{\text{gas}}-M$  relation is shown in the top two panels of Fig. 6.2. Both the GR and nDGP data follow a power-law relation as a function of the mass. From the lower panel of Fig. 6.2, we see that the median temperature in N5 agrees very closely with GR, typically within a couple of percent. This is consistent with the fact that the fifth force has a very small amplitude in this model (see the discussion below Eq. (1.23)). However, the temperature in N1 is enhanced by about 5% relative to GR on average. This result is quite surprising:

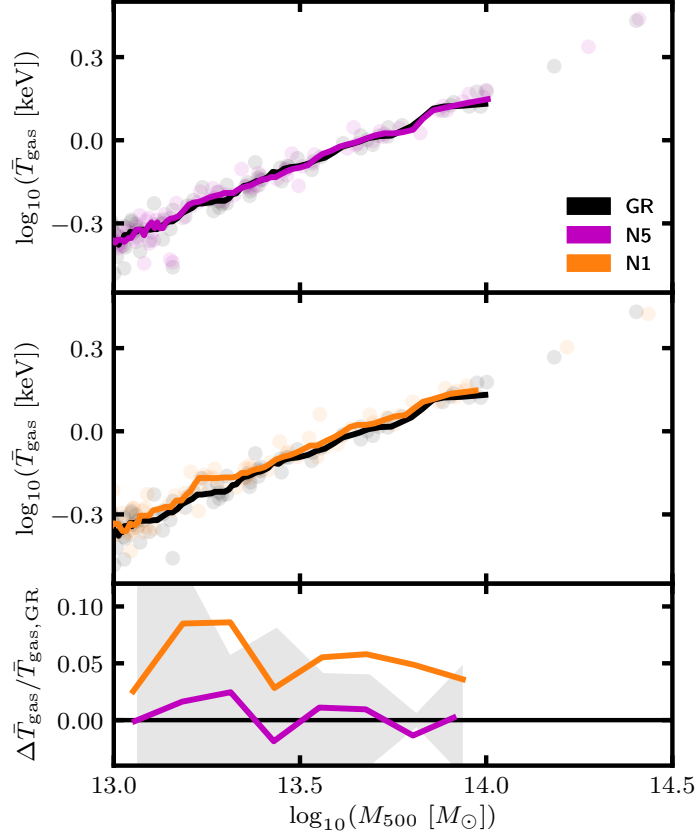


Figure 6.2: Gas temperature as a function of mass for haloes from our full-physics SHYBONE simulations (see Sec. 6.2.1) at  $z = 0$ . Data is included for GR (*black*) and the nDGP models N5 (*magenta*) and N1 (*orange*). The data points correspond to individual haloes. The lines show the median temperature and mean logarithm of the mass which have been computed using a moving window. *Bottom panel*: relative difference between the median temperatures in nDGP and GR; the grey shaded region shows the size of the GR halo scatter.

using the same full-physics simulations, [Hernández-Aguayo et al. \(2020\)](#) found that the N1 fifth force reaches just 2%-3% of the strength of the Newtonian force at the radius  $R_{500}$  for galaxy group-sized haloes and is even more efficiently screened at smaller radii. Therefore, the total gravitational potential at radius  $R_{500}$ , within which we have calculated the gas temperature, is expected to be just a few percent deeper than the Newtonian potential. We would therefore expect the temperature to be enhanced by just a few percent rather than the 5% that we observe. However, we note that, in nDGP, gravity is enhanced at the outer halo regions even at redshift  $z = 2$  (see, for example, Fig. 7 of [Hernández-Aguayo et al., 2020](#)). Therefore, between  $0 < z < 2$ , gas at the outer halo radii will undergo a gravitational acceleration in nDGP that is enhanced compared to GR. Consequently, it



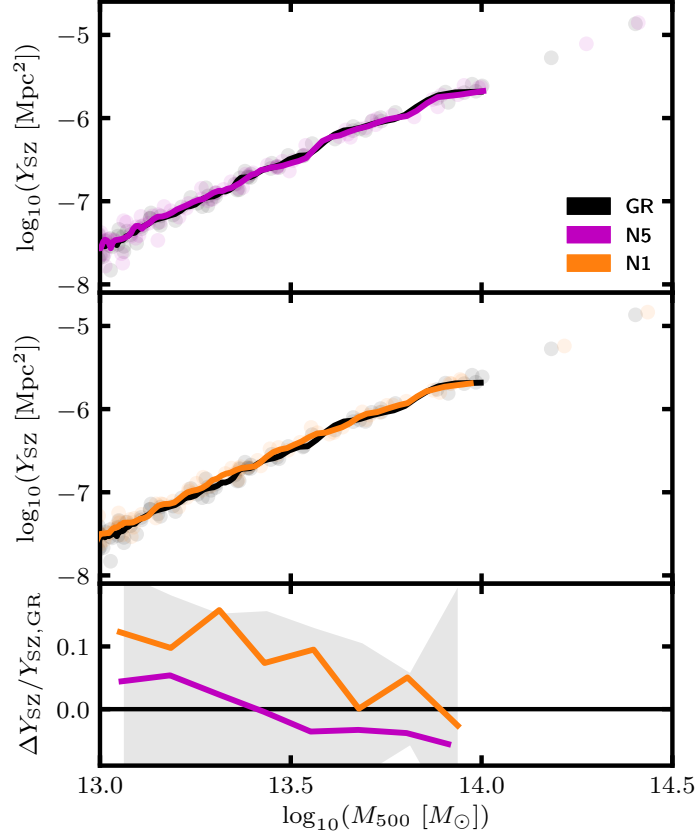


Figure 6.3: SZ Compton  $Y$ -parameter as a function of mass for haloes from our full-physics SHYBONE simulations (see Sec. 6.2.1) at  $z = 0$ . Data is included for GR (*black*) and the nDGP models N5 (*magenta*) and N1 (*orange*). The data points correspond to individual haloes. The lines show the median  $Y$ -parameter and mean logarithm of the mass which have been computed using a moving window. *Bottom panel*: relative difference between the median  $Y$ -parameters in nDGP and GR; the grey shaded region shows the size of the GR halo scatter.

will have a higher speed than in GR as it reaches smaller radii where it gets shock-heated. The fact that this happens over a long period of time can potentially explain how the gas temperature is enhanced by as much as 5% within  $R_{500}$ .

Our results for the  $Y_{\text{SZ}}-M$  and  $Y_{\text{X}}-M$  scaling relations are shown in Figs. 6.3 and 6.4, respectively. The  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  parameters are closely related to each other, and so the results appear similar for both: the enhancement of the  $Y$ -parameters in the N1 model ranges from zero at high masses to 10%-15% at low masses, while in N5 it ranges between a 5% suppression at high masses and 5% enhancement at low masses. The low-mass enhancement in N1 can in part be explained by the enhanced temperature seen in Fig. 6.2. Even for N5,

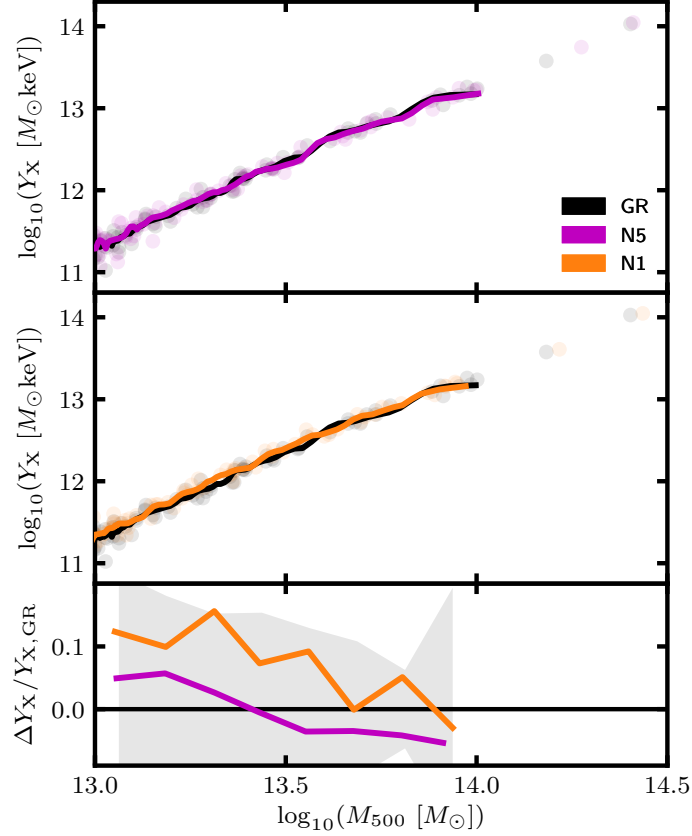


Figure 6.4: X-ray analogue of the Compton  $Y$ -parameter as a function of mass for haloes from our full-physics SHYBONE simulations (see Sec. 6.2.1) at  $z = 0$ . Data is included for GR (*black*) and the nDGP models N5 (*magenta*) and N1 (*orange*). The data points correspond to individual haloes. The lines show the median  $Y$ -parameter and mean logarithm of the mass which have been computed using a moving window. *Bottom panel*: relative difference between the median  $Y$ -parameters in nDGP and GR; the grey shaded region shows the size of the GR halo scatter.

the temperature appears to be enhanced on average for masses  $M_{500} \lesssim 10^{13.4} h^{-1} M_{\odot}$ , so this can also partly explain the  $\sim 5\%$  enhancement of the  $Y$ -parameters at these masses. The  $Y$ -parameters are also correlated with the gas density. In the top row of Fig. 6.5, we show the median gas density profiles for haloes from two mass bins (annotated). For the low-mass bin, both the N5 and N1 gas profiles appear to be enhanced, on average, with respect to GR, while for the high-mass bin the profiles appear to be suppressed. This can help explain why the  $Y$ -parameters are enhanced in nDGP at lower masses and closer to GR or suppressed at higher masses.

The physical origin of these effects on the gas density is not entirely clear. They

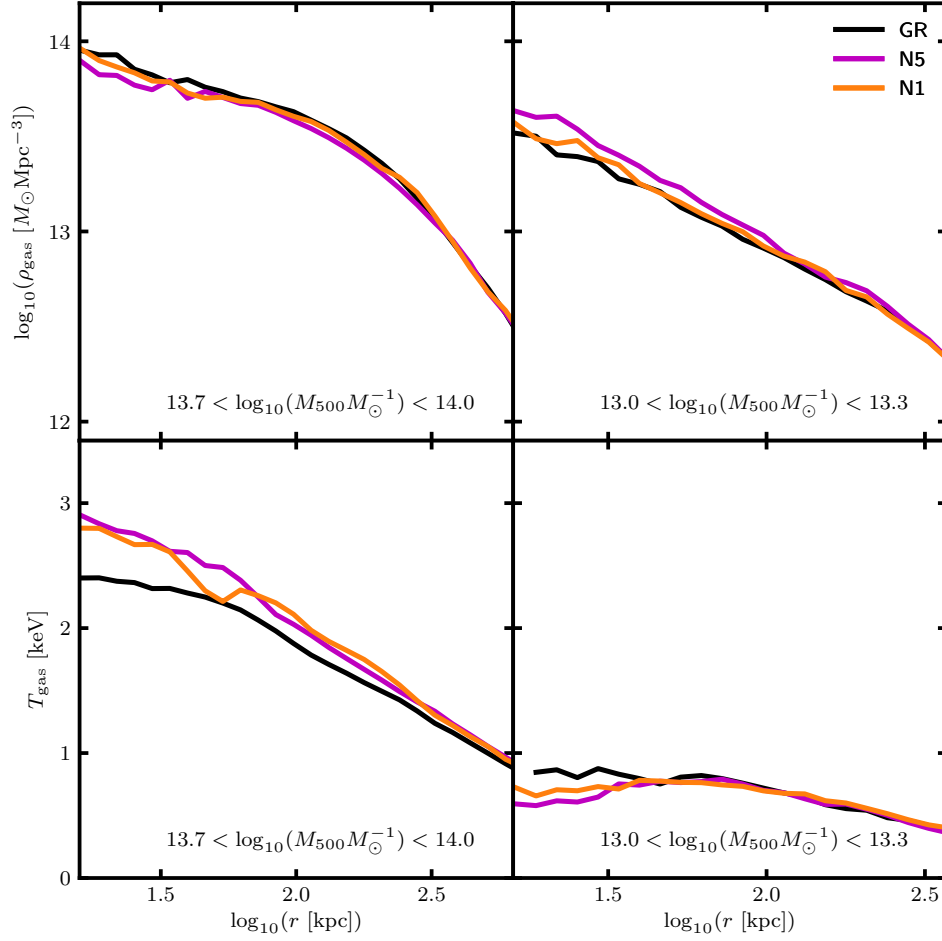


Figure 6.5: Median gas density (*top row*) and temperature (*bottom row*) profiles of haloes from our full-physics SHYBONE simulations (see Sec. 6.2.1) at  $z = 0$ . Data is shown for GR (*black*) and the nDGP models N5 (*magenta*) and N1 (*orange*). The two mass bins used to measure the median profiles are annotated. The maximum radius shown for each column corresponds to  $R_{500}$ .

could be related to the complex interrelations between the nDGP fifth force and baryonic processes such as cooling and feedback. For example, if the fifth force leads to a larger amount of feedback, this would heat up and blow out surrounding gas. This would be consistent with the results shown for the gas density and temperature profiles in the high-mass bin in Fig. 6.5, where the gas density is suppressed and the temperature is enhanced in nDGP compared to GR. The opposite trend is present in the low-mass bin, which would be consistent with a lowering of feedback efficiency in nDGP compared to GR. Another possibility is that the enhancement of the gas speeds due to the fifth force leads to differences in the density profiles between nDGP and GR. This effect can be inherited from times before

the gas falls into haloes and is screened from the fifth force. Haloes of different mass will experience this effect to a different extent as larger haloes are formed from matter and gas further afield.

In the lower panels of Fig. 6.5 we show the halo gas temperature profiles. For the higher mass bin, the profiles in N5 and N1 are both enhanced compared to GR. For N1, this is consistent with the result for the mass-weighted temperature discussed above; however, for N5, the enhancement relative to GR appears to contradict Fig. 6.2. This is actually related to the difference in binning: while the median mass-weighted temperature has been computed using a moving window containing a fixed number of haloes, the temperature profile is computed within a single wide bin. The mean mass of this bin is actually higher in N5 than in GR, indicating that this bin contains a greater number of high-mass haloes in N5 which also have a higher temperature. This supports our decision to use a moving average in Figs. 6.2-6.4, which avoids the issues that arise from having a fixed set of bins for each model. For the lower mass bin, the nDGP temperature profiles are suppressed for radii  $r \lesssim 100\text{kpc}$  and the N1 profile is just slightly enhanced at higher radii. We note that, because there are more particles at the outer radii, which cover a larger volume, these regions have a greater overall contribution to the mass-weighted temperature, which can explain why the latter is enhanced in N1 even though the temperature profile is suppressed at lower radii compared to GR. And, as described above, the difference in binning can make it difficult to directly compare Figs. 6.2 and 6.5.

We finally note that, due to the small box size of our full-physics simulations, we can only rigorously study the scaling relations for halo masses corresponding to galaxy groups. A larger box will be required to rigorously probe the interplay between the fifth force and baryonic physics in galaxy clusters. Galaxy groups, particularly low-mass groups, are typically more susceptible to feedback than cluster-sized objects. This is why, in Fig. 6.2, the scatter in the GR halo temperature is above 10% for low-mass groups and less than 5% for high-mass groups. It will therefore be interesting to see how the nDGP scaling relations compare to GR at these larger masses, where the unpredictable effects from feedback are not as significant. In Chapter 8, we will present a retuned baryonic physics model which we

plan to use for running large-box full-physics simulations of the nDGP model in the future.

### 6.3.2 Concentration-mass relation

In Sec. 6.3.2.1, we discuss the concentration results from our DMO simulations (dashed lines in Figs. 6.6-6.8). Then, in Sec. 6.3.2.2, we summarise the results from full-physics simulations (solid lines in Figs. 6.6-6.8), including the effect of baryons on the model differences. Finally, in Sec. 6.3.2.3, we present a general model for the concentration-mass relation in nDGP.

#### 6.3.2.1 Dark-matter-only concentration

In order to study the concentration over a wide and continuous halo mass range, we have combined the data from our DMO simulations into a single catalogue. In order to avoid resolution issues with the concentration measurement, we exclude haloes which have fewer than 2000 particles (within the radius  $R_{200}$ ) and we leave out L1000 due to its low mass resolution. The resulting catalogue consists of haloes spanning masses  $3.04 \times 10^{11} h^{-1} M_{\odot} \lesssim M_{200} \lesssim 10^{15} h^{-1} M_{\odot}$ . We note that, because L62 has not been run for N2 and N0.5, the data for these models only extends down to mass  $1.278 \times 10^{12} h^{-1} M_{\odot}$  ( $\equiv$  2000 particles from L200). Throughout this section, we will only refer to the results from this combined catalogue; however, in Appendix C.1, we also compare the concentration predictions from each of our DMO simulations, including L1000.

The top row of Fig. 6.6 shows the median concentration as a function of mass for redshifts 0, 1 and 2 (from left to right). The median has been computed using mass bins containing a minimum of 100 GR haloes each: the bins all have equal width in logarithmic mass apart from the highest-mass bin, which is wide enough to enclose the 100 highest-mass haloes. The same set of bins is used for each gravity model. As expected from literature (e.g., Duffy et al., 2008), the median concentration appears to follow a descending power-law relation with the mass. This behaviour arises due to the hierarchical nature of structure formation: higher-mass haloes form at later times when the background density is lower. Therefore, the concentration of these haloes is also typically lower.

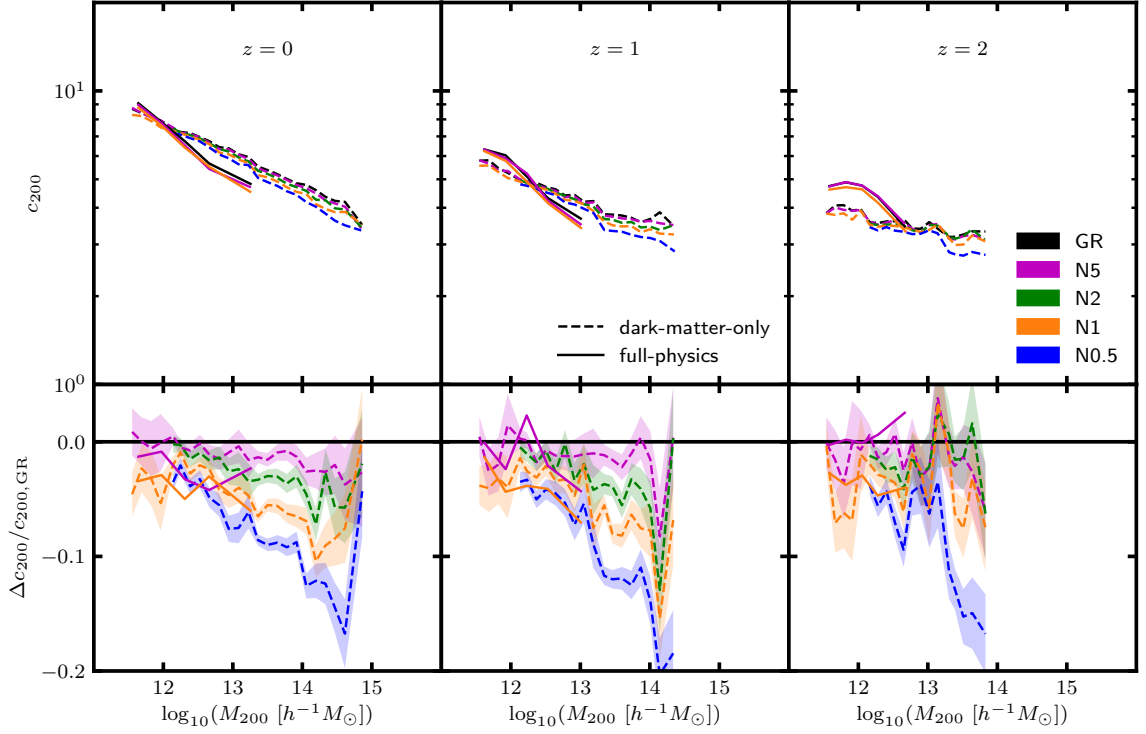


Figure 6.6: Median halo concentration (*top row*) and its relative difference with respect to GR (*bottom row*), as a function of the mean logarithm of the halo mass at redshifts 0, 1 and 2. The data is generated using the dark-matter-only simulations L62, L200 and L500 (*dashed lines*) and our full-physics simulation (*solid lines*), the specifications of which are given in Table 6.1. These have been run for GR (*black*) and the nDGP models N5 (*magenta*), N2 (*green*), N1 (*orange*) and N0.5 (*blue*). The shaded regions in the lower panels show the  $1\sigma$  uncertainty in the relative difference.

The bottom row of Fig. 6.6 shows the relative difference between the nDGP and GR median concentrations. The shaded region shows the  $1\sigma$  error. To calculate this, the standard error of the mean (equal to the standard deviation divided by the square root of the halo count) is computed for each mass bin for GR and nDGP, and then combined in quadrature. We note that, although the nDGP and GR simulations are started from the same initial conditions, the differing gravitational forces affect the trajectories of the simulation particles, which end up at different positions with different velocities, essentially losing much of the memory of their initial states. Therefore, the concentration measurements of each model can be treated as independent, so that the errors may be combined as described. Our results show that the nDGP fifth force causes the concentration to be reduced, since particles experience the fifth force and hence have enhanced velocities before

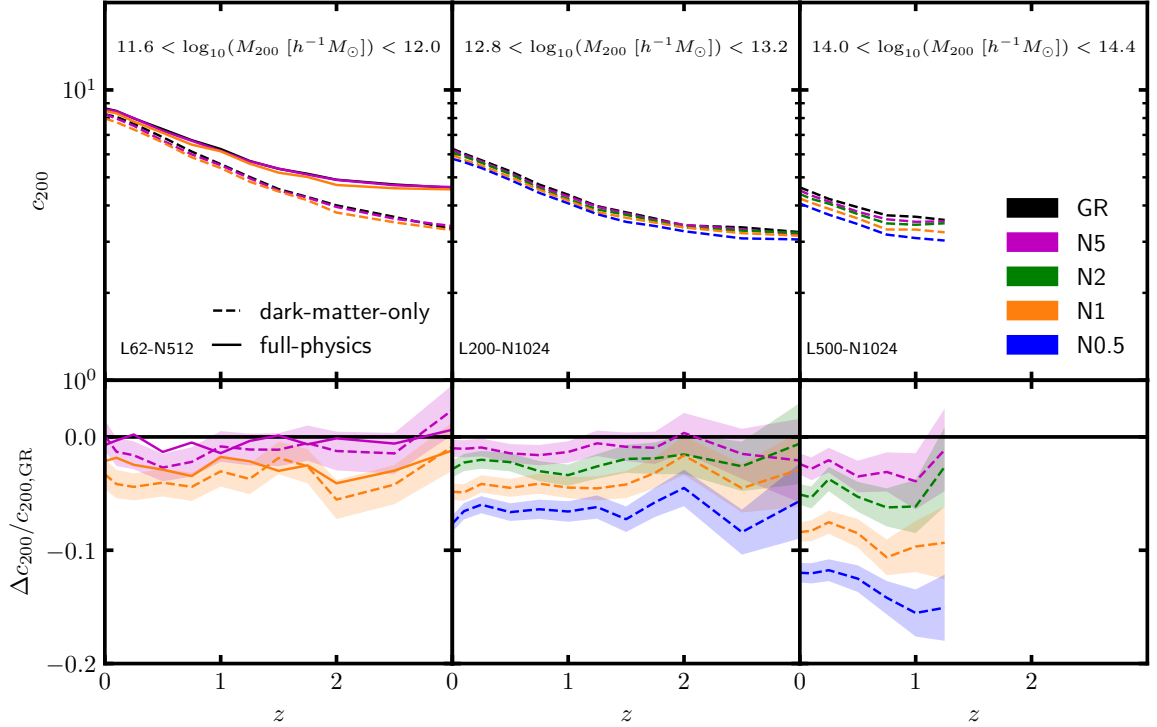


Figure 6.7: Median halo concentration (*top row*) and its relative difference with respect to GR (*bottom row*), as a function of redshift for three mass bins. The data is generated using the dark-matter-only simulations L62, L200 and L500 (*dashed lines*) and the full-physics simulation (*solid lines*), the specifications of which are given in Table 6.1. These have been run for GR (*black*) and the nDGP models N5 (*magenta*), N2 (*green*), N1 (*orange*) and N0.5 (*blue*). The shaded regions in the lower panels show the  $1\sigma$  uncertainty in the relative difference.

they fall into haloes, so that after entering the haloes their higher kinetic energy makes it harder for them to settle towards the central regions. The effect is greater for models which have a stronger fifth force, so the concentration suppression is highest in N0.5 ( $\sim 10\%$  on average) and lowest in N5 (at percent level). At  $z = 0$  and  $z = 1$ , the suppression is greater at higher mass. This appears to be the case for N0.5 at  $z = 2$  as well, but not for weaker models, where the suppression appears to have a much weaker dependence on the halo mass.

To complement these results, we show the median concentration, computed within three mass bins, as a function of redshift in Fig. 6.7. The lower-mass bin,  $10^{11.6}h^{-1}M_{\odot} < M_{200} < 10^{12}h^{-1}M_{\odot}$ , corresponds to galaxy-sized haloes: here, we use haloes from L62, for which we again note that only the GR, N5 and N1 models are available. For the middle-

mass bin,  $10^{12.8}h^{-1}M_{\odot} < M_{200} < 10^{13.2}h^{-1}M_{\odot}$ , we use haloes from L200. For both of these bins, the nDGP suppression of the concentrations appears to be approximately constant over the redshift range  $0 \leq z \leq 3$ , ranging from a couple of percent at most in N5 to about 7% in N0.5.

The higher-mass bin,  $10^{14}h^{-1}M_{\odot} < M_{200} < 10^{14.4}h^{-1}M_{\odot}$ , shown in Fig. 6.7 corresponds to cluster-sized objects; for this, we use haloes from L500. Because clusters typically form at later times, this bin consists of fewer than 100 haloes for redshifts  $z \gtrsim 1.25$ , and we therefore exclude these redshifts from the figure. The suppression of the concentration in nDGP is greater for this bin than for the lower-mass bins, reaching  $\sim 15\%$  in N0.5. This is consistent with the results of Fig. 6.6. As for the other bins, the suppression does not appear to evolve with redshift in N5, N2 and N1. However, for N0.5, the suppression is slightly greater at  $z = 1$  ( $\sim 15\%$ ), than at  $z = 0$  ( $\sim 12\%$ ). We note that the error is also greater at high redshift due to the reduced number of objects, so these results alone do not provide compelling evidence of a redshift evolution of the concentration suppression.

To help make sense of these results, in Fig. 6.8 we show the median density profiles of haloes from a few mass bins at redshifts 0 and 2. These have been computed by measuring the median density, in radial bins spanning  $0.05R_{200}$  to  $R_{200}$ , using the binned haloes. The density has been scaled by  $r^2$  so that the profiles peak at the scale radius,  $R_s$ . This means that the concentration,  $c_{200} = R_{200}/R_s$ , can effectively be read off from the peak radius: a higher (lower) peak radius corresponds to a lower (higher) concentration. In the left column of Fig. 6.8, we show the median profile for haloes from L62 in the mass bin  $10^{11.6}h^{-1}M_{\odot} < M_{200} < 10^{12}h^{-1}M_{\odot}$ . In the right column, we use haloes from L500 within mass bins  $10^{14.4}h^{-1}M_{\odot} < M_{200} < 10^{14.8}h^{-1}M_{\odot}$  and  $10^{13.4}h^{-1}M_{\odot} < M_{200} < 10^{13.8}h^{-1}M_{\odot}$  at redshifts 0 and 2, respectively. We use a lower mass for the  $z = 2$  profile due to the limited number of haloes at higher masses.

For the higher mass bins – where we have seen that there is a greater suppression of the concentration in nDGP models – a clear trend is present: at the outer (inner) regions of haloes, the density is greater (lower) in nDGP than in GR. As mentioned above, this is related to the nature of the Vainshtein screening in nDGP, which suppresses the fifth



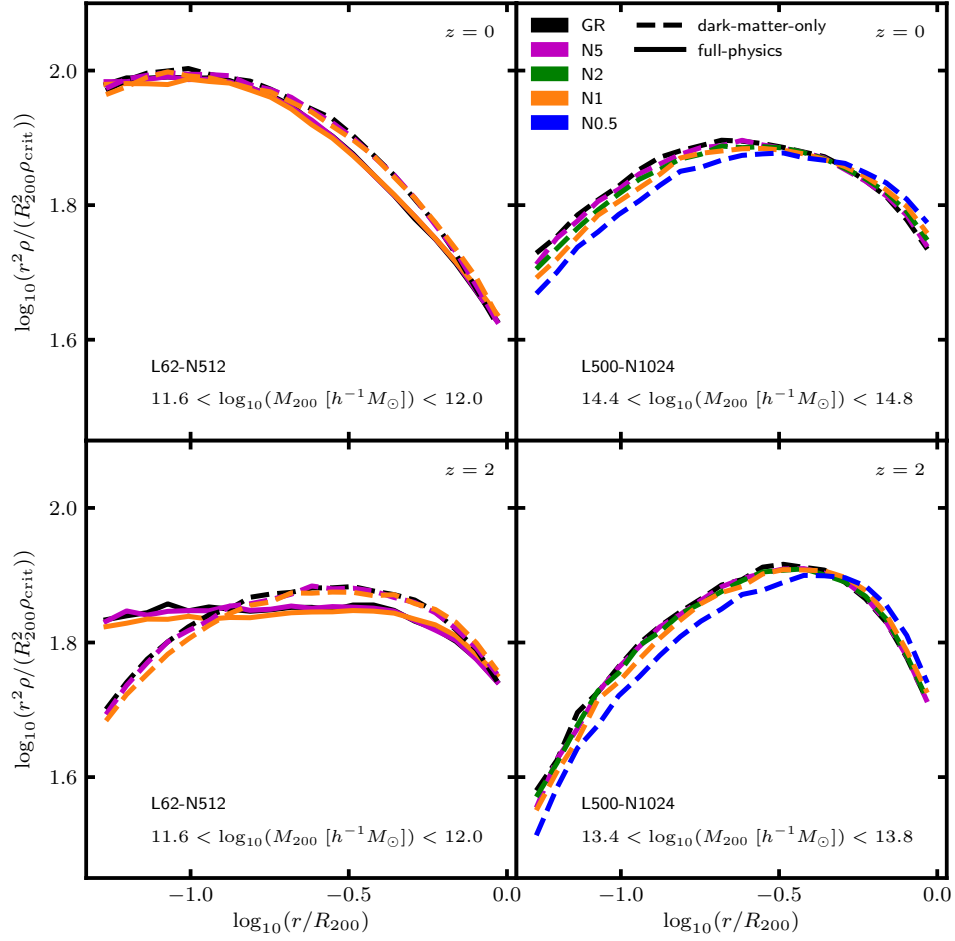


Figure 6.8: Median density profiles of haloes from the AREPO simulations L62 (*left column*) and L500 (*right column*) at redshifts 0 (*top row*) and 2 (*bottom row*). Data from both the full-physics (*solid lines*) and dark-matter-only (*dashed lines*) counterparts of L62 are shown. The L500 simulation includes runs for GR (*black*) and the nDGP models N5 (*magenta*), N2 (*green*), N1 (*orange*) and N0.5 (*blue*), while the L62 simulation includes GR, N5 and N1 only. The mass bins used to measure the median density are annotated.

force on small scales or distances. This means that the fifth force is stronger at large scales, which correspond to the outer regions of these haloes and regions further away from the halo-formation sites. This causes orbiting dark matter particles to undergo an enhanced gravitational acceleration at these regions and have higher kinetic energy, which prevents them from relaxing and settling into lower-radius orbits where the fifth force is suppressed. This causes  $r^2\rho(r)$  to peak at a higher radius in the nDGP models than in GR, resulting in a suppressed concentration. The effect is greatest in N0.5.

For the lower mass bins, we have seen in Figs. 6.6 and 6.7 that the effect of the fifth

force is not as strong. This is consistent with the low-mass density profiles in Fig. 6.8, where the nDGP profiles are closer to GR. However, the density is still slightly reduced at the inner regions and increased at the outer regions, and so the concentration is still suppressed. The reason that the effect is not as strong at low mass is again due to the nature of the Vainshtein screening: lower-mass haloes have a smaller spatial extent, therefore the small-scale suppression of the fifth force is more substantial throughout the range  $r < R_{200}$ . In addition, smaller haloes generally form at higher redshifts, so that the particles inside them have spent less time outside the haloes and are therefore less affected by the fifth force; this is because, once these particles enter haloes, the fifth force is strongly suppressed.

### 6.3.2.2 *Full-physics concentration*

In Figs. 6.6-6.8, we have also included data from our full-physics simulations, which are represented with solid lines. Because these data are only available for the  $62h^{-1}\text{Mpc}$  box, the data only extends to low-mass galaxy clusters (although, we note that the mean logarithmic mass of the rightmost bin shown in Fig. 6.6 is only slightly above  $10^{13}h^{-1}M_{\odot}$ ). Nevertheless, by comparing this to the data from the combined DMO data, we can get an idea of how the results differ when gas and processes such as star formation and feedbacks are included.

From the solid lines in Figs. 6.6 and 6.7, we see that the full-physics concentration is typically greater at lower masses and reduced at higher masses. The full-physics simulations include a gaseous component which, unlike dark matter, is affected by turbulence. This causes the gas cells to slow down and settle at the inner regions of haloes. Also, at the centre of a halo, we are likely to see stellar particles concentrate. This means that the total halo density is enhanced in the inner regions, which is consistent with the stacked density profiles of the full-physics simulations in Fig. 6.8. According to these results, the rescaled density profile becomes approximately flat at the inner regions, corresponding to a  $\rho(r) \propto r^{-2}$  power-law. This clearly deviates from the NFW profile, which follows an  $r^{-1}$  power law in these regions. Because the concentration is a parameter of the NFW profile, we still have to fit Eq. (3.2) in order to measure this. Doing so produces a value that is

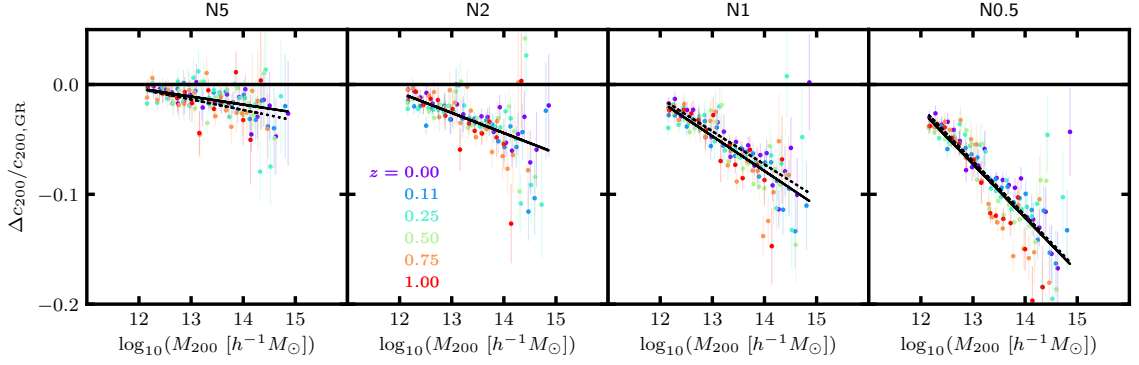


Figure 6.9: Relative difference between the median halo concentration in nDGP with respect to GR, as a function of the mean logarithm of the halo mass. Binned data are shown for all snapshots with redshift  $z \leq 1$ , where the redshift is represented by colour. The data is generated using the dark-matter-only simulations L62, L200 and L500, the specifications of which are given in Table 6.1. These have been run for GR and the nDGP models N5, N2, N1 and N0.5 (shown from left to right). The error bars indicate the  $1\sigma$  uncertainties. The solid lines represent the best-fit linear relations for each panel, while the dashed lines show the predictions from our general model, which is given by Eq. (6.3). For all models, and across a halo mass range of four orders of magnitude, the fitting function gives a percent-level agreement with the simulation measurement of the concentration decrement at  $0 \leq z \leq 1$ .

either higher or lower than for DMO haloes with the same mass.

Despite the difference in the absolute concentrations, the suppression of the concentration in nDGP appears to have a similar magnitude in the full-physics and DMO simulations, according to the bottom-left panel of Fig. 6.7, for galaxy-sized haloes in the redshift range  $0 \leq z \leq 3$ . In Fig. 6.6, the dashed and solid lines in the lower panels also appear to have a similar magnitude; however, we are unable to rigorously test this for masses  $M_{200} \gtrsim 10^{13} h^{-1} M_{\odot}$ , which would require full-physics simulations of nDGP that have a much larger box size. Such simulations are highly expensive, and are therefore left for future work.

### 6.3.2.3 Modelling the concentration in nDGP

From Figs. 6.6 and 6.7, it appears that the suppression of the DMO halo concentration in nDGP grows with mass and is approximately constant as a function of redshift. In Fig. 6.9, we show the binned relative difference data from our combined DMO simulation data for all snapshots at  $z \leq 1$ . The data appears to follow a linear trend as a function of the mass,

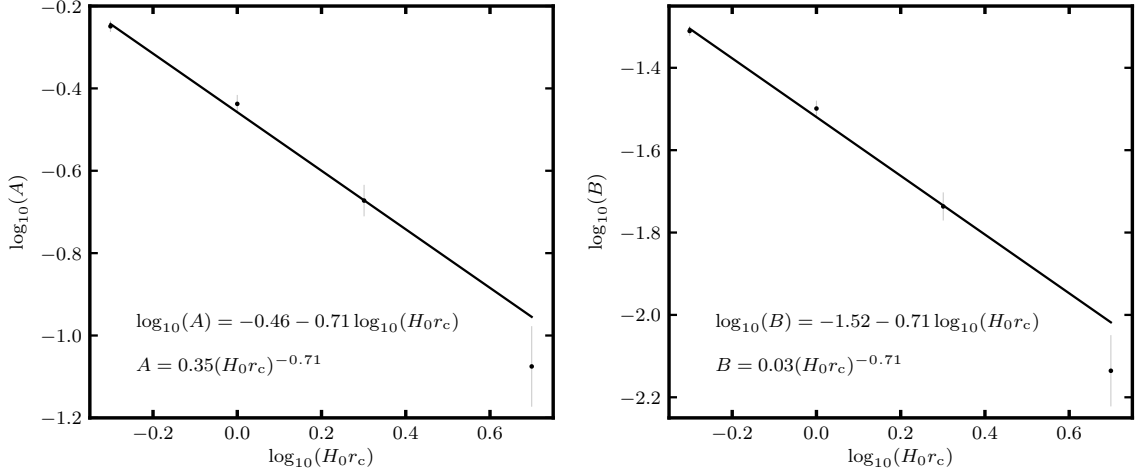


Figure 6.10: Best-fit values of the parameters  $A$  and  $B$  of Eq. (6.1) as a function of the logarithm of  $H_0 r_c$ , where  $r_c$  is the cross-over scale of nDGP gravity. The best-fit values of the data points have been computed by fitting Eq. (6.1) to the data shown in the four panels of Fig. 6.9, which correspond to the models N5, N2, N1 and N0.5. The error bars represent the  $1\sigma$  uncertainties in the data, obtained from the weighted least squares fits. The solid lines show best-fit power-law fits of the four data points, Eq. (6.2), which are annotated.

therefore we can model this using:

$$\Delta c/c_{\text{GR}} = A - B \log_{10}(M_{200} M_{\odot}^{-1} h), \quad (6.1)$$

where  $A$  and  $B$  are parameters representing the amplitude and slope of the relation, respectively. This does not include any dependence on redshift. For the N0.5 data, there is a clear  $z$ -dependence, with low- $z$  (blue) data having a smaller suppression than high- $z$  (red) data; however, the suppression in different snapshots is still quite close, and there does not appear to be any  $z$ -evolution for the other, more realistic, models of nDGP.

The solid lines in Fig. 6.9 are the best-fit relations for each model. These are created by using weighted least squares to fit Eq. (6.1) to the data points, where points with large (small) error bars are given smaller (larger) weighting. In Fig. 6.10, we show the best-fit values of  $A$  and  $B$  as a function of the  $H_0 r_c$  parameter which characterises the nDGP models. Both  $A$  and  $B$  appear to be well-described by a power-law relation. Using weighted

least squares to fit the four data points, we obtain the following best-fit relations:

$$\begin{aligned} A &= (0.35 \pm 0.01)(H_0 r_c)^{-0.71 \pm 0.05}, \\ B &= (0.0302 \pm 0.0008)(H_0 r_c)^{-0.71 \pm 0.05}. \end{aligned} \quad (6.2)$$

Interestingly, the relations both have power-law slope  $-0.71 \pm 0.05$ . They can therefore be combined with Eq. (6.1) to form the following simple relation:

$$\begin{aligned} \frac{\Delta c}{c_{\text{GR}}} &= [(0.35 \pm 0.01) - (0.0302 \pm 0.0008) \log_{10}(M_{200} M_{\odot}^{-1} h)] \\ &\quad \times (H_0 r_c)^{-0.71 \pm 0.05}. \end{aligned} \quad (6.3)$$

This can be used to predict the suppression of the concentration in nDGP, as a function of the halo mass  $M_{200}$  and model parameter  $H_0 r_c$ . The dashed lines in Fig. 6.9 show the model predictions for our four nDGP models. The agreement with the data is generally very good for the full mass range,  $10^{12} h^{-1} M_{\odot} \lesssim M_{200} \lesssim 10^{15} h^{-1} M_{\odot}$ , of our simulation data. The agreement is particularly good for weaker models, where it appears to match  $z = 0$  and  $z = 1$  data equally well. For N0.5, which is our strongest model, our relation appears to slightly underestimate the concentration suppression for high-redshift data; however, the overall level of agreement is still very good, considering that the model is able to give reasonable predictions for such a wide range of nDGP models and masses.

### 6.3.3 Halo mass function

The HMF does not have a strict mass resolution requirement like the concentration, therefore we use all haloes which have at least 100 particles within the radius  $R_{500}$ . We again combine the halo data from our DMO simulations, and the relaxed resolution requirement means that we can now also include L1000 haloes. The HMF is computed using mass bins with equal logarithmic width 0.2. The halo count in each bin is divided by the total volume from all contributing simulations: for example, the volume is  $62^3 h^{-3} \text{Mpc}^3$  for the lowest-mass bins where only the L62 box has sufficient resolution, and  $(62^3 + 200^3 + 500^3 + 1000^3) h^{-3} \text{Mpc}^3$  for the highest-mass bins where all simulations have sufficient resolution. In Appendix C.1, we also assess the consistency of our DMO simulations by comparing the HMF predictions at different resolutions.

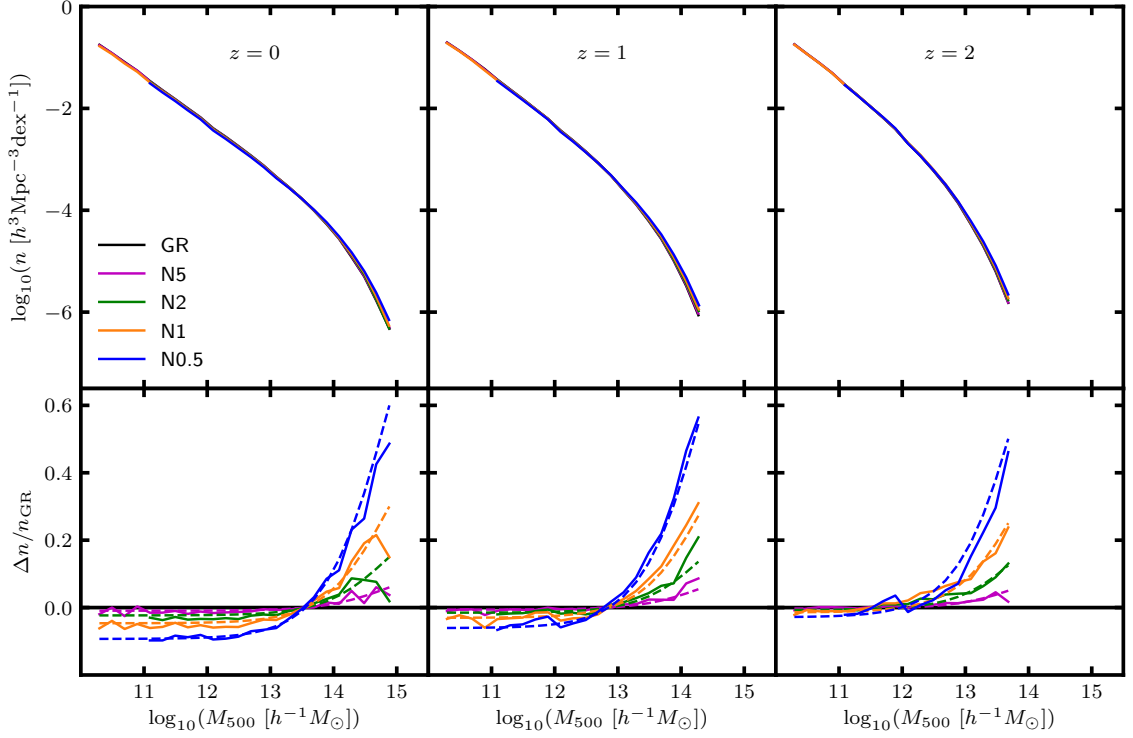


Figure 6.11: Halo mass function (*top row*) and its relative difference in nDGP with respect to GR (*bottom row*), as a function of the mean logarithm of the halo mass at redshifts 0, 1 and 2. The data is generated using our AREPO dark-matter-only simulations, the specifications of which are given in Table 6.1. These have been run for GR (*black*) and the nDGP models N5 (*magenta*), N2 (*green*), N1 (*orange*) and N0.5 (*blue*). The dashed lines show the predictions from our general fitting model, which is given by Eqs. (6.4, 6.5).

The binned HMF is shown in Fig. 6.11 for redshifts 0, 1 and 2, where only mass bins containing at least 100 haloes are displayed. We note that, because our highest-resolution simulation L62 has been run for N5 and N1 only, the data for these models extends to lower masses than the other models. The relative difference between the nDGP and GR results is shown in the lower panels. For all three redshifts, the HMF is significantly enhanced in nDGP relative to GR at high mass: for N0.5, the HMF is enhanced by up to 60%, while for N5 the enhancement is less than 10%. On the other hand, the HMF is suppressed at lower masses in nDGP, by up to  $\sim 10\%$  in N0.5 and a couple of percent in N5. The threshold mass above which the HMF is enhanced and below which it is suppressed is higher at lower redshifts, with values  $\sim 10^{13.5}h^{-1}M_\odot$  at  $z = 0$  and  $\sim 10^{12}h^{-1}M_\odot$  at  $z = 2$ . The low-mass suppression of the HMF also decreases with redshift.

These results can again be explained by the behaviour of the fifth force, which enhances

the overall strength of gravity on large scales, accelerating the formation of high-mass haloes so that there is a greater abundance of these objects in nDGP compared to GR at a given time. On the other hand, the abundance of low-mass haloes, which undergo an increased number mergers, is reduced. The mass threshold between HMF enhancement and HMF suppression is reduced at higher redshifts, which is likely simply because the masses of a given population of haloes are lower at earlier times.

Structure formation is sped up by a greater extent in models which feature a stronger fifth force, so the effects described above are greater for N0.5 than for weaker models. The enhancement of the HMF is greatest at the high-mass end. Therefore, by using observations of high-mass galaxy clusters from ongoing and upcoming galaxy surveys (e.g., [Levi et al., 2013](#); [Laureijs et al., 2011](#); [Ivezic et al., 2008](#)), it will be possible to make powerful constraints of nDGP. However, any tests of the nDGP model of this kind may be affected by the cluster observable-mass scaling relations discussed earlier, and this should be investigated in a future work.

From the lower panels of Fig. 6.11, it appears that, for any model, i.e., for a given choice of  $H_0 r_c$ , the HMF enhancement has a constant shape, but shifts downwards and towards larger  $M_{500}$  as one goes to lower redshifts. Therefore, it can be well-described by the following model:

$$\frac{\Delta n}{n_{\text{GR}}} = A(H_0 r_c) \left[ \tanh \left( \log_{10}(M_{500} M_{\odot}^{-1} h) - B(z) \right) + C(z) \right]. \quad (6.4)$$

We use a portion of a tanh function to represent the mass-dependent shape, which is level at low mass and rises steeply at high mass. We also include the following parameters:  $A(H_0 r_c)$  controls the amplitude, which depends on the model parameter  $H_0 r_c$ ;  $B(z)$  represents the  $z$ -dependent shift along the mass axis; and  $C(z)$  represents the  $z$ -dependent shift along the  $\Delta n/n_{\text{GR}}$  axis. By adopting simple linear models for each of these parameters, and by combining the data from all simulation snapshots in the range  $0 \leq z \leq 2$ , we have used unweighted least squares to obtain the following best-fit results:

$$\begin{aligned} A(H_0 r_c) &= (0.342 \pm 0.014)(H_0 r_c)^{-1}, \\ B(z) &= (14.87 \pm 0.03) - (0.481 \pm 0.010)z, \\ C(z) &= (0.864 \pm 0.008) + (0.047 \pm 0.005)z. \end{aligned} \quad (6.5)$$

The predictions of this calibrated model are indicated by the dashed lines in Fig. 6.11. The agreement with the simulation data is excellent for all models for the mass ranges shown, which span 4–5 decades depending on redshift. At  $z = 0$  and  $z = 1$ , apart from the highest mass bin where data is noisy, the agreement between the fitting function and simulation measurements is within  $\sim 3\%$ ; at  $z = 2$ , the agreement is within  $\sim 3\%$  for all but the strongest model ( $H_0 r_c = 0.5$ ) where we still have a  $\lesssim 5\%$  accuracy. In the limit  $H_0 r_c \rightarrow \infty$ , where nDGP becomes GR, our model predicts a relative difference of zero as expected. However, we note that our model will predict a constant relative difference if extrapolated to higher masses. This behaviour may not be physically accurate, but the high halo masses are beyond the dynamical range of our simulations and so we cannot test this reliably. Therefore, the model in Eqs. (6.4, 6.5) should only be used for the mass range  $10^{11} h^{-1} M_\odot \lesssim M_{500} \lesssim M_{\max}(z)$ , where  $M_{\max}(z)$  is the maximum mass used to calibrate the above model at a given redshift. The latter can be estimated using the relation:

$$\log_{10} (M_{\max} M_\odot^{-1} h) = 14.81 - 0.54z, \quad (6.6)$$

which we have calibrated using snapshots in the range  $0 \leq z \leq 2$ .

In this section, we have focused on the mass definition  $M_{500}$ , which is commonly used in cluster number counts studies (e.g., Ade et al., 2016). For completeness, we also present, in Appendix C.2, results and modelling for mass definition  $M_{200}$ .

## 6.4 Summary, Discussion and Conclusions

In this chapter, we have extended our general framework for cluster tests of gravity (Fig. 1.2) to the popular nDGP model, in which a fifth force is able to act over sufficiently large scales.

Using the first cosmological simulations that simultaneously incorporate full baryonic physics and the nDGP model, we have studied the observable-mass scaling relations for three mass proxies (see Sec. 6.3.1). For groups and clusters in the mass range  $M_{500} \lesssim 10^{14.5} M_\odot$ , our results show that for the N1 model, the  $\bar{T}_{\text{gas}}(M)$  relation is enhanced by about 5% with respect to GR, while the  $Y_{\text{SZ}}(M)$  and  $Y_X(M)$  relations are both enhanced by 10%-15% at low masses but more closely match the GR relations at high masses. For N5,



which is much weaker than N1, the  $\bar{T}_{\text{gas}}(M)$  relation closely resembles the GR relation, while the  $Y_{\text{SZ}}(M)$  and  $Y_{\text{X}}(M)$  relations are enhanced by up to 5% at low mass and suppressed by up to 5% at high mass. These deviations from GR could be related to the effect of the fifth force on gas velocities during cluster formation, and they also hint at an interplay between the fifth force and stellar and black hole feedback.

Using a suite of DMO  $N$ -body simulations, which cover a wide range of resolutions and box sizes, we have found that, in nDGP, the concentration is typically suppressed relative to GR, varying from a few percent in N5 to up to  $\sim 15\%$  in N0.5 (see Sec. 6.3.2). Using stacked density profiles at different mass bins, we have shown that this behaviour is caused by a reduced (increased) density at the inner (outer) halo regions. Including full baryonic physics significantly affects the concentration-mass relation; however, our results show that, for masses  $M_{200} \lesssim 10^{13} h^{-1} M_{\odot}$ , the model differences between nDGP and GR still have a similar magnitude compared to the DMO simulations.

By combining the data from our  $z \leq 1$  simulation snapshots, we have calibrated a general model, given by Eq. (6.3), which is able to accurately predict the suppression of the halo concentration with respect to the GR results as a function of the halo mass and the  $H_0 r_c$  parameter of nDGP over ranges  $10^{12} h^{-1} M_{\odot} \lesssim M_{200} \lesssim 10^{15} h^{-1} M_{\odot}$  and 0.5-5, respectively. This model can be included in our MCMC pipeline for converting between mass definitions in case, for example, the theoretical predictions and observables are defined with respect to different spherical overdensities. Our model can also be used, along with the HMF, to predict the nonlinear matter power spectrum, which can also be used to constrain gravity.

We have also used our DMO simulations to study the HMF over the mass range  $1.52 \times 10^{10} h^{-1} M_{\odot} \leq M_{500} \lesssim 10^{15} h^{-1} M_{\odot}$  at redshifts 0, 1 and 2 (see Sec. 6.3.3). Our results (Fig. 6.11), indicate that the nDGP HMF is enhanced at high masses (by up to  $\sim 60\%$  in N0.5) and suppressed at low masses (by  $\sim 10\%$  in N0.5) compared to GR. These results indicate the potential constraining power from using the observed mass function to probe the  $H_0 r_c$  parameter of nDGP. By combining the data from our  $z \leq 2$  snapshots, we have calibrated a general model, given by Eq. (6.4), which can accurately reproduce the

HMF enhancement as a function of the halo mass, redshift and  $H_0 r_c$  parameter. This model can be used for theoretical predictions of the nDGP HMF (using a parameter-dependent GR calibration) in our MCMC pipeline.

In Chapter 4, we showed that a model for the  $f(R)$  dynamical mass enhancement can be used to predict observable-mass scaling relations in  $f(R)$  gravity using their GR counterparts. Such a model in nDGP could similarly be useful to help understand the enhancements of the temperature and SZ and X-ray  $Y$ -parameters observed in this work. This is left to a future study. For now, though, we note that the scaling relations in nDGP still appear to follow power-law relations as a function of the mass: the  $\bar{T}_{\text{gas}}(M)$  relation in N1 can be related to the GR relation by a simple rescaling of the amplitude, whereas the  $Y_{\text{SZ}}(M)$  and  $Y_{\text{X}}(M)$  relations appear to have shallower slopes in N5 and N1 than in GR. Therefore, in our future MCMC pipeline for obtaining constraints of nDGP, we can still assume the GR power-law form of the scaling relations by allowing the parameters controlling the amplitude and slope to vary along with the cosmological and nDGP parameters (e.g., [de Haan et al., 2016](#); [Bocquet et al., 2019](#)).

Although our simulations have only been run for a single choice of cosmological parameters, we expect that our models for the enhancements of the halo concentration and HMF will have a reasonable accuracy for other (not too exotic) parameter values. The gravitational force enhancement in nDGP, given by  $[1 + 1/(3\beta)]$ , has only a weak dependence on  $\Omega_{\text{M}}$ : for the N1 model ( $\Omega_{\text{rc}} = 0.25$ ), the force enhancement varies within a very small range (roughly 12.1% – 12.6%) for  $\Omega_{\text{M}} \in [0.25, 0.35]$  at the present day, and the range of variation is even smaller at higher redshifts. Therefore, for now we assume that the effects of the cosmological parameters on the concentration and HMF are approximately cancelled out in the ratios  $\Delta c/c_{\text{GR}}$  and  $\Delta n/n_{\text{GR}}$ . However, we will revisit this in a future work, using a large number of nDGP simulations that are currently being run for different combinations of cosmological parameters, before these models are used in tests of gravity using observational data.

Finally, we note that, because the SHYBONE simulations have a small box size ( $62h^{-1}\text{Mpc}$ ), it is difficult to robustly model the observable-mass scaling relations for

cluster-sized objects ( $M_{500} \gtrsim 10^{14} M_{\odot}$ ). It would therefore be useful to revisit this study using full-physics nDGP simulations with a larger box. We have been fine-tuning a new baryonic model which can allow TNG-like simulations to be run at a much lower resolution, making it possible to run large simulations with reduced computational cost. We will present this model in Chapter 8, in which we will also revisit our  $f(R)$  scaling relation results using much larger simulations.

# The impact of modified gravity on the Sunyaev-Zel'dovich effect

## 7.1 Introduction

The SZ effect is caused by the inverse Compton scattering of CMB photons off of high-energy electrons within ionised gas. The effect is made up of two measurable components: a thermal (tSZ) component which arises due to the random thermal motions of the electrons; and a (much smaller) kinematic (kSZ) component resulting from the bulk motion of the gas relative to the CMB rest frame (e.g., [Sunyaev and Zeldovich, 1972, 1980](#)). The tSZ and kSZ signals are both highly correlated with the presence of large-scale structures such as groups and clusters of galaxies. Their power spectra are therefore extremely sensitive to the values of cosmological parameters which affect the growth of large-scale structure, offering the possibility of probing a wide range of cosmological models, including MG theories in which the strength of gravity is enhanced.

A number of works have made use of the tSZ power spectrum to constrain cosmological parameters including  $\Omega_M$ ,  $\sigma_8$ , the dark energy equation of state parameter and the neutrino mass (e.g., [Horowitz and Seljak, 2017](#); [Hurier and Lacasa, 2017](#); [Bolliet et al., 2018](#); [Salvati et al., 2018](#)). Meanwhile, as the precision of measurements of the kSZ power continues to improve, a number of works have identified this as a promising probe for future constraints

of dark energy and MG theories (e.g., [Ma and Zhao, 2014](#); [Bianchini and Silvestri, 2016](#); [Roncarelli et al., 2018](#)). The wealth of high-quality observational data coming from current and upcoming surveys (e.g., [Sievers et al., 2013](#); [Aghanim et al., 2016](#); [George et al., 2015](#); [Reichardt et al., 2020](#); [Ade et al., 2019](#); [Abazajian et al., 2016](#)) for both the tSZ and kSZ effects make it an exciting time for this growing area.

In addition to their sensitivity to cosmology, the tSZ and kSZ power spectra are also highly sensitive to non-gravitational processes, such as star formation, cooling and stellar and black hole feedback, which can alter the thermal state of the intra-cluster medium (e.g., [McCarthy et al., 2014](#); [Park et al., 2018](#)). Without a careful consideration of these processes, which are still not fully understood, this could pose a barrier to making reliable constraints. Incorporating full-physics baryonic models in numerical simulations, along with the cosmological model of interest, is now a vital step in order to understand the potential sensitivity of the constraints to baryonic physics.

In this chapter, we study the effects of  $f(R)$  gravity and nDGP on the tSZ and kSZ power spectra. These are expected to be altered by the effects of the fifth force on the abundance and peculiar motion of large-scale structures, and on the temperature of the intra-cluster gas via the enhancement of the halo gravitational potential. We make use of the full-physics SHYBONE simulations, which were also used in Chapters 4 and 6 to study the observable-mass scaling relations in  $f(R)$  gravity and nDGP, respectively. We measure the power spectra using mock maps of the tSZ and kSZ effects, which are generated using the simulation data. We also study non-radiative simulations (using the same cosmological parameters and initial conditions), allowing us to single out fifth force and baryonic feedback effects.

The chapter is structured as follows: in Sec. 7.2, we describe the simulations used in this chapter and our methods for predicting the SZ power spectra; our main results are presented in Sec. 7.3; and, finally, we present a summary of our findings and their significance in Sec. 7.4.

## 7.2 Simulations and methods

In Sec. 7.2.1, we describe the simulations used in this chapter. Then, in Sec. 7.2.2, we present our procedure for generating SZ maps from the simulation data.

### 7.2.1 Simulations

The results discussed in this chapter have been produced using the SHYBONE simulations (Arnold et al., 2019b; Hernández-Aguayo et al., 2020), which have already been described in Chapters 4 and 6. In addition to these full-physics simulations, we have also used non-radiative counterparts for the  $f(R)$  model, using identical initial conditions and cosmological parameters. We have not run non-radiative counterparts for the nDGP model since these are computationally expensive to perform and there is already a lot of information provided by the existing simulations. Particle data has been saved at various snapshots: the  $f(R)$  data consists of 46 snapshots between  $z = 3$  and  $z = 0$ , whereas the nDGP data includes 100 snapshots between  $z = 20$  and  $z = 0$ .

### 7.2.2 SZ maps

In the generation of each SZ map, a light cone is first constructed using our simulation snapshots with  $z \leq 3$ . We use a field of view of  $1^\circ \times 1^\circ$  for the light cone, which is aligned along a specified direction from an imaginary observer placed at the centre of the simulation box at  $z = 0$ . The box is repeated along this direction and, at a given distance from the observer, the snapshot that is closest to the corresponding redshift is used. Each snapshot is randomly rotated and shifted in order to reduce statistical correlations caused by the repetition of the box.

The  $1^\circ \times 1^\circ$  field of view is split into a  $512 \times 512$  grid of pixels, and an imaginary light ray is fired along the central axis of each pixel from  $z = 3$  to the observer. For each gas cell, an effective size,  $s$ , is defined, which can be used to determine whether it intersects with the light ray. By approximating the gas cells as spherical, the radius,  $r_{\text{cell}}$ , of a gas

cell can be estimated using:

$$r_{\text{cell}} = 2.5 \left( \frac{3V_{\text{cell}}}{4\pi} \right)^{\frac{1}{3}}, \quad (7.1)$$

where  $V_{\text{cell}}$  is the volume of the gas cell. This quantity is similar to the smoothing radius in smoothed-particle hydrodynamics, with the factor 2.5 used to smooth the gas distribution. However, in the mock SZ map, the minimum length scale that is resolved (at a given distance) is the pixel side-length,  $r_{\text{pixel}}$ . The effective size of a gas cell is therefore set as follows:

$$s = \begin{cases} r_{\text{pixel}} & \text{if } r_{\text{cell}} < r_{\text{pixel}}. \\ r_{\text{cell}} & \text{if } r_{\text{cell}} \geq r_{\text{pixel}}. \end{cases} \quad (7.2)$$

A gas cell contributes to the SZ signal of a pixel if the distance between its centre of mass and the light ray is smaller than  $s$ .

The tSZ effect is quantified by the Compton  $y$ -parameter, which can be computed via an integral of the electron pressure along the line of sight as follows:

$$y = \frac{\sigma_{\text{T}}}{m_{\text{e}}c^2} \int n_{\text{e}} T_{\text{gas}} dl. \quad (7.3)$$

This is evaluated for each pixel  $ij$  via a summation over all gas cells that intersect with the light ray:

$$y_{ij} = \frac{\sigma_{\text{T}}}{m_{\text{e}}c^2} \sum_{\alpha} p_{\alpha} w_{\alpha,ij}, \quad (7.4)$$

where  $w_{\alpha,ij}$  is a normalised smoothing kernel. The quantity  $p_{\alpha}$  is given by:

$$p_{\alpha} = \frac{N_{\text{e},\alpha}}{s_{\alpha}^2} T_{\alpha}, \quad (7.5)$$

where  $N_{\text{e},\alpha}$ ,  $s_{\alpha}$  and  $T_{\alpha}$  are the electron number count, the effective size and the temperature of gas cell  $\alpha$ , respectively. Note that we have not accounted for the relativistic SZ (rSZ) effect in our calculations. The rSZ effect can induce a significant bias in the measurement of the  $y$ -parameter for the most massive clusters (see, e.g., [Erler et al., 2018](#)). However, the effect is much smaller for lower-mass objects which have a lower gas temperature. Since our simulations contain only galaxy groups and low-mass clusters ( $M_{500} \lesssim 10^{14.5} M_{\odot}$ ), we expect that including the rSZ effect would have a modest impact on our tSZ power spectrum results. In particular, we expect the effect on the model differences to be very small, but

this is something that should be tested in the future with large simulations that contain a fair sample of cluster-sized objects.

The kSZ effect is quantified by the  $b$ -parameter:

$$b = \sigma_T \int \frac{n_e v_r}{c} dl, \quad (7.6)$$

where  $v_r$  is the radial component of the gas peculiar velocity and  $b$  is positive (negative) for gas that is moving away from (towards) the observer. The  $b$ -parameter is equivalent to the CMB temperature fluctuation due to the kSZ effect:  $b = -\Delta T/T$ . This is evaluated for each pixel as follows:

$$b_{ij} = \frac{\sigma_T}{c} \sum_{\alpha} q_{\alpha} w_{\alpha,ij}. \quad (7.7)$$

The quantity  $q_{\alpha}$  is given by:

$$q_{\alpha} = \frac{N_{e,\alpha}}{s_{\alpha}^2} v_{r,\alpha}, \quad (7.8)$$

where  $v_{r,\alpha}$  is the radial component of the peculiar velocity of gas cell  $\alpha$ .

We have generated 14 independent light cones, each aligned along a unique direction. The same set of directions has been used to construct the maps for each gravity model and for both the full-physics and non-radiative data. This means that for any two maps aligned in the same direction, the only differences are caused by the contrasting gravity models and hydrodynamics schemes. The tSZ and kSZ maps corresponding to one of the light cones, generated using the  $f(R)$  simulations, are shown in Figs. 7.1 and 7.2, respectively. For each figure, the GR maps are shown in the left column, with the map from the full-physics run in the top row and the map for the non-radiative simulation in the bottom row.

The bright yellow peaks in the tSZ maps, which correspond to a high  $y$ -parameter, trace hot gas within groups and clusters of galaxies. These peaks are found at the same positions in both the full-physics and non-radiative maps. However, the addition of feedback mechanisms, which create winds that heat up and blow gas out of haloes, cause the peaks to appear more diffuse in the full-physics map. The kSZ map is made up of dark and bright regions, which correspond to negative and positive values of the  $b$ -parameter, respectively.

Rather than the absolute maps of F6 and F5, which are visually very similar to the GR maps, we display residual maps to indicate the main differences. These are shown



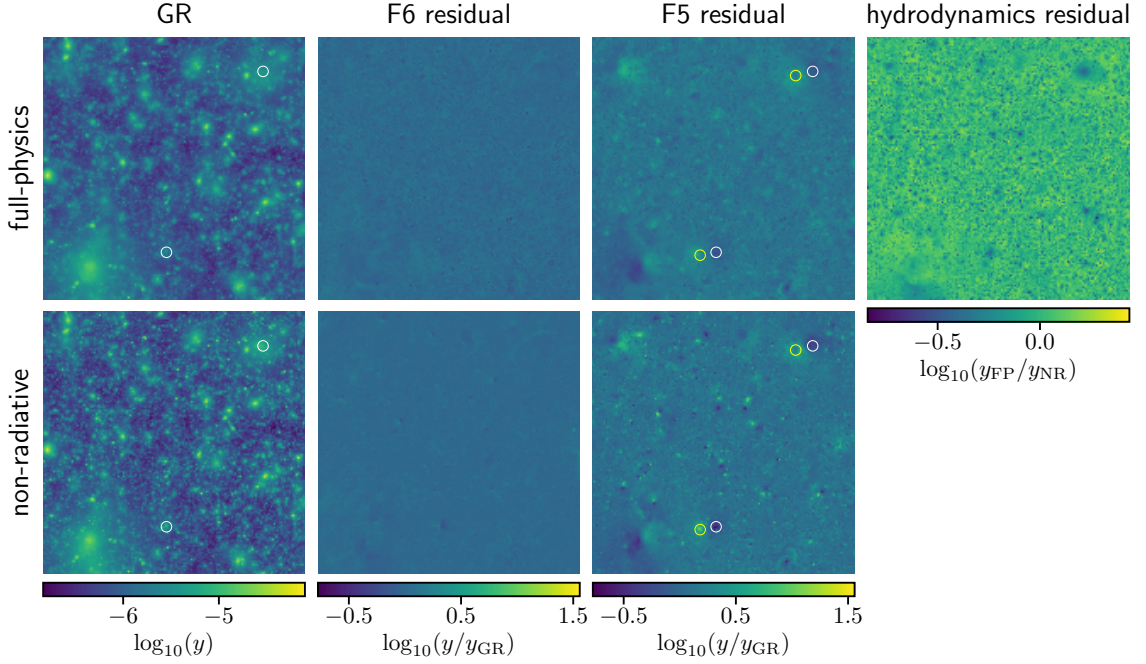


Figure 7.1: Maps of the thermal SZ effect in GR (*first column*), its relative enhancement in F6 (*second column*) and F5 (*third column*) with respect to GR, and the relative difference between the full-physics and non-radiative GR maps (*fourth column*). The maps have a side length of  $1^\circ$  and a  $512 \times 512$ -pixel resolution, and have been constructed from the SHYBONE simulations (see Sec. 7.2). Both the full-physics (*top row*) and non-radiative (*bottom row*) runs are shown. The  $y$ -parameter is computed for each pixel using Eq. (7.4). The rings indicate two haloes whose positions are shifted in F5 (*yellow*) relative to GR (*white*).

in the second and third columns of Figs. 7.1 and 7.2. The tSZ residuals represent the enhancement of the  $f(R)$   $y$ -parameter with respect to GR for each pixel. The F6 residuals are quite close to zero across the field of view, owing to the efficient screening of the fifth force in galaxy groups and clusters for this model. However, for the F5 model, for which the fifth force is more prominent, the residuals appear more complex. Pairs of bright and dark regions, two of which are indicated by rings placed in Fig. 7.1, are visible throughout the images. These are caused by the shift of halo positions in F5 compared to GR, with each dark (bright) region corresponding to the position in GR (F5). While this in itself does not provide useful information about the effect of the fifth force on the tSZ effect, we note that at the extremes the positive residuals ( $\log_{10}(y/y_{\text{GR}}) \approx 1.5$ ) are greater in magnitude than the negative residuals ( $\log_{10}(y/y_{\text{GR}}) \approx -0.8$ ), indicating that the tSZ effect is strengthened on average in F5 compared to GR.

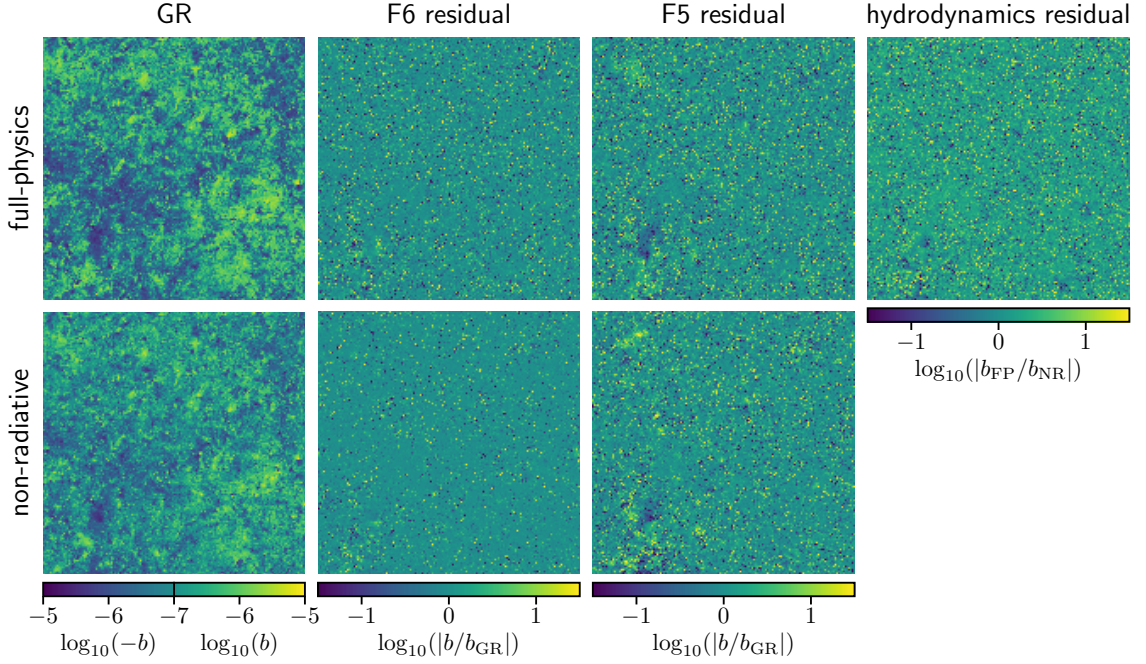


Figure 7.2: Maps of the kinetic SZ effect in GR (*first column*), its absolute relative enhancement in F6 (*second column*) and F5 (*third column*) with respect to GR, and the absolute relative difference between the full-physics and non-radiative GR maps (*fourth column*). The maps have a side length of  $1^\circ$  and a  $512 \times 512$ -pixel resolution, and have been constructed from the SHYBONE simulations (see Sec. 7.2). Both the full-physics (*top row*) and non-radiative (*bottom row*) runs are shown. The  $b$ -parameter is computed for each pixel using Eq. (7.7).

For the kSZ signal, the  $f(R)$  gravity residuals correspond to the enhancement of the absolute value of the  $b$ -parameter with respect to GR. A higher value of  $b$  indicates that gas is moving faster with respect to the CMB rest-frame. Many individual pixels gain much higher and much lower values of  $b$ , seemingly at random, across the field of view. This is caused by the effect of the fifth force on the motion of the gas. Pairs of bright and dark regions are also just visible in the F5 residual map, again corresponding to the relative shifts in halo position with respect to GR.

In the rightmost columns of Figs. 7.1 and 7.2, we show the relative difference between the full-physics and non-radiative GR maps. The tSZ results indicate that within haloes the tSZ signal is suppressed (dark blue regions) by up to 86% and boosted outside haloes (bright yellow regions) by up to 173%. This is caused by the ejection of gas from haloes by feedback mechanisms, causing the electron pressure to be lowered within haloes and

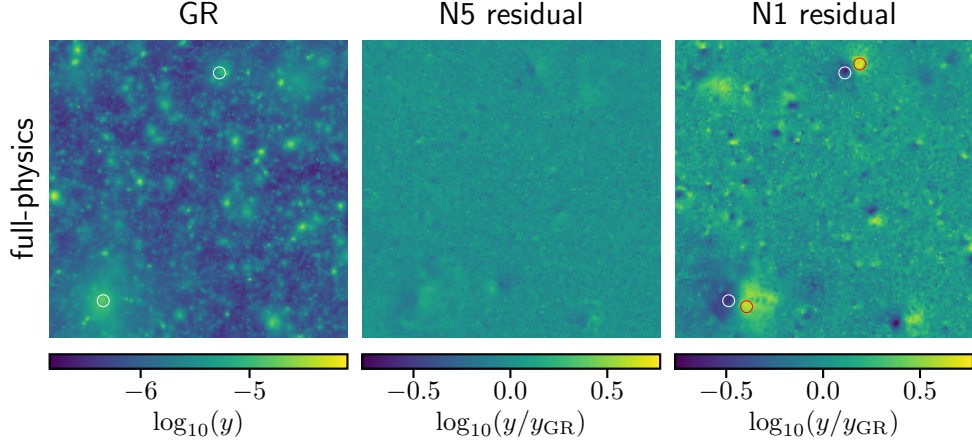


Figure 7.3: Map of the thermal SZ effect in GR (*left column*) and its relative enhancement in N5 (*middle column*) and N1 (*right column*). The maps have a side-length of  $1^\circ$  and a  $512 \times 512$ -pixel resolution, and have been constructed from the SHYBONE simulations (see Sec. 7.2). The  $y$ -parameter is computed for each pixel using Eq. (7.4). The rings indicate two haloes whose positions are shifted in N1 (*red*) relative to GR (*white*).

raised outside haloes. For the kSZ results, as for the  $f(R)$  gravity residuals, the value of  $b$  is increased and reduced seemingly at random, owing to the unpredictable effects of the full-physics processes on the motion of the gas.

The nDGP tSZ maps for the same light cone are shown in Fig. 7.3, where recall that we do not have non-radiative runs. Again, the fifth force causes a shift in halo positions with respect to GR, and this is clearly visible for both nDGP models. The effect is greater in the N1 model, which is a stronger modification of GR than N5. We do not show the kSZ maps for nDGP, since these appear similar to the  $f(R)$  maps and do not offer extra information.

## 7.3 Results

This section gives the main results of this chapter. In Sec. 7.3.1, we analyse the effects of baryonic processes and the fifth force on the stacked electron pressure profiles of FOF groups from our simulations. Then, in Sec. 7.3.2, we discuss the effects on the tSZ and kSZ angular power spectra. Finally, in Sec. 7.3.3, we examine the effects on the power spectrum of the transverse component of the electron momentum field, which is closely related to the

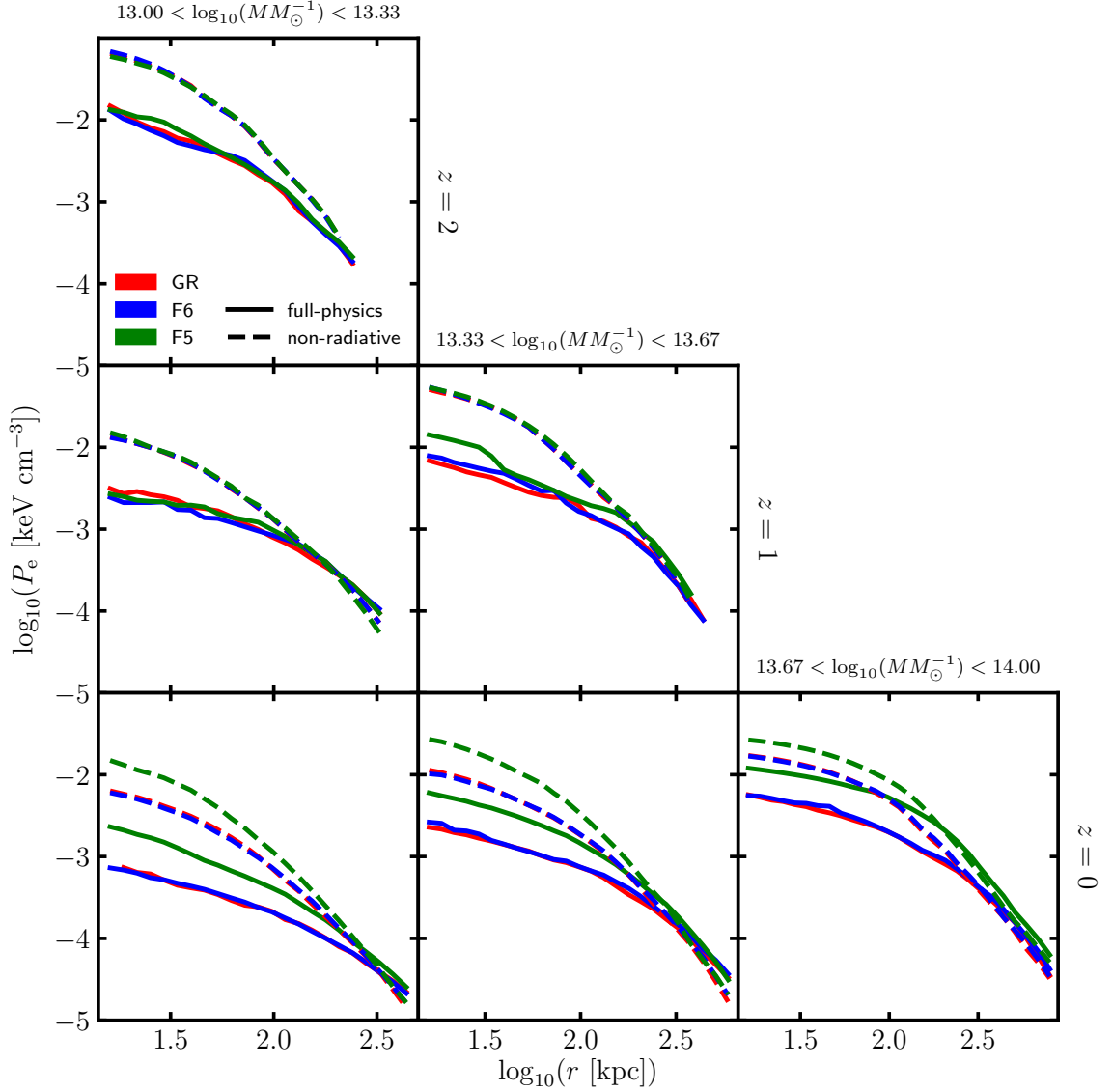


Figure 7.4: Stacked electron pressure profiles for haloes from three mass bins in the range  $10^{13} M_{\odot} < M_{500} < 10^{14} M_{\odot}$  and redshifts 0, 1 and 2. The haloes have been identified from the SHYBONE simulations (see Sec. 7.2), and have been generated for the GR (red), F6 (blue) and F5 (green) gravity models, and for both the full-physics (solid lines) and non-radiative (dashed lines) hydrodynamics schemes.

kSZ angular power spectrum.

### 7.3.1 Electron pressure profiles

We show the stacked electron pressure profiles at  $z = 0$ ,  $z = 1$  and  $z = 2$  in Fig. 7.4 and Fig. 7.5 for  $f(R)$  gravity and nDGP, respectively. Three equally spaced logarithmic

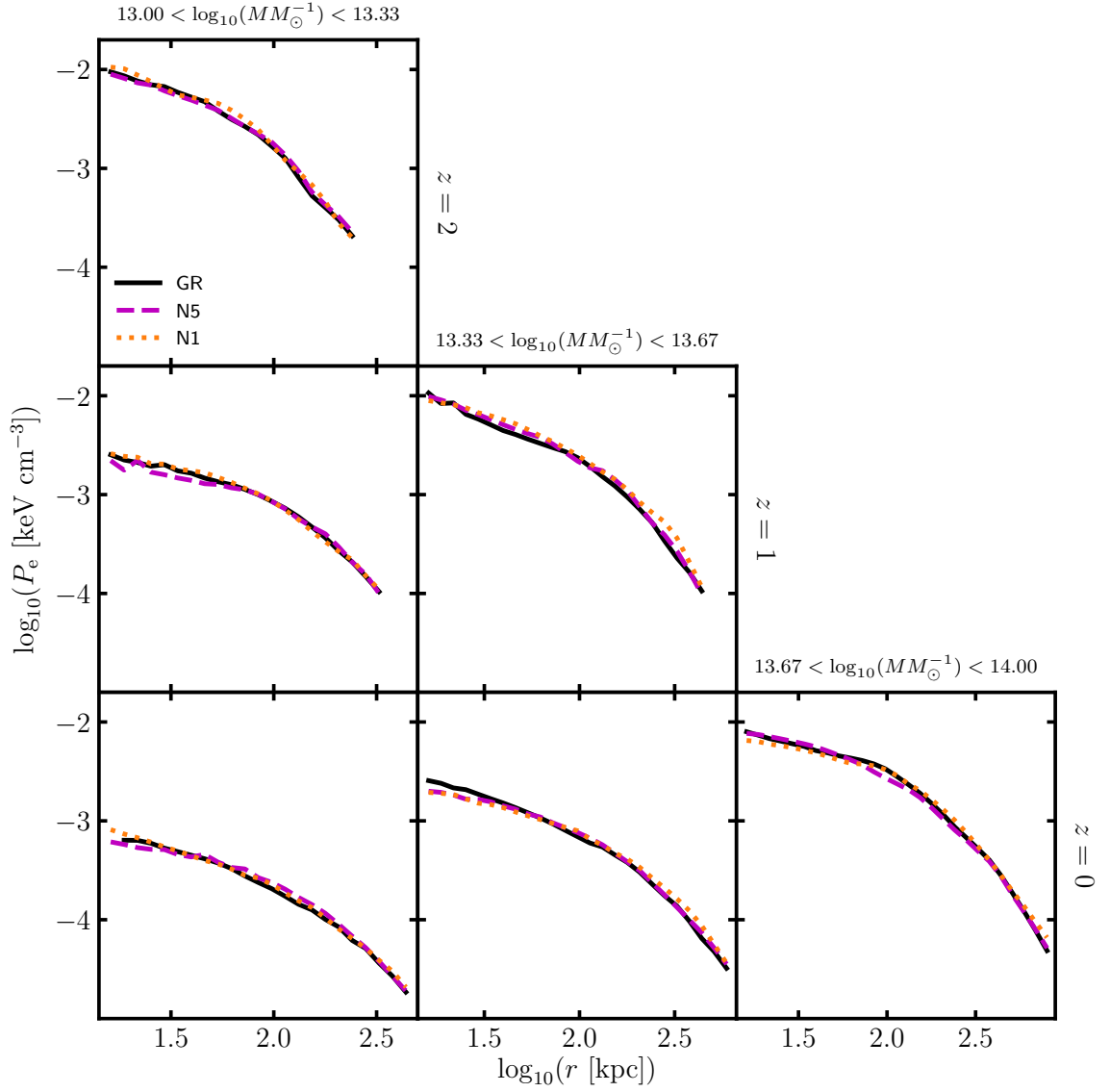


Figure 7.5: Stacked electron pressure profiles for haloes from three mass bins in the range  $10^{13}M_{\odot} < M_{500} < 10^{14}M_{\odot}$  and redshifts 0, 1 and 2. The haloes have been identified from the full-physics SHYBONE simulations (see Sec. 7.2), and have been generated for the GR (black solid), N5 (magenta dashed) and N1 (orange dotted) gravity models.

mass bins, which span the range  $13 < \log_{10}(M_{500}M_{\odot}^{-1}) < 14$ , are considered. The volume-weighted electron pressure is measured in radial bins for each halo. This is given by the following:

$$\bar{P}_e = \frac{\sum_i P_{e,i} V_i}{\sum_i V_i}, \quad (7.9)$$

where  $P_{e,i}$  and  $V_i$  are the electron pressure and volume of gas cell  $i$ , and the summations are evaluated over all gas cells whose centres of mass are within the current bin. The median profile is measured for each radial bin using the haloes enclosed in each mass bin, and is displayed in the figures. Because of the limitations in the size of the halo population at higher masses, only the lowest-mass bin is shown at  $z = 2$  and the highest-mass bin is not shown for  $z = 1$ . For the highest-mass bin shown at each redshift, because the halo number is relatively small, some haloes are also excluded from each model to ensure that the same halo population is used in all models. Any small difference in population could otherwise have a significant effect in these bins, which contain only  $\sim 10$  haloes each. This consideration is not required for the other bins, which have  $\gtrsim 25$  haloes each. We are also unable to include data at  $M_{500} > 10^{14}M_{\odot}$ , for which there are only a few haloes for each model.

For haloes in F5 at sufficiently low redshift, we find that the fifth force boosts the electron pressure. This is caused by the increase in the temperature of the intra-cluster gas, which results from the deepened gravitational potential well (see Chapter 4). This indicates that the tSZ signal from individual haloes is expected to be significantly enhanced in F5. The magnitude of the background scalar field,  $|f_R|$ , increases with time, and as a result the chameleon mechanism is more efficient at screening the fifth force at earlier times. This explains why the enhancement of the pressure in F5 vanishes for  $z \gtrsim 1$ . On the other hand, the background scalar field in the F6 model is 10 times weaker than in F5, and as a result the fifth force is efficiently screened within group- and cluster-sized haloes even at  $z = 0$ .

Meanwhile, Fig. 7.5 shows much smaller differences between the electron pressure profiles in GR and nDGP than in Fig. 7.4, especially at lower  $z$  ( $z \lesssim 1$ ). This is because the Vainshtein mechanism is much more efficient than the chameleon mechanism at screening out the fifth force within haloes at low redshifts — for the latter, depending on the value



of  $|f_{R0}|$  in the two  $f(R)$  models studied here, group-sized objects could be partially or completely unscreened at low  $z$ , while for the former the screening efficiency is similar for haloes of different masses (see, e.g., Fig. 8 of [Hernández-Aguayo et al., 2020](#)), including the ones as small as  $\sim 10^{11.7} h^{-1} M_{\odot}$ , with the fifth force always being strongly suppressed in the inner regions of haloes, at all redshifts.

By comparing the non-radiative and full-physics data in Fig. 7.4, we can see that the additional baryonic processes that are present in the latter act to suppress the pressure at the inner halo regions. This can be caused by, for example, the blowing out of gas by black hole feedback which lowers the density of electrons. Note that, while the electron pressure profiles differ significantly between the full-physics and non-radiative runs, the relative enhancement of F5 with respect to GR seems to be consistent in both cases.

### 7.3.2 tSZ and kSZ power spectra

We have used our SZ maps (see Sec. 7.2.2) to generate the tSZ and kSZ angular power spectra for the  $f(R)$  and nDGP models. The power has been measured for each of the 14 maps in bins of the angular wavenumber  $l$ . For each bin, the mean power and the mean relative difference in the power between gravity models and hydrodynamics schemes have been measured.

From the  $f(R)$  gravity results, shown in Fig. 7.6, we find that the fifth force and the extra baryonic processes that are found in the full-physics simulations have different effects: the middle column shows that, for the non-radiative data, the tSZ and kSZ power spectra are both enhanced in  $f(R)$  gravity relative to GR; and the right column shows that the power is suppressed in the full-physics runs relative to the non-radiative runs, particularly at smaller scales. The latter is consistent with literature: [McCarthy et al. \(2014\)](#) showed that, at scales  $l \gtrsim 1000$ , the tSZ power is suppressed by the ejection of gas by black hole feedback; and [Park et al. \(2018\)](#) found that the kSZ power is suppressed by both the locking away of free electrons in stars, black holes and neutral gas (at all scales), and the ejection of gas through black hole feedback (at smaller scales). For our data, this suppression by baryonic processes occurs at  $l \gtrsim 3000$  for the tSZ power and at  $l \gtrsim 500$  for the kSZ power.

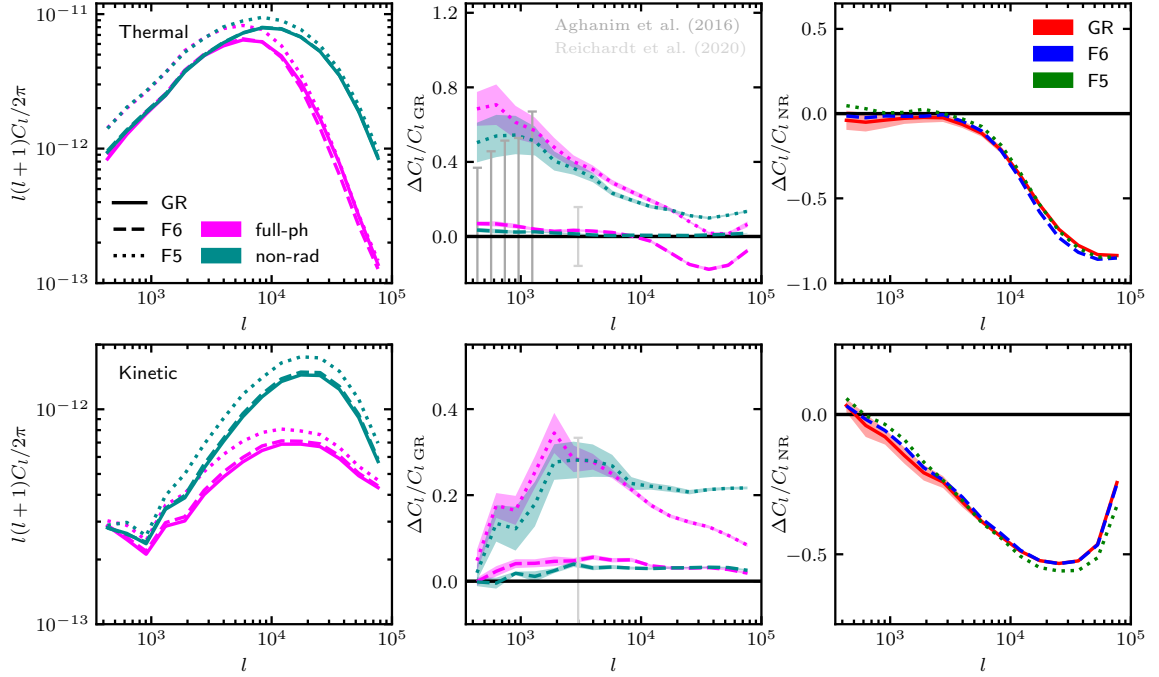


Figure 7.6: Angular power spectra and their relative differences plotted as a function of the angular wavenumber. The data has been generated from maps of the thermal (*top row*) and kinetic (*bottom row*) SZ signals, which have been created using the SHYBONE simulations (see Sec. 7.2). *Left column*: mean angular power spectrum plotted for GR (*solid lines*), F6 (*dashed lines*) and F5 (*dotted lines*), including data from the full-physics (*magenta*) and non-radiative (*cyan*) simulations. *Middle column*: mean relative enhancement of the F6 (*dashed lines*) and F5 (*dotted lines*) angular power spectra with respect to GR, plotted for the full-physics (*magenta*) and non-radiative (*cyan*) simulations. The standard error of the mean is indicated by the shaded regions. The error bars indicate the precision of the latest observations from the Planck (Aghanim et al., 2016) and SPT (Reichardt et al., 2020) collaborations. *Right column*: mean relative enhancement of the full-physics angular power spectra with respect to the non-radiative data, plotted for GR (*red*), F6 (*blue*) and F5 (*green*). For clarity, the standard error is shown for GR only.

The shape and amplitude of this suppression is very similar for each gravity model, as shown in the right column: the tSZ power is suppressed by up to  $\sim 85\%$  and the kSZ power is suppressed by up to  $\sim 50\%$ .

With the extra baryonic processes of cooling, star formation and feedback absent, the tSZ power is enhanced by the fifth force on all scales. The enhancement is greater in F5 than in F6, with peaks of  $\sim 50\%$  and a few percent, respectively, at  $l < 1000$ . However, due to the relatively small size of the fields of view in our light cones, we cannot measure the angular power spectra at  $l \lesssim 500$ , and so it is unclear what the asymptotic behaviour



at large angular scales is, for which future works with larger simulations are needed. For the kSZ power, a roughly constant enhancement is observed (of  $\sim 22\%$  for F5 and  $\sim 3\%$  for F6) at scales  $l \gtrsim 3000$ , with a downturn at larger scales ( $l \simeq 2000$ ). The presence of the fifth force speeds up the formation of large-scale structures, boosting the abundance and peculiar velocity of groups and clusters of galaxies and, in turn, the tSZ and kSZ power spectra. In addition to this, the electron pressure profiles of haloes at a given mass are also enhanced, as discussed in Sec. 7.3.1, which could further boost the tSZ signal and tSZ power spectrum at small angular scales (the relation between the latter and halo electron pressure profiles, however, is more complicated, as we will discuss toward the end of this subsection).

The enhancement of the kSZ power by  $\sim 22\%$  in F5 is higher than predicted by [Bianchini and Silvestri \(2016\)](#) and [Roncarelli et al. \(2018\)](#), who estimated an enhancement of about 15% for the same model using analytical predictions and hydrodynamical simulations, respectively. We remark that our results use only the redshift range  $z \leq 3$  while these works used redshifts up to 9.9 and 15, including the epoch of reionisation which can have a substantial contribution to the total kSZ power. The fifth force is expected to be screened for  $z \gtrsim 3$ , which can explain why the kSZ signal (an integral over the redshift range) shows less deviation from GR in these works. Our smaller redshift range  $z \leq 3$  also explains why the amplitude of the kSZ power in Fig. 7.6 is lower than is typically predicted in literature (e.g., [Park et al., 2018](#)).

The SHYBONE simulations are the first to simultaneously compute the fifth force of  $f(R)$  gravity while incorporating full baryonic physics. The interplay between these two competing mechanisms in the full-physics simulations is therefore of particular interest. According to the middle column of Fig. 7.6, the extra processes in the full-physics simulations have a non-negligible effect on the relative difference between  $f(R)$  gravity and GR. For the tSZ power, a suppression of the  $f(R)$  enhancement is observed at very small scales ( $l \gtrsim 10000$ ), such that the F5 power is brought close to the GR power, and the F6 power becomes  $\sim 20\%$  lower than GR. For the kSZ power, the F5 enhancement is again suppressed at these scales, while there appears to be little change for F6.

We note that these results are likely to be sensitive to the choice of full-physics parameters implemented by SHYBONE. Given that feedback is not fully understood theoretically or from observations, there is a non-negligible uncertainty in the results at small scales. In order to avoid potentially biased results, constraints should instead be made using large scales where the details of baryonic processes are not as prominent. For the tSZ power, these scales correspond to  $l \lesssim 3000$ , although we note that even this range could be sensitive to the full-physics parameters. Our simulations predict enhancements of  $\sim 40\text{-}70\%$  in F5 and less than  $10\%$  in F6, relative to GR, at these scales. For the kSZ power spectrum, again star formation, feedback and cooling appear to have a non-negligible effect at all of the scales tested here. However, the model differences between  $f(R)$  gravity and GR do not differ significantly in the non-radiative and full-physics simulations for scales  $l \lesssim 10^4$ . In this scale range, we observe relative differences of up to  $\sim 35\%$  and  $\sim 5\%$  between GR and the F5 and F6 models, respectively. Note that the non-radiative runs could be considered as an extreme case of the hydrodynamics scheme, with the most interesting physical processes neglected, and for this reason we expect that slight variations of the baryonic model should produce milder differences from the IllustrisTNG model than what is observed between the full-physics and non-radiative curves here. Also note that, due to the small box size, the full-physics runs used in this thesis could suffer from significant sample variance, e.g., due to a few large haloes experiencing unusually strong feedback in one model and not another; again, having a large simulation box in the future will help to address this question.

The tSZ and kSZ power spectra for the nDGP model (for full-physics only), are shown in Fig. 7.7. As for the  $f(R)$  model, the fifth force of nDGP enhances the power on all probed scales: the tSZ power is enhanced by up to  $\sim 40\%$  in N1 and less than  $10\%$  in N5; and the kSZ power is enhanced by up to  $\sim 35\%$  in N1 and  $\sim 5\%$  in N5. However, given the absence of a non-radiative simulation for the nDGP model, we note that it is possible that these differences could be sensitive to baryonic physics, as in  $f(R)$  gravity.

Interestingly, the tSZ power spectrum at high  $l$  is significantly enhanced in N1, even though the pressure profiles (Fig. 7.5) do not appear to show a clear deviation from GR. There are a few reasons why this can happen. First of all, the tSZ power receives contribu-

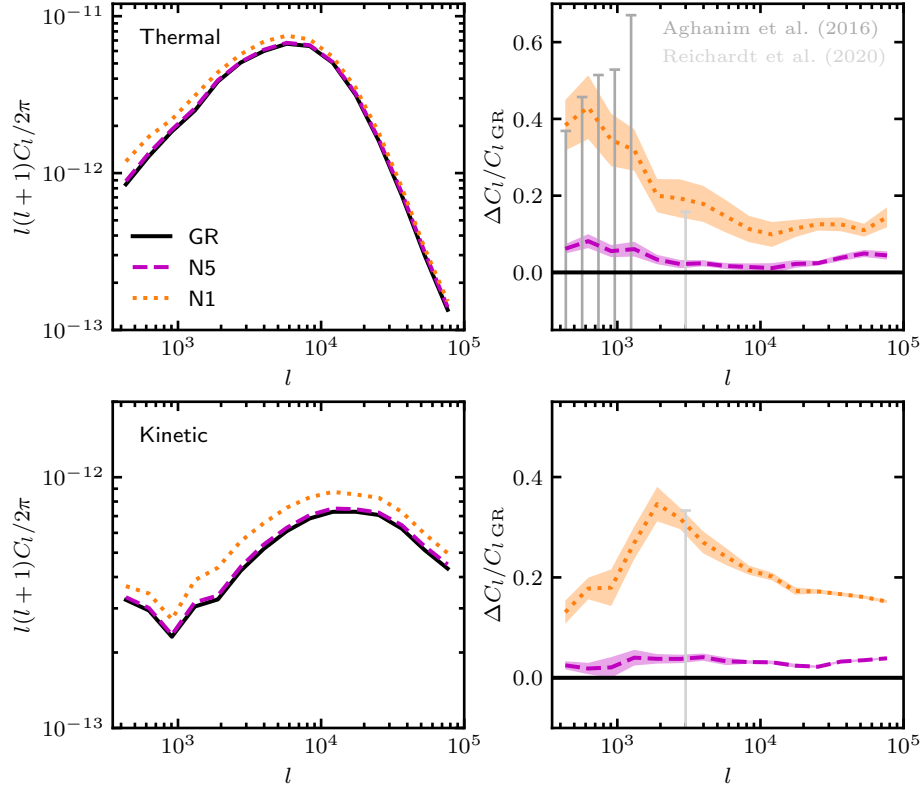


Figure 7.7: Angular power spectra and their relative differences plotted as a function of the angular wavenumber. The data has been generated from maps of the thermal (*top row*) and kinetic (*bottom row*) SZ signals, which have been created using the SHYBONE simulations (see Sec. 7.2). *Left column*: mean angular power spectrum plotted for GR (*solid lines*), N5 (*dashed lines*) and N1 (*dotted lines*). *Right column*: mean relative enhancement of the N5 and N1 angular power spectra with respect to GR. The standard error of the mean is indicated by the shaded regions. The error bars indicate the precision of the latest observations from the Planck (Aghanim et al., 2016) and SPT (Reichardt et al., 2020) collaborations.

tions from outside haloes as well as from within. The fourth column of Fig. 7.1 indicates that outside haloes the tSZ signal can be boosted by the ejection of gas by feedback. The presence of the fifth force is expected to result in the feedback being triggered earlier, which can cause the power to be enhanced relative to GR at angular scales corresponding roughly to halo sizes\*. Secondly, smaller angular scales receive a greater contribution from higher redshifts (see, e.g., McCarthy et al., 2014). In F5, the fifth force is efficiently screened for  $z \gtrsim 1$ , but in N1 it can still reach a few percent of the strength of the Newtonian force at

\*The fifth force also enhances matter clustering on large scales overall, and this is expected to be reflected in the clustering of free electrons.

the radius  $R_{200}$  and  $\sim 8\%$  outside it at  $z \sim 2$  (e.g., [Hernández-Aguayo et al., 2020](#)). This means that the SZ power at high  $l$  can be enhanced by a greater amount in N1 than in F5. In fact, for the mass bin shown at  $z = 2$ , the N1 pressure profile is enhanced by  $\sim 9\%$  with respect to GR, and we also find that the nDGP pressure is enhanced in lower-mass bins which are not shown in Fig. 7.5. We have further verified (though not shown here) that at  $z \gtrsim 1.5$ , the 3D electron pressure power spectrum is significantly enhanced even at high  $k$  values well inside the 1-halo regime.

The results discussed in this section indicate that the tSZ and kSZ power have the potential to effectively probe  $f(R)$  gravity and nDGP at large scales. To demonstrate this, we have included error bars in Figs. 7.6 and 7.7 to indicate the uncertainties of the latest tSZ and kSZ observations from the Planck ([Aghanim et al., 2016](#)) and SPT (SPT, [Reichardt et al., 2020](#)) collaborations. The  $\sim 16\%$  precision of the tSZ measurement by SPT is sufficient to distinguish the F5 model from GR at  $l = 3000$ , while the Planck measurements have sufficient precision to distinguish F5 at large angular scales ( $l \lesssim 500$ ). The 33% precision of the kSZ measurement by SPT has a similar magnitude to the relative enhancements of the F5 and N1 models with respect to GR, indicating that more precise measurements from future surveys will be capable of ruling out these models. However, in order to avoid bias from uncertain baryonic physics, it will be necessary to use a range of full-physics parameters to confirm that reliable constraints can be achieved at these angular scales. It will also be important to revisit this study using simulations with a greater box size that can accurately probe the tSZ power up to angular scales  $l \sim 100$ , where the precision of the Planck measurements is particularly high ([Aghanim et al., 2016](#)). Finally, understanding the degeneracies between MG and variations in other cosmological parameters is also critical in order to have unbiased constraints.

Before finishing this section, let us note that, despite the qualitative difference in their respective screening mechanisms – Vainshtein screening is always efficient inside dark matter haloes while the same cannot be said about the chameleon mechanism (cf. Figs. 7.4 and 7.5) – the enhancements of both the tSZ and kSZ power spectra are very similar in these two models.

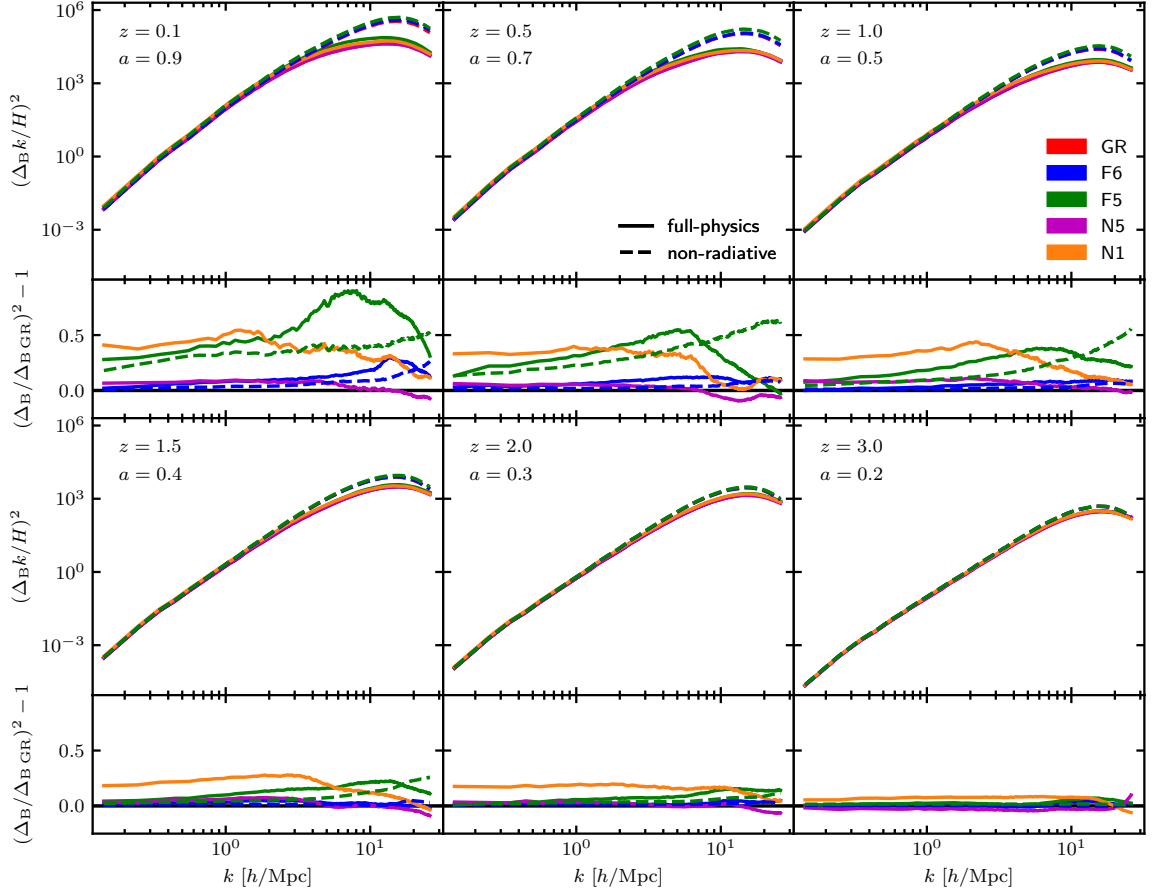


Figure 7.8: Power spectrum of the transverse component of the electron momentum plotted against the wavenumber. The data has been generated for six different redshifts (annotated) using the SHYBONE simulations (see Sec. 7.2.1) for both the non-radiative (*dashed lines*) and full-physics (*solid lines*) hydrodynamics schemes. In addition to GR (*red lines*), data is plotted for F6 (*blue lines*), F5 (*green lines*), N5 (*magenta lines*) and N1 (*orange lines*). The lower sub-panels show the relative enhancement of the MG (F6, F5, N5 and N1) power spectra with respect to GR.

### 7.3.3 Transverse momentum power spectrum

In order to understand the (similar) effects of  $f(R)$  gravity and nDGP on the kSZ power in more detail, we have measured the power spectrum of the transverse component of the electron momentum field which, in the small-angle limit, can be related to the kSZ angular power spectrum using the Limber approximation (e.g., [Shaw et al., 2012](#)):

$$C_l^{\text{kSZ}} = \frac{8\pi^2}{(2l+1)^3} \left( \frac{\sigma_T \bar{\rho}_{\text{gas},0}}{\mu_e m_p} \right)^2 \times \int_0^{z_{\text{re}}} \frac{dz}{c} (1+z)^4 \chi^2 \Delta_B^2(k, z) e^{-2\tau(z)} \frac{x(z)}{H(z)}, \quad (7.10)$$

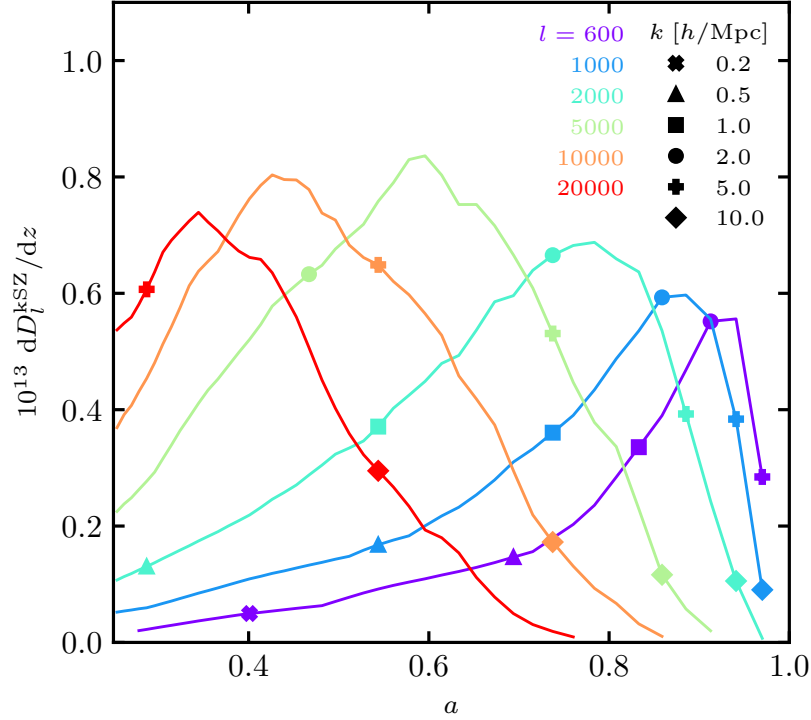


Figure 7.9: Derivative of the GR kSZ angular power spectrum as a function of the cosmological scale factor  $a$  for six different values of the angular wavenumber  $l$ . The values are computed from the full-physics SHYBONE simulation (see Sec. 7.2.1) using the Limber approximation (Eq. 7.10). The scales that are spanned by each  $l$  value are indicated by markers which represent unique values of the wavenumber  $k$ .

where  $\bar{\rho}_{\text{gas},0}$  is the present-day mean background gas density,  $\mu_e m_p$  is the mean gas mass per electron,  $z_{\text{re}}$  is the redshift at the epoch of reionisation,  $\chi$  is the fraction of electrons that are ionised,  $k = l/x$  is the wavenumber,  $x(z) = \int_0^z (cdz'/H(z'))$  is the comoving distance at redshift  $z$ , and the optical depth,  $\tau$ , is given by:

$$\tau(z) = \sigma_{\text{TC}} \int_0^z dz' \frac{\bar{n}_e(z')}{(1+z')H(z')}. \quad (7.11)$$

We have computed the transverse momentum power,  $\Delta_{\text{B}}^2(k, z)$ , using the electron momentum field  $\mathbf{q}$  for a sample of snapshots from our simulations. This is defined  $\mathbf{q} = \mathbf{v}(1 + \delta) = \mathbf{v}(n_e/\bar{n}_e)$ , where  $\mathbf{v}$  is the velocity field of the gas. The power spectrum of the transverse momentum component,  $\mathbf{q}_{\perp}$ , is related to the power spectrum of the curl of the momentum field,  $\nabla \times \mathbf{q}$ , by  $P_{q_{\perp}} = P_{\nabla \times q}/k^2$ , and can be converted to the more commonly used definition  $\Delta_{\text{B}}^2 = P_{q_{\perp}} k^3/(2\pi^2)$ .

In Fig. 7.8, we show the dimensionless quantity  $(\Delta_{\text{B}} k/H)^2$  at six different redshifts

for all gravity models and hydrodynamics schemes. In the lower sub-plots at each redshift, we show the MG enhancements of  $\Delta_{\text{B}}^2$  with respect to GR. For the non-radiative data, we see that the  $f(R)$  enhancement is always increasing from large to small scales. The lowered enhancement at large scales is caused by the limited range of the fifth force, which is set by the Compton wavelength (Eq. (1.10)). For the full-physics data, the  $f(R)$  enhancement follows a similar pattern at large scales but drops off for smaller scales ( $k \gtrsim 5h/\text{Mpc}$ ) where baryonic processes are particularly prominent. For the nDGP data, the enhancement is roughly constant at large scales, since the fifth force in this model is long-range in the linear regime. There is again a suppression at small scales, which is likely to be caused by Vainshtein screening but could also be related to baryonic physics (we do not have non-radiative simulations to confirm the latter). For both  $f(R)$  gravity and nDGP, the enhancement vanishes at higher redshifts where the amplitude of the scalar field is lower and the fifth force is screened out.

We also show, in Fig. 7.9, the derivative  $dD_l^{\text{kSZ}}/dz$ , where

$$D_l^{\text{kSZ}} = l(l+1)C_l^{\text{kSZ}}/(2\pi). \quad (7.12)$$

This has been computed using the integrand and pre-factors in Eq. (7.10), and indicates the cosmic times and range of  $k$ -modes that have the greatest contribution to the kSZ angular power spectrum for different angular scales  $l$ . The enhancement of the kSZ power in F5 and N1 is observed to peak at  $l = 2000$  in Figs. 7.6 and 7.7. From Fig. 7.9, we see that  $D_{l=2000}^{\text{kSZ}}$  receives a significant contribution from times  $0.4 \lesssim a \lesssim 0.9$  and scales  $1h/\text{Mpc} \lesssim k \lesssim 5h/\text{Mpc}$ . At these scales, the enhancement of  $\Delta_{\text{B}}^2$  peaks for N1 and has a similar magnitude for F5. The enhancements in these models span  $\sim 30\%$ - $60\%$  over these scales and times, which is consistent with the peak enhancement of  $C_l^{\text{kSZ}}$ . Going to larger angular scales ( $600 < l < 1000$ ),  $D_l^{\text{kSZ}}$  is affected by lower  $k$ -modes (down to  $\sim 0.2h/\text{Mpc}$ ) and lower redshifts. This then probes the larger scales (in Fig. 7.8) where the F5 fifth force is suppressed and the N1 enhancement levels off. This is consistent with Figs. 7.6 and 7.7, where the kSZ power appears to be suppressed by a greater amount in F5 than in N1. Interestingly, this also implies that the enhancement of  $C_l^{\text{kSZ}}$  could be constant at angular scales larger than those available from our mock SZ maps. At smaller angular scales

( $10000 \lesssim l \lesssim 20000$ ),  $D_l^{\text{kSZ}}$  receives a significant contribution from high- $k$  modes ( $2h/\text{Mpc} \lesssim k \lesssim 10h/\text{Mpc}$ ), where Vainshtein screening suppresses the nDGP fifth force and the  $f(R)$  fifth force is suppressed by baryonic processes (for the full-physics runs). In addition to this,  $D_l^{\text{kSZ}}$  is probing earlier times  $0.25 \leq a \lesssim 0.6$ , where the scalar field amplitude is reduced in both models. This is therefore consistent with the lowered enhancement of  $C_l^{\text{kSZ}}$  at these angular scales. Note that the non-radiative  $f(R)$  runs produce a higher kSZ power at high  $l$  than the full-physics runs, which also agrees with the observation in Fig. 7.8 that at large  $k$  and  $a \lesssim 0.6$  the former has a larger transverse-momentum power spectrum.

The amplitude of  $\Delta_{\text{B}}k/H$  in Fig. 7.8 appears to agree reasonably well with literature results (e.g., Zhang et al., 2004; Shaw et al., 2012; Bianchini and Silvestri, 2016), although it is slightly lower at large scales. We note that this is likely because the relatively small size  $62h^{-1}\text{Mpc}$  of our simulation box misses off longer-wavelength modes. It will therefore be useful to revisit this study with a larger box. The inclusion of longer-wavelength modes is expected to further suppress the F5 enhancement of  $\Delta_{\text{B}}^2$  at low- $k$  and to have little effect on the N1 enhancement.

We also note that for the entire  $l$  range studied in Figs. 7.6 and 7.7 the kSZ power spectrum is dominated by modes with  $k \gtrsim 0.2h/\text{Mpc}$  in the transverse-momentum power spectrum. From Fig. 7.8, we can see that in this regime galaxy formation has a non-negligible impact on  $\Delta_{\text{B}}$ , which means that uncertainties in the subgrid physics can be an important theoretical systematic effect in using the kSZ power to test gravity models. Using kSZ data at  $l < 600$  may help reduce this effect, but the current simulation size does not allow a study of that range of  $l$ .

## 7.4 Summary, Discussion and Conclusions

Over the past couple of decades, great advances have been made in the measurement of the secondary anisotropies of the CMB caused by the SZ effect, including its thermal component and even its much smaller kinematic component. The angular power spectrum of the tSZ effect has been increasingly adopted as a probe of cosmological parameters that influence the growth of large-scale structures. Also, as observations of the kSZ power spectrum



continue to improve, the latter has been identified as another potentially powerful probe of cosmology. The next generation of ground-based observatories (Ade et al., 2019; Abazajian et al., 2016) look set to revolutionise the constraining power of these probes.

In this chapter, we have looked at the viability of using the angular power spectra of the tSZ and kSZ effects as large-scale probes of HS  $f(R)$  gravity and nDGP, which are representative of a wide-range of MG theories which exhibit the chameleon and Vainshtein screening mechanisms, respectively. In order to do so, we have made use of the SHYBONE simulations (cf. Sec. 7.2.1), which are the first cosmological simulations that simultaneously incorporate full-physics plus HS  $f(R)$  gravity (Arnold et al., 2019b) and nDGP (Hernández-Aguayo et al., 2020). The simulations employ the IllustrisTNG full-physics model, which incorporates calibrated sub-resolution recipes for star formation and cooling as well as stellar and black hole feedback and allows realistic galaxy populations to be produced in hydrodynamical simulations.

Using these simulations, we have generated mock maps of the tSZ and kSZ signals (Sec. 7.2.2), and used these maps to measure the angular power spectra. Our results (Figs. 7.6 and 7.7) indicate that the fifth force, present in  $f(R)$  gravity and nDGP, and the subgrid baryonic physics have different effects on the tSZ and kSZ power spectra. The former enhances the power on all scales probed by our maps ( $500 \lesssim l \lesssim 8 \times 10^4$ ) by boosting the abundance and peculiar velocity of large-scale structures (e.g., dark matter haloes and free electrons inside them), while the latter brings about a suppression on scales  $l \gtrsim 3000$  for the tSZ effect and on all tested scales for the kSZ effect. Even with both of these effects present, we find that the power can be significantly enhanced in  $f(R)$  gravity and nDGP: by up to 60% for the tSZ effect and 35% for the kSZ effect for the F5 and N1 models; and by 5%-10% for F6 and N5, which correspond to relatively weak modifications of GR. In addition, we have computed the power spectrum of the transverse component of the electron momentum field (Sec. 7.3.3), which is closely related to the kSZ angular power spectrum. In particular, we show in Fig. 7.9 that at angular sizes  $l \geq 600$  the kSZ signal is dominantly contributed by  $k$ -modes in the transverse-momentum power spectrum which are in the non-linear regime, and which are affected strongly by MG. The  $k$ -modes in the

linear regime may contribute more to smaller  $l$ , but at least for  $f(R)$  gravity the impact of MG at those  $l$  values will be much less significant due to the finite range of the fifth force, as we can already see in Fig. 7.6.

We find that the relative difference between the MG models and GR is significantly affected by the additional baryonic processes that act in the full-physics simulations. Given that these processes are still relatively less well-constrained, this adds to the uncertainty in our theoretical predictions of the kSZ angular power spectra on small angular scales, e.g.,  $l > 600$ . Therefore, further work should be carried out using a range of full-physics parameters to precisely identify the scales on which constraints can be reliably made before the tSZ and kSZ power are used to probe  $f(R)$  gravity and nDGP.

Finally, we note that the reason we are unable to study larger scales is the relatively small box size of the SHYBONE simulations. In Chapter 8, we will present a re-calibrated full-physics model that can be used to run larger simulations, which can potentially be used to analyse larger scales in a future work.

# Realistic galaxy formation simulations to study clusters in modified gravity

## 8.1 Introduction

In Chapters 4 and 6, we used full-physics simulations to study the observable-mass scaling relations of galaxy groups and clusters in  $f(R)$  gravity and nDGP. These simulations, which are from the SHYBONE suite (Arnold et al., 2019b; Hernández-Aguayo et al., 2020), have a box size of  $62h^{-1}\text{Mpc}$  (‘L62’) and a high mass resolution. This is more suited for studying galaxies than galaxy clusters; indeed, these simulations contain just  $\sim 100$  galaxy group-sized objects and only 5-10 cluster-sized objects, with no objects above mass  $M_{500} \sim 10^{14.5} M_{\odot}$ . The L62 predictions of the cluster scaling relations may therefore suffer from poor statistics and be potentially subjected to a significant influence by sample variance. Unfortunately, high-resolution simulations which incorporate both screened modified gravity and full baryonic physics are very expensive to run for larger cosmological volumes, which has made it difficult to study the interplay between baryonic physics and the fifth force at higher masses.

The SHYBONE simulations make use of the IllustrisTNG baryonic physics model (Weinberger et al., 2017; Pillepich et al., 2018c), which can be used to generate galaxy populations whose stellar and gaseous properties closely match observations. However, it has been shown

that a sufficiently high mass resolution is required to achieve this level of agreement (e.g., [Pillepich et al., 2018c](#)). For example, a lowered mass resolution means that the gas cells will have larger volumes, resulting in a smoothed out density field which can miss out the high-density peaks where star formation would be highest. The L62 SHYBONE simulations, with  $512^3$  gas cells, have  $\sim 15$  times lower mass resolution than that used to calibrate the TNG model ( $25h^{-1}\text{Mpc}$  box with the same number of gas cells). The gas and stellar properties of haloes from the L62 simulations still give a reasonable agreement with observational data, however the lowered resolution means that there is less star formation, resulting in the amplitudes of the stellar mass fraction, the stellar mass function and the star formation rate density being reduced compared to the fiducial TNG results (for example, see Fig. 4 in [Arnold et al., 2019b](#)).

For this work, running simulations with a substantial number of galaxy clusters (with masses  $10^{14}M_{\odot} \lesssim M_{500} \lesssim 10^{15.5}M_{\odot}$ ) would require a box size of at least  $\sim 300h^{-1}\text{Mpc}$ , which necessitates going to even lower resolutions. In order to make this possible without losing the good agreement with observational data, we have retuned the parameters of the IllustrisTNG model, including parameters which control the density threshold for star formation and the energy released by the stellar and black hole feedback mechanisms. This retuning was a significant undertaking which involved running over 200 simulations with a reduced box size, and our new model can be used to run low-resolution full-physics simulations for both standard gravity and modified gravity scenarios. We will present the model and describe the simulations used to tune it in this chapter. We have used this model to run GR and  $f(R)$  gravity simulations with a significantly increased box size of  $301.75h^{-1}\text{Mpc}$ , and we will present the predictions for the observable-mass scaling relations over an extended mass range  $10^{13}M_{\odot} \leq M_{500} \lesssim 10^{15}M_{\odot}$ .

In Sec. 8.2, we provide a detailed description of the baryonic physics retuning and the large-box simulations, including the agreement with galaxy observations. We then present our results for the observable-mass scaling relations in Sec. 8.3. Finally, we provide a summary of this chapter in Sec. 8.4.

## 8.2 Simulations and methods

In Sec. 8.2.1, we discuss our retuning of the IllustrisTNG model for low-resolution simulations. Our large-box simulations, which are used for the main results of this chapter, are presented in Sec. 8.2.2.

### 8.2.1 Baryonic physics fine-tuning

To recalibrate the baryonic physics at our desired mass resolution, we have run a large number of realisations, using AREPO (Springel, 2010), which have a small box size of  $68h^{-1}\text{Mpc}$  (‘L68-N256’). These have  $256^3$  dark matter particles with mass  $1.35 \times 10^9 h^{-1} M_\odot$  and, initially, the same number of gas cells with mass  $\sim 2.6 \times 10^8 h^{-1} M_\odot$  on average. The calibration was carried out using GR (although, as we will show, the retuned model works equally well for F5) with the same cosmological parameter values as the SHYBONE simulations. The runs were all started from the same set of initial conditions at redshift  $z = 127$ . These have been generated using the N-GENIC code (e.g., Springel et al., 2005), which uses the Zel’dovich approximation to displace an initially homogeneous particle distribution and obtain an initial density field with a prescribed linear power spectrum. Each of the input particles is then split into a dark matter particle and a gas cell, with the ratio of masses set by the values of the cosmic density parameters  $\Omega_M$  and  $\Omega_b$ .

We used the SUBFIND code to locate FOF groups and the bound substructures of each group. By adjusting the baryonic physics of these calibration runs, we have aimed for reasonable agreement with observational data and empirical constraints for the six galaxy properties shown in Fig. 8.1, which were also used to calibrate the IllustrisTNG model (Pillepich et al., 2018c). These are: the stellar mass fraction (FOF groups), with empirical constraints from Behroozi et al. (2013) and Kravtsov et al. (2018); the stellar mass function (subhaloes), with observations from D’Souza et al. (2015), Bernardi et al. (2013), Baldry et al. (2012) and Li and White (2009); the star formation rate density (SFRD) as a function of redshift, with observations from Behroozi et al. (2013); the gas mass fraction (FOF groups), with observations from Lovisari et al. (2015), Gonzalez et al. (2013), Pratt

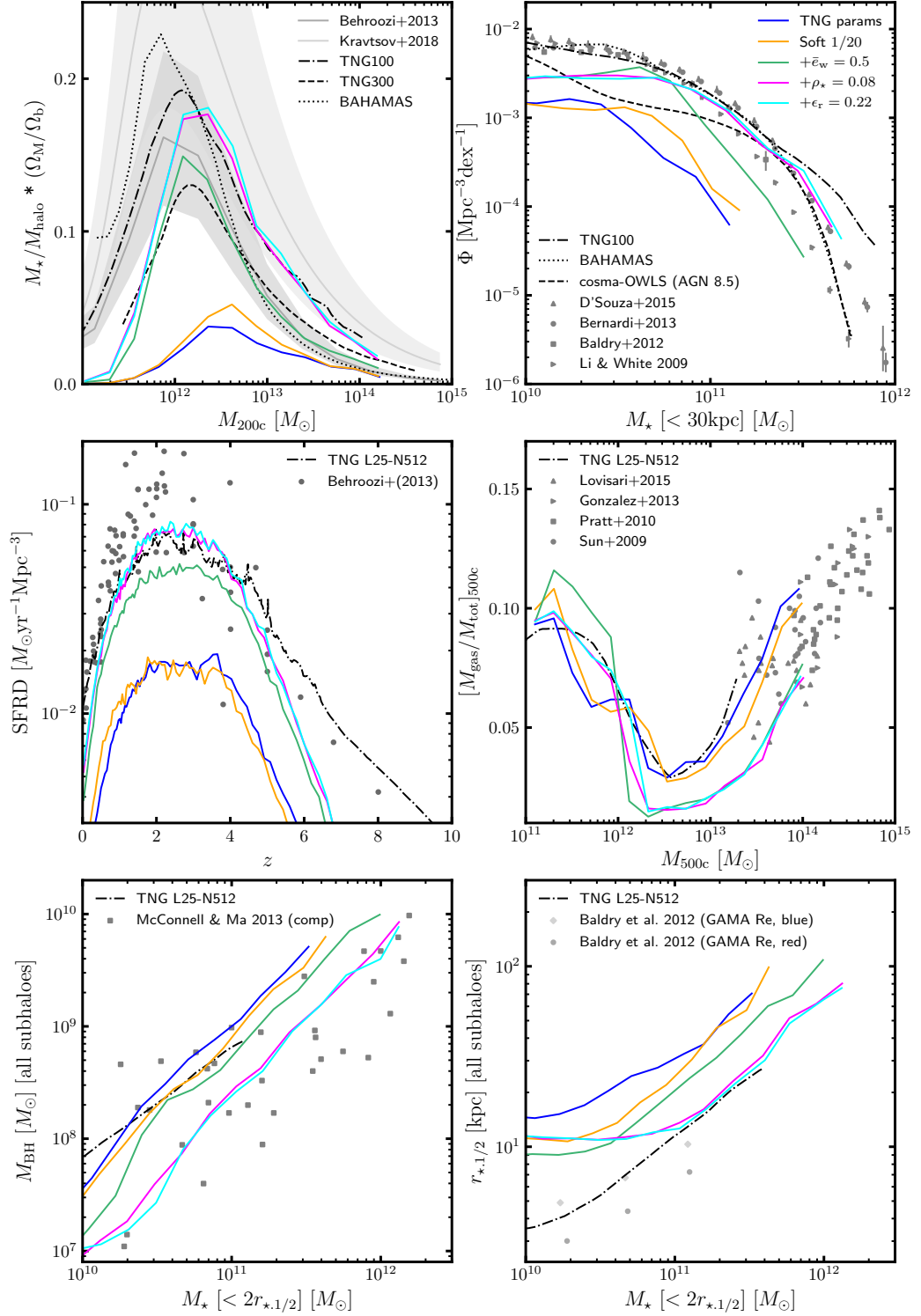


Figure 8.1: Stellar, gas and black hole properties in a sample of the L68-N256 calibration runs (*coloured solid lines*). The properties are (*clockwise from top-left*): stellar mass fraction; stellar mass function; gas mass fraction; stellar half-mass radius; black hole mass; star formation rate density. A selection of results from previous literature are shown as a comparison. See Sec. 8.2.1 for further details.

et al. (2010) and Sun et al. (2009); the black hole mass versus the stellar mass (subhaloes), with the compilation of observations from McConnell and Ma (2013); and the galaxy size versus the stellar mass (subhaloes), with observational data from Baldry et al. (2012). The results for a selection of our calibration runs are represented by the coloured solid lines. Apart from the SFRD, which is a direct output of the simulations, these lines are generated using mass-binning of either FOF groups or subhaloes (see the parentheses above). The black lines show predictions from the TNG simulations (e.g., Nelson et al., 2018; Springel et al., 2018; Marinacci et al., 2018; Pillepich et al., 2018b; Naiman et al., 2018) as well as the BAHAMAS and cosmo-OWLS simulations (McCarthy et al., 2017; Brun et al., 2014).

The dark blue line in Fig. 8.1 shows the predictions using the fiducial TNG model at our lowered resolution. Star formation is significantly reduced at this resolution compared to the fiducial TNG resolution, which is used by the ‘TNG L25-N256’ simulation (‘TNG100’ has a similar resolution, while ‘TNG300’ has  $\sim 5$  times lower resolution). Consequently, the stellar mass fraction, the stellar mass function and the SFRD are significantly lower. The primary objective of our retuning is therefore to achieve a greater amount of star formation in order to obtain a closer match with the observational data. Our changes are described in the sections below, and the effects of these changes are shown in Fig. 8.1. We note that the calibration runs discussed in this section are only a very small subset of the  $\sim 200$  simulations which were run for this calibration study: we provide further details on these simulations and the calibration procedure in Appendix D.

#### 8.2.1.1 *Gravitational softening*

In low-resolution simulations, where the gas cells have higher masses, there is a larger risk of two-body heating: this occurs when two particles come close together and incur a significant gravitational boost, which can raise the internal energy and subsequently the temperature of the gas. We have therefore increased the gravitational softening length to  $1/20$  times the mean inter-particle separation, which is about twice the length used for the SHYBONE simulations. The gravitational force is dampened when gas cells come within this distance, preventing extreme interactions. This change alone causes an overall reduction of the gas

temperature in our simulations, which results in more cool gas that is capable of forming stars: see the orange lines in Fig. 8.1, which have a greater amplitude than the dark blue lines for the stellar mass fraction and stellar mass function.

We also considered fractions of 1/30 and 1/10 for the softening length. For higher masses (e.g.,  $M_{200} \gtrsim 10^{12} M_{\odot}$ ), we observed that using a larger softening length results in greater star formation (for the reasons discussed above). However, we were unable to significantly boost star formation at lower masses; in fact, we observed that a large softening length (for example, a fraction 1/10 of the mean inter-particle separation) can even lead to less star formation at low masses. A potential effect of using a larger softening is that the gravitational potential well of haloes effectively becomes shallower. For low-mass haloes, where the gravitational potential well is already shallower than for high-mass haloes, this could potentially lead to a lower gas density (e.g., the gas is now less gravitationally bound) which in turn could reduce star formation. This is a motivation for using the fraction 1/20, for which we never observed the above effect, rather than using larger fractions.

It is evident from Fig. 8.1 that, while it can increase star formation, changing the gravitational softening length alone is not enough to produce stellar contents that match observational data.

### 8.2.1.2 *Stellar feedback*

In the TNG model, a portion of the gas mass in star-forming gas cells is converted into wind particles which are launched in random directions (Pillepich et al., 2018c). For a star-forming gas cell with metallicity  $Z$ , the available wind energy is:

$$e_w = \bar{e}_w \left[ f_{w,Z} + \frac{1 - f_{w,Z}}{1 + (Z/Z_{w,\text{ref}})^{\gamma_{w,Z}}} \right] \times N_{\text{SNII}} E_{\text{SN11,51}} 10^{51} \text{erg} M_{\odot}^{-1}, \quad (8.1)$$

where  $\bar{e}_w$  is a dimensionless free parameter,  $E_{\text{SNII,51}}$  is the available energy from core-collapse supernovae in units of  $10^{51} \text{erg}$ ,  $N_{\text{SNII}}$  is the number of supernovae per stellar mass that is formed, and  $f_{w,Z}$ ,  $Z_{w,\text{ref}}$  and  $\gamma_{w,Z}$  are additional parameters of the model. A wind particle will eventually donate its thermal energy (along with its mass, momentum, and metal content) to a gas cell that is outside its local dense inter-stellar medium. This heats



the gas and subsequently reduces the efficiency of star formation (gas must be sufficiently cool in order to form stars).

Star formation efficiency is already reduced by our lowered gas cell resolution, therefore reducing the thermal heating of the gas by wind feedback can help to rectify this. We have achieved this in our retuning of the model by reducing the value of  $\bar{e}_w$  from the fiducial TNG value 3.6 to 0.5, which lowers the energy of the winds. As can be seen from the green lines in Fig. 8.1, this change significantly boosts star formation over a wide range of masses. The stellar mass fraction now has a reasonable amplitude for  $M_{200} \gtrsim 10^{12} M_\odot$ , while the amplitudes of the SFRD and stellar mass function are much closer to the observational data.

We have also tried varying the wind speed. This is given by (Pillepich et al., 2018c):

$$v_w = \max \left[ \kappa_w \sigma_{\text{DM}} \left( \frac{H_0}{H(z)} \right)^{1/3}, v_{w,\text{min}} \right], \quad (8.2)$$

where  $\kappa_w$  is a dimensionless factor,  $\sigma_{\text{DM}}$  is the local one-dimensional velocity dispersion of the dark matter particles and  $v_{w,\text{min}}$  is the minimum wind velocity allowed in the model. For our calibration runs, we tried reducing the  $\kappa_w$  and  $v_{w,\text{min}}$  parameters. This reduces the speed of the wind particles, which now take longer to transfer the thermal energy to the surrounding gas. Gas is therefore heated up at a slower rate, resulting in an increased amount of star formation. We found that reducing these parameters has a similar effect to reducing the  $\bar{e}_w$  parameter, with star formation boosted over a wide mass range. However, we could find no clear advantage in varying the wind speed parameters instead of  $\bar{e}_w$ , or in varying all three of these parameters in combination. For simplicity, we therefore decided to adjust the stellar feedback using only the  $\bar{e}_w$  parameter.

### 8.2.1.3 *Star formation model*

The star formation rate in IllustrisTNG is computed for gas cells using the Springel and Hernquist (2003) model. Stars can only be formed by gas cells which exceed a particular density threshold, which is approximately  $n_{\text{H}} \approx 0.1 \text{ cm}^{-3}$ . We will refer to the threshold gas density as  $\rho_\star$  in this work. At our reduced resolution, gas cells have a larger volume and therefore a smoothed density which can miss out high-density peaks in galaxies. In

order to account for this, we have reduced  $\rho_\star$  from  $\approx 0.1\text{cm}^{-3}$  to a fixed value of  $0.08\text{cm}^{-3}$ , allowing gas cells with lower density to form stars.

The effect of making this change, in addition to the changes listed above, is shown by the magenta lines in Fig. 8.1. This further boosts the stellar mass fraction and SFRD, which are now both in good agreement with the TNG100 results for  $M_{200} \gtrsim 10^{12}M_\odot$  and  $z \lesssim 5$ , respectively, and there is also now a good agreement with the D’Souza et al. (2015), Bernardi et al. (2013) and Baldry et al. (2012) observations of the stellar mass function for  $M_\star \gtrsim 10^{11}M_\odot$ .

#### 8.2.1.4 *Black hole feedback*

The TNG model employs two types of black hole feedback, depending on the accretion state of the central supermassive black hole (Weinberger et al., 2017): in the low accretion state, a kinetic feedback model is employed which produces black hole-driven winds; and in the high accretion state, a thermal feedback model is employed which heats up the surrounding gas. The rate of accretion  $\dot{M}$  is set by the Eddington-limited Bondi accretion:

$$\begin{aligned}\dot{M}_{\text{Bondi}} &= \frac{4\pi G^2 M_{\text{BH}}^2 \rho}{c_s^3}, \\ \dot{M}_{\text{Edd}} &= \frac{4\pi G M_{\text{BH}} m_p}{\epsilon_r \sigma_{\text{TC}}}, \\ \dot{M} &= \min(\dot{M}_{\text{Bondi}}, \dot{M}_{\text{Edd}}),\end{aligned}\tag{8.3}$$

where  $M_{\text{BH}}$  is the black hole mass,  $\rho$  represents the ambient density around the black hole,  $c_s$  represents the ambient sound speed and  $\epsilon_r$  is the black hole radiative efficiency. The feedback mode depends on whether or not the ratio  $\dot{M}/\dot{M}_{\text{Edd}}$  exceeds the following threshold:

$$\chi = \min \left[ \chi_0 \left( \frac{M_{\text{BH}}}{10^8 M_\odot} \right)^\beta, 0.1 \right],\tag{8.4}$$

where  $\chi_0$  and  $\beta$  are parameters. If  $\dot{M}/\dot{M}_{\text{Edd}} > \chi$ , the resulting thermal feedback will inject thermal energy into the surrounding gas at a rate  $\dot{E}_{\text{therm}} = \epsilon_{\text{f,high}} \epsilon_r \dot{M} c^2$ , where  $\epsilon_{\text{f,high}}$  is another parameter; and if  $\dot{M}/\dot{M}_{\text{Edd}} < \chi$ , the resulting kinetic feedback will inject energy into the surroundings at a rate  $\dot{E}_{\text{kin}} = \epsilon_{\text{f,kin}} \dot{M} c^2$ , where the factor  $\epsilon_{\text{f,kin}}$  depends on the ambient density  $\rho$ . Both of these feedback modes will reduce the efficiency of star formation

in the surrounding gas, either by blowing gas out, so that less gas will exceed the density threshold for star formation, or by heating up gas which, as for stellar feedback, reduces the amount of cool gas capable of forming stars. As discussed above, the star formation efficiency is already reduced by our lowered gas resolution; reducing the overall effect of black hole feedback on star formation therefore provides another means of rectifying this.

For our retuning of the black hole feedback, we have increased  $\epsilon_r$  from the fiducial TNG value 0.2 to 0.22. The effect of this change on the overall energy release is quite complex: the energy injected by thermal feedback will be boosted, unless  $\dot{M} = \dot{M}_{\text{Edd}}$  (i.e.,  $\dot{M}_{\text{Bondi}} > \dot{M}_{\text{Edd}}$ ) in which case the  $\epsilon_r$  factors will cancel and there will be no effect; on the other hand, from Eq. (8.3) we see that  $\dot{M}_{\text{Edd}}$  is lowered if  $\epsilon_r$  is increased, and subsequently the ratio  $\dot{M}/\dot{M}_{\text{Edd}}$  will be greater and there will then be less kinetic feedback. From this discussion, increasing  $\epsilon_r$  is therefore expected to increase the heating of the gas by thermal feedback and reduce the blowing out of gas by kinetic feedback: two effects which would have competing impacts on the star formation efficiency. For our calibration runs, we have observed that increasing  $\epsilon_r$  to 0.22 boosts the amount of star formation. Therefore, it seems that the reduced blowing out of gas by kinetic feedback has the dominant effect here.

The result of making this final adjustment to the baryonic physics model is shown by the cyan lines in Fig. 8.1. The stellar mass fraction and stellar mass function are both slightly boosted for high-mass haloes. From the upper-right panel of Fig. 8.1, our model now appears to slightly overshoot the observed stellar mass function at higher masses; this is actually a consequence of sample variance which results from using a small box-size. As we will show in Sec. 8.2.2, the agreement is very good for the much larger  $301.75h^{-1}\text{Mpc}$  box size. The change to  $\epsilon_r$  also brings the galaxy size relation into closer agreement with the TNG L25-N512 runs, while the good agreement with observations for the black hole mass relation, the gas mass fraction and the SFRD is unaffected.

We also considered the minimum halo mass for black hole seeding. Central black holes are only found in haloes with mass above this threshold. Increasing the threshold means that, at a given time, black holes will have been growing for a shorter period of time and will consequently have a lower mass. This results in lower accretion and therefore reduces the

energy released through feedback. We found that this can significantly boost star formation in higher-mass haloes (which contain larger black holes and are therefore more susceptible to black hole feedback) but has very little effect on the stellar content of low-mass haloes. We found no clear advantage to vary this in addition to the other parameters varied in this work, therefore we adjusted the black hole feedback using only the  $\epsilon_r$  parameter.

#### 8.2.1.5 *Summary and further comments*

In summary, our retuned baryonic model uses updated parameter values  $\rho_\star = 0.08\text{cm}^{-3}$ ,  $\bar{e}_w = 0.5$  and  $\epsilon_r = 0.22$ , in addition to a larger gravitational softening length, to get sufficient star formation.

While our retuning of the baryonic physics has significantly boosted star formation across the full mass range shown in Fig. 8.1, it is still unable to give sufficient star formation at lower masses compared to observational data. Therefore, the stellar mass fraction and stellar mass function are both underestimated at the low-mass end, and the SFRD is underestimated at redshifts  $z \gtrsim 5$  (where there are only low-mass haloes). We attempted to rectify this by using even lower values of  $\rho_{\text{gas}}$  and reduced stellar and black hole feedback, but found that this offered little improvement overall. We even tried switching off feedback entirely, by setting the stellar wind energy to zero ( $\bar{e}_w = 0$ ) and by preventing the seeding of black holes: while this resulted in a huge amount of star formation at masses  $M_{200} \gtrsim 10^{12}M_\odot$ , there was still insufficient star formation at masses  $M_{200} \lesssim 10^{11.5}M_\odot$  to match observations. Therefore, the only way to have sufficient star formation across the full mass range appears to be by increasing the mass resolution. Interestingly, the BAHAMAS simulations (McCarthy et al., 2017) are able to achieve sufficient star formation for the full mass range (see the dotted lines in the top panels of Fig. 8.1) despite having  $\sim 3\times$  lower mass resolution than our simulations. The BAHAMAS simulations were run using the GADGET-3 code (Springel, 2005), which uses SPH rather than the Voronoi mesh. Perhaps the contrasting treatments of the gas by the two codes can explain the different levels of star formation at these lowered resolutions. One possible way to further boost star formation in low-mass haloes is by having a halo mass dependency for some of the baryonic

parameters, but this approach is beyond the scope of this work. On the other hand, we note that our low-resolution simulations are designed primarily for studying galaxy groups and clusters ( $M_{500} \gtrsim 10^{13} M_{\odot}$ ), for which the predictions of our model appear to be very reasonable.

While retuning these parameters, we came across a number of degeneracies. For example, as discussed above in Sec. 8.2.1.2, we found that the stellar-induced wind feedback can be lowered by reducing the speed of the winds rather than the wind energy. And, in our final model, we could have instead used a slightly increased  $\rho_{\star}$  (e.g.,  $\rho_{\star} = 0.09 \text{ cm}^{-3}$ ) and reduced  $\epsilon_r$  (e.g.,  $\epsilon_r = 0.18$ ) to achieve similar results. Therefore, we note that different combinations of parameter values could have been used to achieve a similar level of agreement with the observational data.

### 8.2.2 Large-box simulation

Our full simulation (‘L302-N1136’), which has been run using the retuned baryonic model at the same mass resolution as the L68-N256 calibration runs, has a box size of  $301.75 h^{-1} \text{ Mpc}$  and contains  $1136^3$  dark matter particles and (initially) the same number of gas cells. The simulation has been run for both GR and F5.

The red lines in Fig. 8.2 show the GR predictions of the six observables used to calibrate the baryonic model. The results are slightly different compared to the cyan lines in Fig. 8.1, which use the same baryonic model: the predicted amplitudes of the stellar mass fraction and stellar mass function are slightly lower, which actually improves the high-mass agreement with observations of the latter; and the amplitude of the galaxy size relation is greater for  $10^{11} M_{\odot} \lesssim M_{\star} \lesssim 3 \times 10^{11} M_{\odot}$ , leading to slightly worse agreement with the TNG L25-N512 predictions at these masses. These effects are likely to be a consequence of using a much larger box size, which is less susceptible to sample variance. The L302-N1136 simulation also extends to higher masses ( $M_{500} \sim 10^{15} M_{\odot}$  and  $M_{\star} \sim 10^{12} M_{\odot}$ ) than the L68-N256 runs. At these masses, the agreement with the observational data in Fig. 8.2 looks excellent.

The predictions for the F5 model, shown by the green lines in Fig. 8.2, agree with the

GR predictions for the galaxy size and black hole mass relations; however, the amplitudes of the other four observables are slightly boosted in F5 compared to GR. The SFRD is boosted for redshifts  $0.5 \lesssim z \lesssim 3$ : this is consistent with the results for the SHYBONE simulations (Arnold et al., 2019b). The stellar mass fraction and stellar mass function are boosted at  $M_{200} \sim 10^{12} M_{\odot}$  and  $M_{\star} \sim 10^{11} M_{\odot}$ , respectively, and the gas mass fraction is slightly enhanced for masses  $M_{500} \gtrsim 10^{13} M_{\odot}$ . The F5 predictions are still in excellent agreement with the observations, therefore it is not necessary to carry out a separate retuning of the baryonic physics for this model.

### 8.3 Results

Using the L302-N1136 simulations (see Sec. 8.2), we have measured the observable-mass scaling relations in GR and F5 for FOF groups in the mass range  $10^{13} M_{\odot} \leq M_{500} \lesssim 10^{15} M_{\odot}$ . In Sec. 8.3.1, we will discuss the relations for the mass-weighted gas temperature  $\bar{T}_{\text{gas}}$ , the SZ  $Y$ -parameter  $Y_{\text{SZ}}$ , the X-ray analogue of the  $Y$ -parameter  $Y_{\text{X}}$  and the X-ray luminosity  $L_{\text{X}}$ , which have all been computed in the same way as in Chapter 4. We again exclude gas cells found in the core region, defined by the radial range  $r < 0.15 R_{500}$ , and we will test the ‘true density’ mappings (see Chapter 4) between the F5 and GR relations for redshifts  $0 \leq z \leq 1$ . We will also discuss scaling relations which don’t involve the cluster mass in Sec. 8.3.2; these can potentially be used to test gravity using galaxy groups and clusters with no requirement to measure or infer the mass.

#### 8.3.1 Observable-mass scaling relations

The top rows of Figs. 8.3-8.6 show the F5 and GR scaling relations for redshifts 0, 0.5 and 1, with data points representing individual haloes. At  $z = 0$ , there are  $\sim 8000$  GR haloes with  $M_{500} > 10^{13} M_{\odot}$ , including  $\sim 500$  clusters with  $M_{500} > 10^{14} M_{\odot}$ . This is a significant improvement on the L62 SHYBONE simulations, which only had  $\sim 100$  haloes with  $M_{500} > 10^{13} M_{\odot}$  at  $z = 0$ . The curves in the top rows of the figures show the median observable as a function of the mean logarithmic mass computed within mass bins; the ‘true density’ rescalings of the F5 relation, which are computed using our analytical tanh

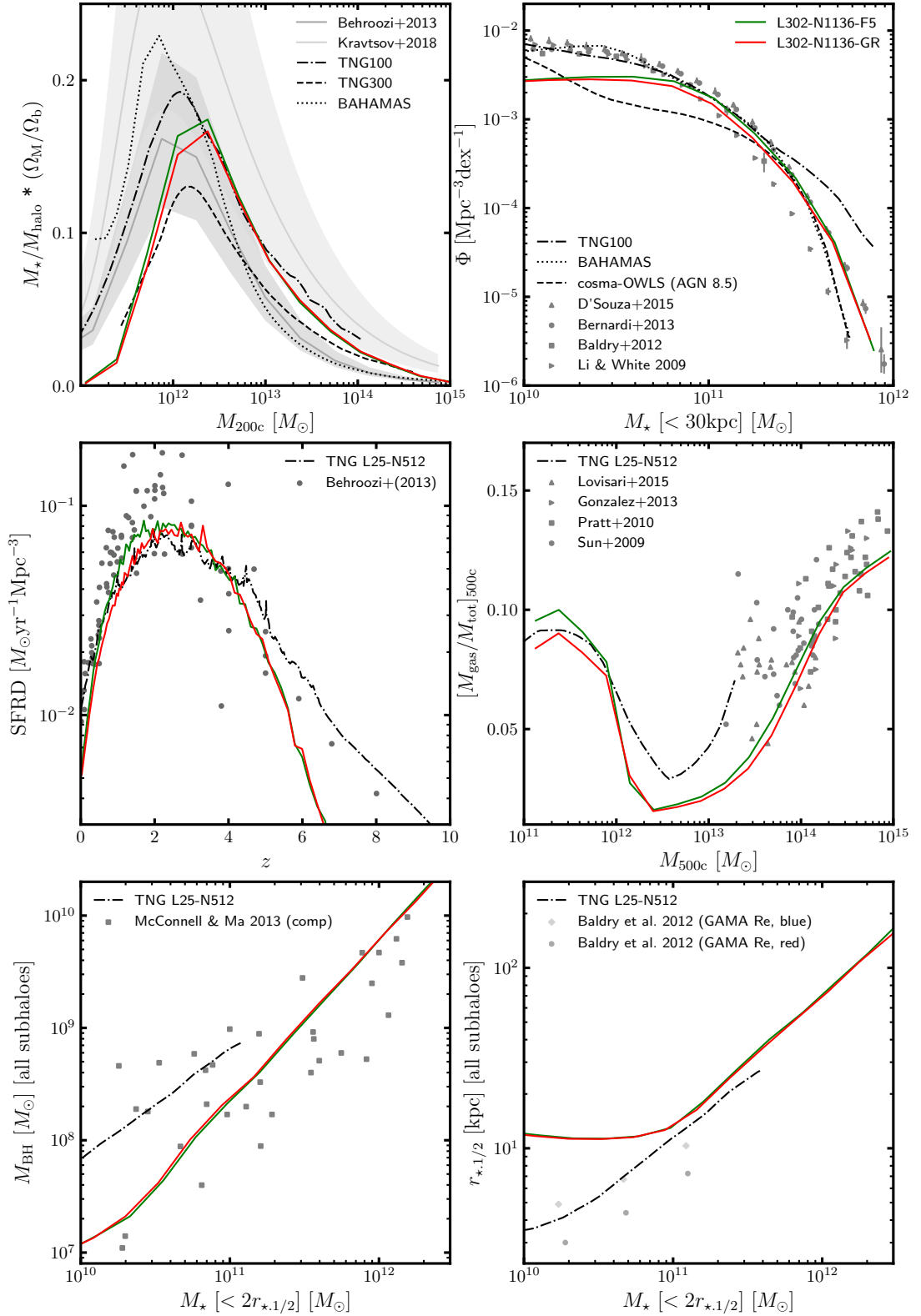


Figure 8.2: Stellar, gas and black hole properties of haloes in the L302-N1136 simulations for GR (red lines) and F5 (green lines). Apart from the coloured lines, the contents of this figure are identical to Fig. 8.1.

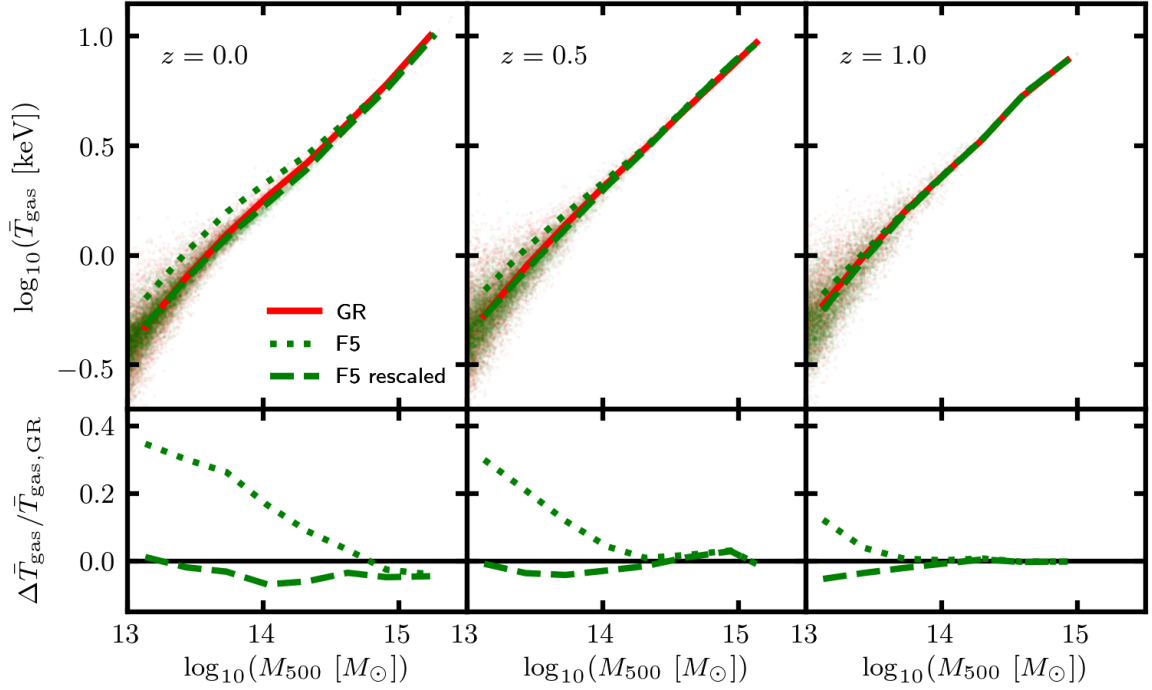


Figure 8.3: Gas temperature as a function of the halo mass for the full-physics L302 simulation (see Sec. 8.2) at redshifts 0, 0.5 and 1. The curves correspond to the median temperature and the mean logarithm of the mass computed within mass bins. Data has been included for GR (*red solid lines*) and F5 (*green lines*). A rescaling to the F5 temperature has been carried out as described in Sec. 8.3.1. Data points are displayed, with each point corresponding to a GR halo (*red points*) or to a halo in F5 (*green points*), including the rescaling. *Bottom row*: the relative difference between the F5 and GR curves in the above plots.

formula (Eq. (2.17)), are indicated by the dashed lines. We use eight mass bins with constant logarithmic width over the range  $10^{13}M_{\odot} \leq M_{500} \leq 10^{15.4}M_{\odot}$ . All bins are shown regardless of the halo count. Although there may only be a few haloes in the highest-mass bins, we note that these correspond to high-mass clusters for which scatter in the scaling relations, especially in the model difference between F5 and GR, is expected to be very low. The relative difference between the F5 and GR binned data is shown in the lower panels of the figures.

### 8.3.1.1 Temperature scaling relation

The results for the  $\bar{T}_{\text{gas}}(M)$  scaling relation are shown in Fig. 8.3. The GR data appears to follow the well-known power-law behaviour for cluster-sized objects; however, the relation



appears to curve at lower masses, where processes such as feedback can cause additional gas heating and break the power-law scaling. In F5, haloes are mostly screened from the fifth force for masses  $M_{500} \gtrsim 10^{14.5} M_{\odot}$  at  $z = 0$ ,  $M_{500} \gtrsim 10^{14} M_{\odot}$  at  $z = 0.5$  and  $M_{500} \gtrsim 10^{13.5} M_{\odot}$  at  $z = 1$ , and here the F5 temperature closely follows the GR temperature. At lower masses, the F5 temperature becomes significantly enhanced, as the total gravitational potential of the halo is raised by the fifth force.

Our rescaling of the F5 data, which we recall involves dividing the temperature by the ratio of the dynamical mass to the true mass, can successfully account for this offset at lower masses, restoring  $< 7\%$  agreement with the GR relation. However, the rescaled F5 relation now slightly underestimates the GR relation on average. We note that at  $z = 0$ , this offset appears to be roughly constant for cluster-sized masses; therefore, as long as the GR scaling relation parameters are allowed to vary in MCMC sampling (which can account for small differences in the amplitude), this rescaling is still expected to work well in our constraint pipeline presented in Chapter 5.

### 8.3.1.2 $Y_{\text{SZ}}$ and $Y_{\text{X}}$ scaling relations

The  $Y_{\text{SZ}}(M)$  and  $Y_{\text{X}}(M)$  relations are shown in Figs. 8.4 and 8.5, respectively. The GR relation appears to follow a weakly broken power-law, with a slightly steeper slope for group-sized haloes ( $M_{500} \lesssim 10^{14} M_{\odot}$ ) than for cluster-sized haloes ( $M_{500} \gtrsim 10^{14} M_{\odot}$ ). Again the low-mass behaviour can be explained by feedback, which, in addition to heating up gas, also blows gas out from the inner regions which in turn can lower the  $Y$  values. For example, in Chapter 4, we observed that the  $Y$ -parameter was lower in the full-physics simulations than in the non-radiative simulations, which did not include feedback.

For lower (unscreened) masses, we observe an enhancement of the F5 relations by up to  $\sim 50\%$  compared to GR. This is mostly corrected by our rescaling, after which the agreement is within  $\sim 12\%$  for group-sized haloes and is within a few percent on average for cluster-sized objects. This is positive news for our constraint pipeline in Chapter 5, which used this rescaling to model the  $Y_{\text{SZ}}(M)$  relation for clusters in the redshift range  $0 < z < 0.5$ .

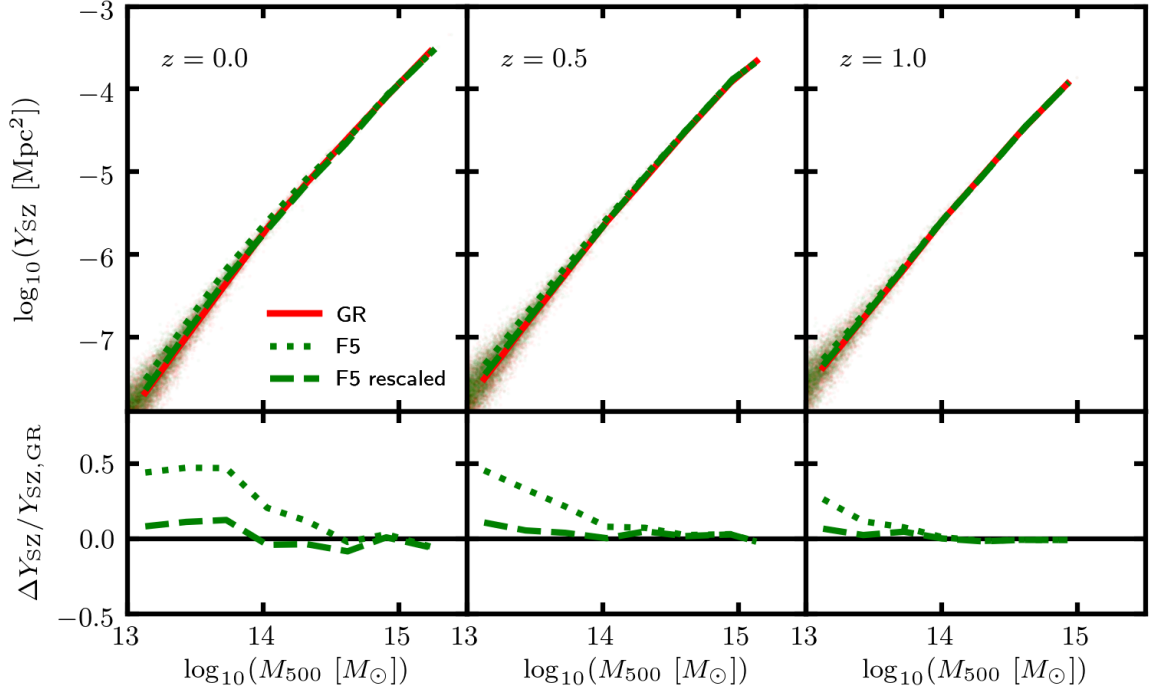


Figure 8.4: SZ Compton  $Y$ -parameter as a function of the halo mass for the full-physics L302 simulation (see Sec. 8.2) at redshifts 0, 0.5 and 1. Apart from the observable, this figure has the same layout as Fig. 8.3.

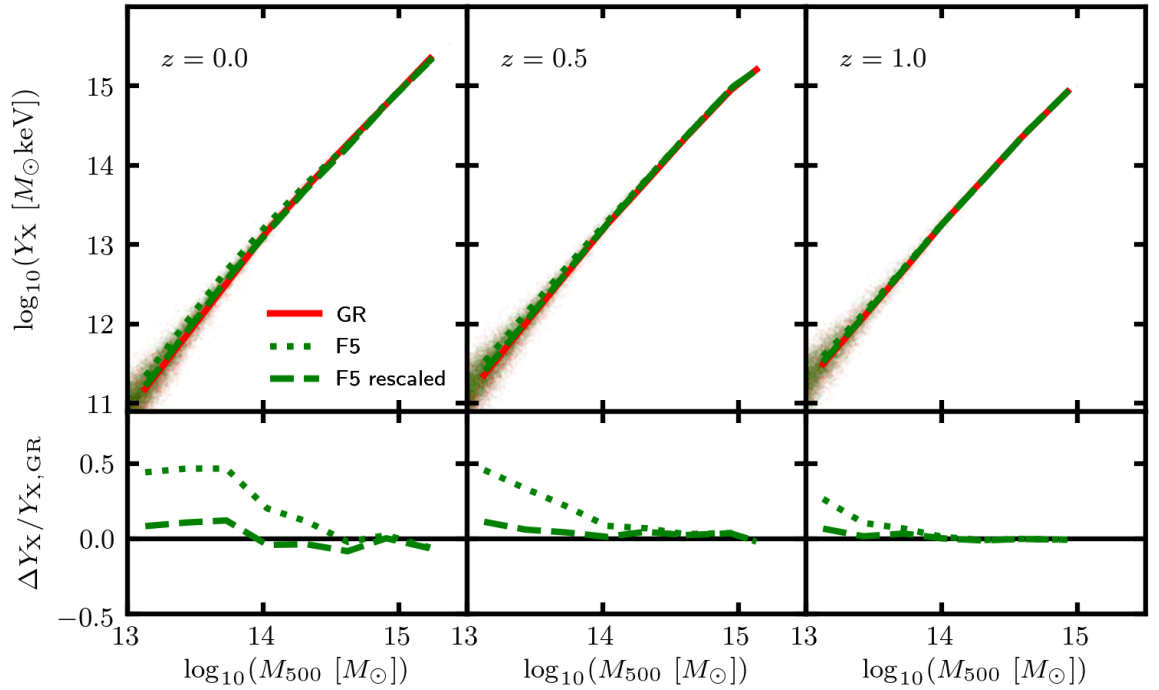


Figure 8.5: X-ray analogue of the  $Y$ -parameter as a function of the halo mass for the full-physics L302 simulation (see Sec. 8.2) at redshifts 0, 0.5 and 1. Apart from the observable, this figure has the same layout as Fig. 8.3.

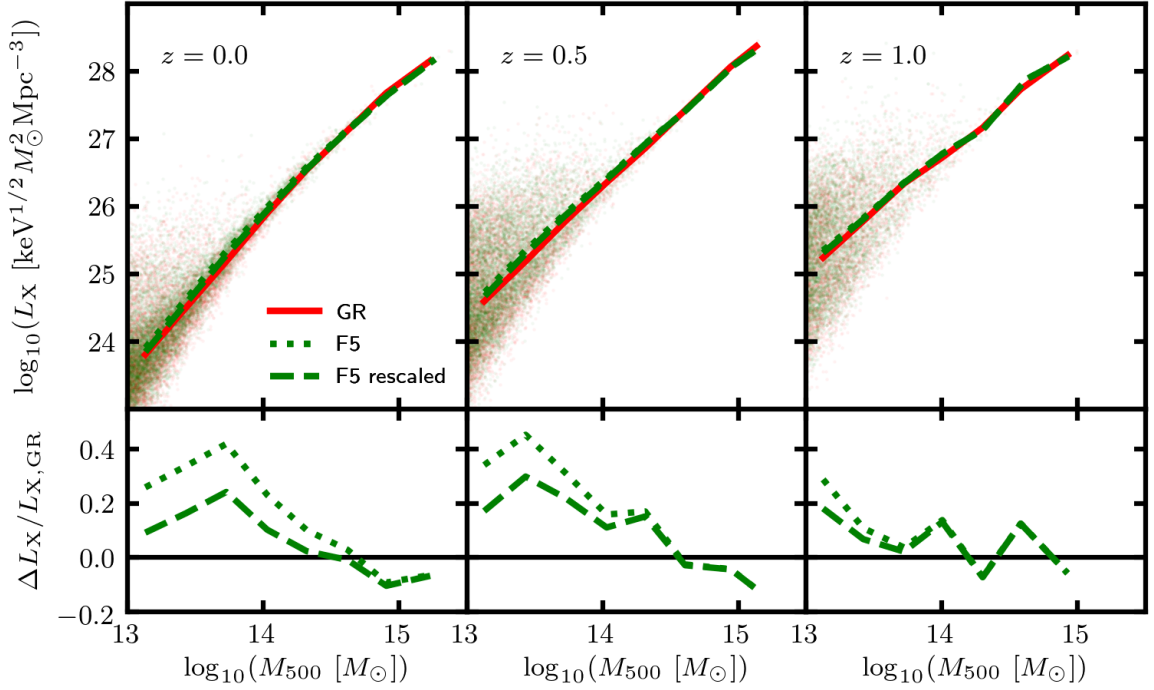


Figure 8.6: X-ray luminosity as a function of the halo mass for the full-physics L302 simulation (see Sec. 8.2) at redshifts 0, 0.5 and 1. Apart from the observable, this figure has the same layout as Fig. 8.3.

### 8.3.1.3 X-ray luminosity scaling relation

The  $L_X(M)$  relation is shown in Fig. 8.6. As we showed in Chapter 4, our rescalings were unable to accurately account for the difference between GR and  $f(R)$  gravity. That study was carried out primarily for group-sized haloes, and for these new results the rescaling is again unsuccessful for the mass range  $10^{13} M_\odot \lesssim M_{500} \lesssim 10^{14} M_\odot$ . The X-ray luminosity varies as  $T_{\text{gas}}^{1/2} \rho_{\text{gas}}^2$ . For the ‘true density’ rescaling, which is applied here, it is assumed that the gas temperature is enhanced by the fifth force while the gas density is unchanged. This may be the case in non-radiative simulations, however in full-physics simulations it is not necessarily true. For example, it is likely that there are different levels of feedback in F5 and GR. A greater amount of feedback in one model would result in the blowing out of gas and subsequent lowering of the gas density. This is expected to have a much greater effect on  $L_X$ , which varies as  $\rho_{\text{gas}}^2$ , than on the other observables considered in this work. For the  $Y$ -parameters, which vary as  $T_{\text{gas}} \rho_{\text{gas}}$ , the effects of feedback on the gas density and the temperature can roughly balance out (e.g., Fabjan et al., 2011), allowing our rescaling to

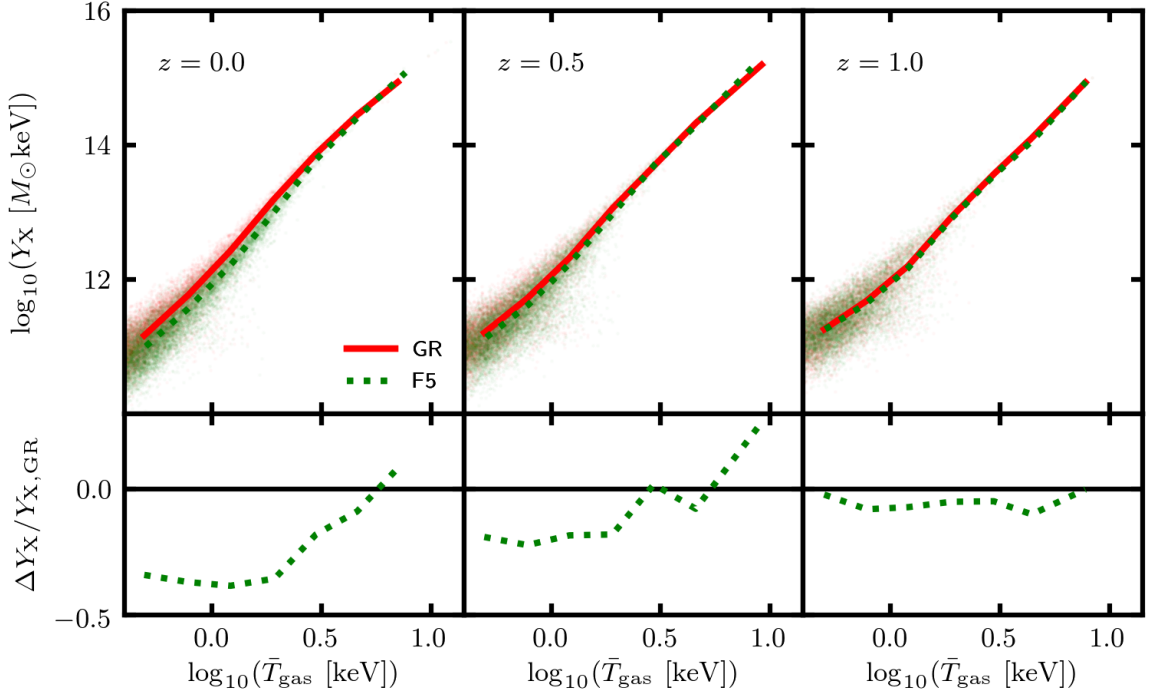


Figure 8.7: X-ray analogue of the  $Y$ -parameter as a function of gas temperature for haloes from the full-physics L302 simulation (see Sec. 8.2) at redshifts 0, 0.5 and 1. The curves correspond to the median luminosity and the mean logarithm of the temperature computed within temperature bins. Data has been included for GR (*red solid lines*) and F5 (*green dotted lines*). Data points are displayed, with each point corresponding to a GR halo (*red points*) or to a halo in F5 (*green points*). *Bottom row*: the relative difference between the F5 and GR curves in the above plots.

perform better for these observables as we saw in Figs. 8.4 and 8.5.

While the above is particularly problematic for galaxy groups, which are more susceptible to feedback, our rescaling appears to work reasonably well for cluster-sized haloes in Fig. 8.6, where the rescaling brings the agreement to within 10% at  $z = 0$ . However, the  $L_X(M)$  relation is also highly scattered compared to the other relations considered in this work. For example, the agreement between F5 and GR has a large  $\sim 20\%$  fluctuation at  $z = 1$  for  $M_{500} > 10^{14} M_\odot$ , even though clusters are completely screened at this redshift.

Based on this discussion, the  $\bar{T}_{\text{gas}}(M)$ ,  $Y_{\text{SZ}}(M)$  and  $Y_X(M)$  relations are more suitable than the  $L_X(M)$  relation for tests of gravity that involve the cluster mass.

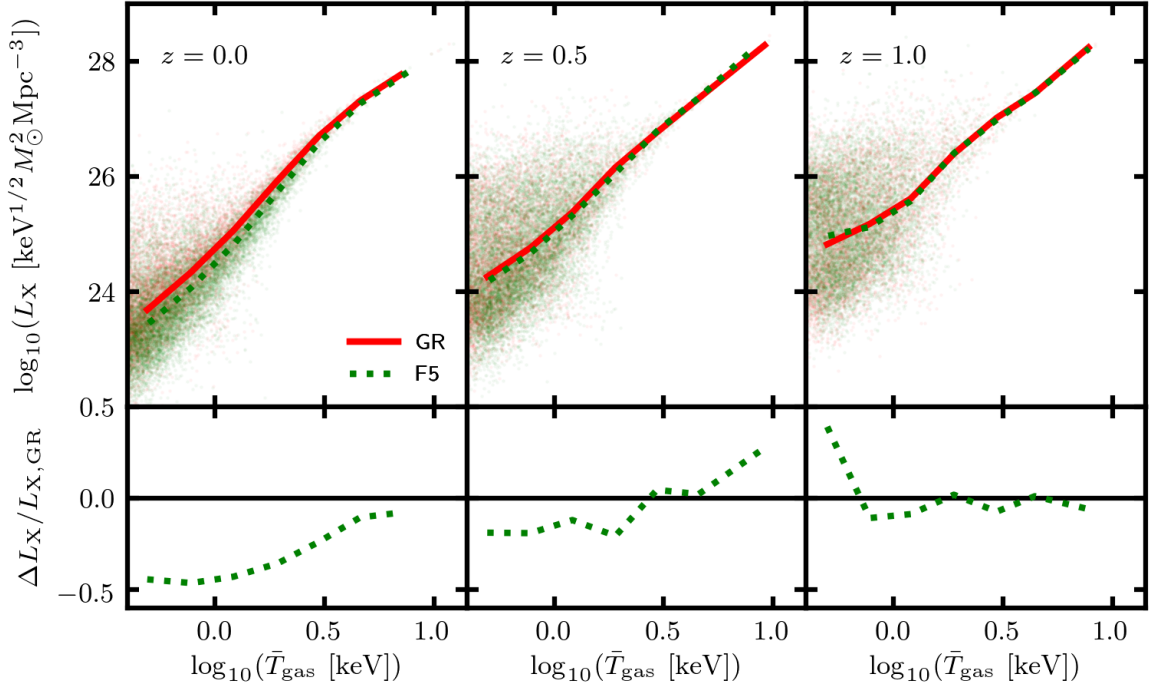


Figure 8.8: X-ray luminosity as a function of gas temperature for haloes from the full-physics L302 simulation (see Sec. 8.2) at redshifts 0, 0.5 and 1. Apart from the observable used in the vertical axis, this figure has the same format as Fig. 8.7.

### 8.3.2 $Y_X$ -temperature and $L_X$ -temperature relations

In Figs. 8.7 and 8.8, we show the  $Y_X(\bar{T}_{\text{gas}})$  and  $L_X(\bar{T}_{\text{gas}})$  relations, respectively, at redshifts 0, 0.5 and 1. The curves show the median  $Y_X$ -parameter and mean logarithmic temperature computed within seven temperature bins, with logarithmic width 0.2, spanning the range  $10^{-0.4} \text{keV} \leq \bar{T}_{\text{gas}} \leq 10^1 \text{keV}$ .

Haloes in F5 and GR with the same temperature are expected to have a similar dynamical mass; in this case, the F5 haloes would have a lower true mass than the GR haloes, and therefore a lower gas density (e.g., recall the ‘effective density’ rescalings in Chapter 4). This explains why, for  $Y_X(\bar{T}_{\text{gas}})$ , the amplitude of the F5 relation is suppressed by up to  $\sim 40\%$  compared to GR, while for  $L_X(\bar{T}_{\text{gas}})$  the F5 relation is suppressed by up to  $\sim 45\%$  (the differences may also be partly due to differences in the levels of feedback in the two models). For both relations, the difference is greater for lower redshifts and lower temperatures, where more haloes are unscreened.

Neither of these relations involve the cluster mass. Therefore, these could potentially

be used to test gravity using galaxy groups and clusters without the risk of bias from mass measurements. This demonstrates that, besides their abundances inferred from observables such as  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$ , the combination of different internal observational properties for a population of galaxy clusters or groups can also offer useful, possibly complementary, constraints on the theory of gravity.

## 8.4 Summary

Running large-box cosmological simulations which simultaneously incorporate screened modified gravity and full baryonic physics can be computationally expensive, necessitating the use of lower mass resolutions so that the calculations can involve fewer particles. However, this means that the gas density field is smoothed, resulting in high-density peaks being lost and consequently an overall reduction in star formation. This can result in poor agreement with observations of the stellar and gaseous properties of galaxies.

In this chapter, we have retuned the IllustrisTNG baryonic model so that it can be used to run full-physics simulations at a much lower resolution while still retaining a high level of agreement with galaxy observations. Calibrated using runs with a box size of  $68h^{-1}\text{Mpc}$  and, initially,  $256^3$  gas cells, our model uses updated values for the following parameters (Sec. 8.2.1): the threshold gas density for star formation,  $\rho_{\star}$ , is reduced from  $\approx 0.1\text{cm}^{-3}$  to  $0.08\text{cm}^{-3}$ ; the parameter  $\bar{e}_{\text{w}}$  which controls the energy released by the stellar-driven wind feedback is reduced from 3.6 to 0.5; and the black hole radiative efficiency  $\epsilon_{\text{r}}$  is increased from 0.2 to 0.22. In addition to these changes, we have also increased the gravitational softening to a factor 1/20 of the mean interparticle separation. By reducing the heating and blowing out of gas by feedback and two-body interactions, and lowering the threshold density of star formation, these changes boost the amount of star formation at our lowered resolution, resulting in good agreement with observations of galaxy properties including the stellar mass fraction, the stellar mass function, the SFRD and the gas mass fraction (Fig. 8.1).

Using our retuned model, we have run GR and F5 simulations with a box size  $301.75h^{-1}\text{Mpc}$  (Sec. 8.2.2). The predictions of stellar and gaseous properties in both grav-

ity models show a very good match with galaxy observations, particularly for group- and cluster-sized masses (Fig. 8.2), which shows that for F5 it is not necessary to further retune the baryonic parameters. Using these simulations, we have studied, for redshifts  $0 \leq z \leq 1$  and masses  $10^{13} M_{\odot} \leq M_{500} \lesssim 10^{15} M_{\odot}$ , the scaling relations between the cluster mass and four observable mass proxies (Sec. 8.3.1): the SZ Compton  $Y$ -parameter  $Y_{\text{SZ}}$  and its X-ray analogue  $Y_{\text{X}}$ , the mass-weighted gas temperature  $\bar{T}_{\text{gas}}$ , and the X-ray luminosity  $L_{\text{X}}$ .

For the  $Y_{\text{SZ}}(M)$  and  $Y_{\text{X}}(M)$  relations, our mapping between the F5 and GR relations, which involves dividing the F5  $Y$ -parameter by the ratio of the dynamical mass to the true mass, is accurate to within  $\sim 12\%$  for galaxy groups and just a few percent for galaxy clusters. This validates our method for accounting for the effect of the fifth force on the  $Y_{\text{SZ}}(M)$  relation, which is currently used in our  $f(R)$  constraint pipeline (Chapter 4). For the  $\bar{T}_{\text{gas}}(M)$  relation, the same rescaling is again reasonable, with  $\lesssim 7\%$  accuracy for the full range of masses. Our rescaling does not work as well for the  $L_{\text{X}}(M)$  relation, which is likely due to the greater susceptibility of the X-ray luminosity to feedback processes.

We have also shown (Sec. 8.3.2) that the  $Y_{\text{X}}$ -temperature and  $L_{\text{X}}$ -temperature scaling relations can differ in F5 and GR by up to 45%. These relations could potentially be used for large-scale tests of gravity that do not involve measuring the cluster mass, and hence not only eliminating one potential source of uncertainty but also including additional information in the model constraints.

By running large-box full-physics simulations for a range of  $f(R)$  gravity field strengths, it will be possible to test our models for the enhancements of the dynamical mass (Chapter 2) and the halo concentration (Chapter 3) in the presence of full baryonic physics over a wide mass range. Our baryonic model can also potentially be used to run large full-physics simulations for other classes of modified gravity and dark energy models, e.g., the nDGP model using the MG solvers implemented in the AREPO code, since it is likely that a recalibration of the baryonic parameters will not be necessary unless the model studied is extreme and differs strongly from the current best-fit  $\Lambda$ CDM (but in that case the model is likely to have already been ruled out by other observations). The application to the nDGP model will make it possible to validate our models for the nDGP enhance-

ments of the halo concentration and the HMF in addition to extending our results for the observable-mass scaling relations to higher masses (Chapter 6). Finally, in our study of the thermal and kinetic SZ angular power spectra in  $f(R)$  gravity and nDGP (Chapter 7), we were unable to study larger angular scales ( $l \lesssim 500$ ), again due to the relatively small box size of the SHYBONE simulations: this can potentially be rectified by using these larger simulations. These possibilities will be explored in future works.

The ability to run large realistic galaxy and cluster formation simulations for beyond- $\Lambda$ CDM models will prove highly beneficial for research in this field: not only will this endow us with numerical tools to predict observables, such as cluster properties, that cannot be studied using DMO simulations, but the hydrodynamical simulations enabled by such a tool can be used to quantify the impacts of baryons on various other observables, such as weak lensing and galaxy clustering. The lack of such a quantitative assessment would either restrict the amount of data that can be reliably used in model tests, or lead to biased constraints on models and parameters.



# Conclusions and outlook

## 9.1 Thesis summary

### 9.1.1 Modelling the dynamical mass enhancement in $f(R)$ gravity

In Chapter 2, we found a simple model to describe the relationship between the dynamical mass and lensing mass of dark matter haloes in the Hu-Sawicki  $f(R)$  model. As shown by the solid line fits of Fig. 2.3, the tanh fitting formula of Eq. (2.17) has generally shown excellent agreement with AHF halo data, for  $z < 1$ , from three ECOSMOG DMO simulations, which are summarised in Table 2.1. By taking advantage of the variety of resolutions offered by these simulations, and using  $\Lambda$ CDM simulations to produce approximate data for field strengths not covered by the  $f(R)$  gravity simulations, the validity of Eq. (2.17) has been probed vigorously across a wide and continuous range of field values that cover  $10^{-6.5} < |f_{R0}| < 10^{-4}$  within  $z < 1$ .

In addition to this, we used a simple thin-shell model (Sec. 2.2.1) to predict the behaviours of free parameters  $p_1$  and  $p_2$  in Eq. (2.17), which characterise the inverse width and the central logarithmic mass of the tanh-like transition respectively. The predictions, which neglect the effects of environmental screening due to nearby dark matter haloes, are given by Eqs. (2.18, 2.19). Using a stringent criterion to exclude unreliable snapshots in the fitting, the result for  $p_2$ , shown in Fig. 2.4, is  $p_2 = (1.503 \pm 0.006) \log_{10} \left( \frac{|\bar{f}_R|}{1+z} \right) + (21.64 \pm 0.03)$ . The slope value of  $1.503 \pm 0.006$  shows excellent agreement with the prediction of 1.5 by

Eq. (2.19), and the data of Fig. 2.4 shows a clear linear trend as predicted. As shown by Fig. 2.5, the  $p_1$  data is more scattered, but given the size of the one standard deviation error bars, the constant trend predicted by Eq. (2.18) is not unreasonable, resulting in  $p_1 = (2.21 \pm 0.01)$ . As shown by the dashed line fits of Fig. 2.3, these results for  $p_1$  and  $p_2$  show good agreement with the simulation data across the full range of field values and redshifts. We also repeated the analysis using a different approach to utilise the errors in the simulation data, and the results, shown in Appendix A.1, also agree with the thin-shell model prediction very well. In Appendix A.2 we further argue that the results in this chapter apply to models with different cosmological parameters such as  $\sigma_8$  and  $\Omega_M$ .

A generic fitting function for the relationship between the dynamical and lensing masses of dark matter haloes is an essential ingredient of the new framework, proposed in this thesis, to carry out cosmological tests of gravity in an unbiased way. Taking Eq. (2.16) as an example, our general formula for the dynamical mass enhancement allows us to incorporate this particular effect of  $f(R)$  gravity into galaxy cluster scaling relations in a self-consistent way. A key benefit of a fitting function is that it allows a continuous search through the model parameter space without having to run full simulations for every parameter point sampled in MCMC. The results will also be useful for other cluster tests of gravity that employ the difference between dynamical and lensing masses, such as by comparing cluster dynamical and lensing mass profiles, or by looking at measured cluster gas fractions.

The results presented in this chapter indicate that a simple model sometimes works surprisingly well despite the greatly simplified treatment of the complicated nonlinear physics of (modified) gravity. This has become a common theme in this thesis, and fitting simple models to the predictions from numerical simulations is an approach that we have repeated when studying the halo concentration in  $f(R)$  gravity and the cluster and halo properties in nDGP.

Although we made a very specific choice of  $f(R)$  gravity in this chapter, the theoretical model and the procedure we followed to calibrate it are expected to be applicable to general chameleon gravity theories (e.g., Gronke et al., 2015, 2016). As discussed briefly

in Appendix A.2, in other  $f(R)$  models the transition between screened and unscreened regimes can be different from the Hu and Sawicki (2007) model with  $n = 1$ , which may cause the exact fitted values of  $p_i$  to differ from what we presented in the above. Therefore, other  $f(R)$  models may require a re-calibration based on simulations. However, given that all  $f(R)$  models are phenomenological, it is perhaps more sensible to focus on a representative example, such as that by Hu and Sawicki (2007), to make precise observational constraints. The pipeline and methodology can then be applied to any other models following general parameterisation schemes (e.g., Brax et al., 2012b,a; Lombriser, 2016), which are useful for capturing the essential features of large classes of models using a few parameters. Should a preferred one emerge, the conclusion for the Hu-Sawicki model can serve as a rough guideline as to what level future cluster observations can constrain scalar-tensor-type screened theories. For this reason we decided not to explore other forms of  $f(R)$  in this thesis.

### 9.1.2 Universal model for the halo concentration in $f(R)$ gravity

In Chapter 3, we calibrated a model for the enhancement of the halo concentration in HS  $f(R)$  gravity with  $n = 1$  using a suite of simulations that are summarised by Table 2.1. The model is shown in Fig. 3.6, and is given by Eq. (3.9) with the parameter values listed in Table 3.2. It has been defined in terms of a useful rescaling of the halo mass,  $M_{500}/10^{p_2}$ , such that the data from three different  $f(R)$  gravity models can satisfy a universal description. These models have  $\log_{10}(|f_{R0}|) = (-4, -5, -6)$ , and the fitting was carried out using data from all simulation snapshots with  $\log_{10}(|\bar{f}_R|/(1+z)) \leq -4.5$ . This universal description was shown to have very good agreement with simulations for  $M_{500}/10^{p_2}$  covering nearly 7 orders of magnitude, and covering five decades of the halo mass.

Our model has been tested by comparing its predictions of the enhancement of the concentration with an arbitrarily chosen set of snapshots from our simulations, as shown by the lines plotted in Figs. 3.2 and 3.3. These predictions show excellent agreement with the data for all snapshots, apart from the Crystal snapshots with  $\log_{10}(|\bar{f}_R|/(1+z)) > -4.5$ . This is not surprising given that this data was not used in the fit of the model. Having a

general model that works for  $\log_{10} \left( |\bar{f}_R|/(1+z) \right) \leq -4.5$  will prove very useful, particularly given that an analytical theoretical modelling was not available.

The data of Fig. 3.6 shows that in the unscreened regime the enhancement of the concentration reaches a distinct peak as a function of the halo mass, but drops to negative values at lower mass, where the  $f(R)$  concentration is less than the GR concentration. As shown by Fig. 3.1, such negative enhancement occurs because the innermost regions of the haloes are less dense in  $f(R)$  gravity than in GR. This could be caused by the velocity gained by particles in haloes, which makes it difficult for them to settle into orbits at the central regions of the halo. Meanwhile in the screened regime of Fig. 3.6 there is a small dip in the concentration. Fig. 3.1 suggests that this is caused by the halo being only partially screened, so that outer particles are moved further towards the centre of the halo while the inner regions remain screened. The density profile is therefore unaffected at the innermost regions but is greater at intermediate radii. Therefore the scale radius becomes greater, and fitting an NFW profile would then result in an estimate for the concentration that is lower in  $f(R)$  gravity than in GR. All of these effects are incorporated by the fitted model of Eq. (3.9).

We also carried out some further investigations which can be useful for future studies of the concentration in  $f(R)$  gravity, and in other similar modified gravity theories. Firstly, in addition to applying a direct NFW profile fitting to each of the haloes to measure the concentration, two simplified approaches were also used, namely the methods that are used by Prada et al. (2012) and Springel et al. (2008). The resulting enhancement of the concentration from using these two methods (shown in Fig. 3.4) shows a difference from direct NFW fitting. This is due to the effects of  $f(R)$  gravity on the internal density profile, which means that the choice of regions of the halo to use in measuring the concentration becomes important. The method used by Springel et al. (2008) only requires the mass enclosed by the orbital radius with the maximum circular velocity. Being found at the inner regions of a halo, which become more dense as the halo becomes unscreened, this results in the concentration being overestimated by up to 26%. From this, we conclude that only the direct NFW fitting should be used in  $f(R)$  studies. Secondly, we looked at

the validity of the NFW profile fitting in  $f(R)$  gravity and found that, as shown by the bottom-left panel of Fig. 3.5, for most haloes the  $\chi^2$  measure for the fit is within 20% of the GR measure, and for some haloes the fit is even better. Therefore the systematic effects caused by fitting the NFW profile in  $f(R)$  gravity are unlikely to have a significant effect on the scatter of the concentration measure.

### 9.1.3 Observable-mass scaling relations in $f(R)$ gravity

In Chapter 4, we made use of the first full-physics simulations that have been run for both GR and  $f(R)$  gravity (along with non-radiative counterparts), to study the effects of the fifth force of  $f(R)$  gravity on the scaling relations between the cluster mass and four observable proxies: the gas temperature (Fig. 4.2), the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  parameters (Figs. 4.3 and 4.4) and the X-ray luminosity (Fig. 4.5). To understand these effects in greater detail, we have also examined the effects of both  $f(R)$  gravity and full-physics on the gas density and temperature profiles (see Fig. 4.1). In doing so, we have been able to test two methods for mapping between scaling relations in  $f(R)$  gravity and GR.

The first method was proposed by He and Li (2016). This proposes a set of mappings, given by Eqs. (4.4) and (4.6)-(4.8), that can be applied to haloes whose mass and radius are measured using the effective density field (see Sec. 4.2.1). A second, new, approach is proposed in Sec. 4.2.2, and predicts another set of mappings, given by Eqs. (4.9) and (4.11)-(4.13), that can be applied to haloes whose mass and radius are measured using the true density field. Both sets of mappings involve simple rescalings that depend only on the ratio of the dynamical mass to the true mass in  $f(R)$  gravity. As shown by Figs. 4.6 and 4.7, even with the inclusion of full-physics processes this ratio can be computed with high accuracy using our analytical tanh formula, which is given by Eq. (2.17).

For the mass-weighted gas temperature and the  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  observables, we found that the F6 and F5 scaling relations, with appropriate rescaling applied (using either method discussed above), match the GR relations to within a few percent for the full mass-range tested for the non-radiative simulations. With the inclusion of full-physics effects such as feedbacks, star formation and cooling, the rescaled  $Y_{\text{SZ}}$  and  $Y_{\text{X}}$  scaling relations continue

to show excellent agreement with GR for mass  $M_{500} \gtrsim 10^{13.5} M_\odot$ , which includes group- and cluster-sized objects. These proxies also show relatively low scatter as a function of the cluster mass, compared with other observables.  $Y_{\text{SZ}}$  and  $Y_X$  are therefore likely to be suitable for accurate determination of the cluster mass in tests of  $f(R)$  gravity. The mappings for the gas temperature show a very high accuracy for lower-mass objects, but show a small  $\lesssim 5\%$  offset between F5 and GR for higher-mass objects.

The mappings do not work as well for the X-ray luminosity  $L_X$ , for which the F5 relations after rescaling are typically enhanced by  $\sim 30\%$  compared with GR. This is caused by the unique dependency of  $L_X$  on the gas density to power two, and the gas temperature to power half, which means that the inner halo regions have a greater contribution than for the other proxies and the competing effects of feedback on the temperature and gas density profiles are less likely to cancel out. This issue, in addition to the fact that  $L_X$  has a highly scattered correlation with the cluster mass, means that this proxy is unlikely to be suitable for cluster mass determination in tests of  $f(R)$  gravity.

We also considered the  $Y_X - \bar{T}_{\text{gas}}$  scaling relation (Fig. 4.8), and found that this is suppressed by 30-40% in the F5 model relative to GR. This offers a potential new and useful test of gravity with group- and cluster-sized objects which avoids the systematic uncertainties incurred from mass calibration.

Our results also provide insights into the viability of extending cluster tests of gravity to the group-mass regime. An advantage of using lower-mass objects is that these objects can be unscreened (or partially screened) even for weaker  $f(R)$  models, offering the potential for tighter constraints using data from ongoing and upcoming SZ and X-ray surveys (e.g., Merloni et al., 2012; Ade et al., 2016) which are now entering this regime. On the other hand, as we have seen above, the scatter induced by feedback mechanisms becomes more significant in group-sized haloes, which means that additional work will need to be conducted to characterise this effect and to understand its impact on model tests.

Finally, we note that our parameter  $p_2$ , which is used to compute the ratio of the dynamical mass to the true mass, depends only on the quantity  $|f_R|/(1+z)$ , and not on the model parameters  $n$  and  $f_{R0}$  of HS  $f(R)$  gravity. This dependence was originally derived

by using the thin-shell model (Chapter 2), which does not depend on the details of the  $f(R)$  model. We therefore expect our scaling relation mappings to perform similarly for any combination of the HS  $f(R)$  parameters, and potentially other chameleon-type or thin-shell-screened models. However, due to the high computational cost of running full-physics simulations of  $f(R)$  gravity and other models, we do not seek to confirm this conjecture in this thesis.

#### 9.1.4 Constraint pipeline for unbiased $f(R)$ cluster constraints

In Chapter 5, we combined all of our models for the effects of  $f(R)$  gravity on cluster properties into an MCMC pipeline for constraining the amplitude of the present-day background scalar field,  $|f_{R0}|$ , using cluster number counts. We have adopted the model from [Cataneo et al. \(2016\)](#) for the  $f(R)$  enhancement of the HMF, and used this, along with our model for the enhancement of the halo concentration, to produce a model-dependent prediction of the cluster number counts (Sec. 5.3.1). We have also used our model for the enhancement of the dynamical mass in  $f(R)$  gravity to convert a GR power-law observable-mass scaling relation, which is based on the Planck  $Y_{\text{SZ}}(M_{500})$  relation ([Ade et al., 2016](#)), into a form consistent with  $f(R)$  gravity, where the fifth force enhances the relation at sufficiently low masses (Sec. 5.3.2). These models are all incorporated in our log-likelihood (Sec. 5.3.4), which we have used to infer parameter constraints using a set of mock cluster catalogues (Sec. 5.3.3).

Using a combination of GR and F5 mocks, we have shown that our pipeline is able to give reasonable parameter constraints that are consistent with the fiducial cosmology (Figs. 5.5 and 5.6). For the GR mock, the constraints conclusively rule out  $f(R)$  models with  $\log_{10} |f_{R0}| \gtrsim -5$  and favour values in the range  $-7 \leq \log_{10} |f_{R0}| \lesssim -5$  where  $-7$  is the lowest value considered by our MCMC sampling. Meanwhile, the constraints inferred using the F5 mock favour values close to the fiducial value of  $-5$ , with 68% range  $-5.1^{+0.3}_{-1.0}$  and a ‘most likely’ value of  $-4.92$ . We have also shown that the constraints inferred from both mocks can be imprecise and biased if the  $f(R)$  enhancement of the scaling relation is not accounted for (Fig. 5.7). Therefore, this should be properly modelled in future tests of

$f(R)$  gravity in order to prevent biased constraints. This will become particularly relevant as cluster catalogues start to enter the galaxy group regime (e.g., [Pillepich et al., 2018a](#); [Lovisari et al., 2021](#)), where more objects can be unscreened in  $f(R)$  gravity.

Throughout Chapter 5, the main obstacle to precise and unbiased constraints stemmed from degeneracies between  $f_{R0}$ ,  $\Omega_M$ ,  $\sigma_8$  and the scaling relation parameters  $\alpha$  and  $\beta$ , all of which can influence the predicted cluster count. We have shown that the degeneracies can be prevented by using a tighter Gaussian prior for  $\Omega_M$  and by having better knowledge of the scaling relation parameters (Fig. 5.11). The latter can potentially be achieved by including lensing data for a subset of the clusters. If wide or flat parameter priors are used, this may give rise to biased constraints of  $\log_{10} |f_{R0}|$ . For example, we have found that the parameter degeneracies can have a more significant effect for cluster samples that extend to lower masses (Sec. 5.5.2.1).

### 9.1.5 Cluster and halo properties in nDGP

In Chapter 6, we extended our framework to the popular nDGP model, in which a fifth force is able to act over sufficiently large scales.

Using the first cosmological simulations that simultaneously incorporate full baryonic physics and the nDGP model, we studied the observable-mass scaling relations for the same three mass proxies as studied in Chapter 4 (see Sec. 6.3.1). For groups and clusters in the mass range  $M_{500} \lesssim 10^{14.5} M_\odot$ , our results show that for the N1 model, the  $\bar{T}_{\text{gas}}(M)$  relation is enhanced by about 5% with respect to GR, while the  $Y_{\text{SZ}}(M)$  and  $Y_X(M)$  relations are both enhanced by 10%-15% at low masses but more closely match the GR relations at high masses. For N5, which is much weaker than N1, the  $\bar{T}_{\text{gas}}(M)$  relation closely resembles the GR relation, while the  $Y_{\text{SZ}}(M)$  and  $Y_X(M)$  relations are enhanced by up to 5% at low mass and suppressed by up to 5% at high mass. These deviations from GR could be related to the effect of the fifth force on gas velocities during cluster formation, and they also hint at an interplay between the fifth force and stellar and black hole feedback.

Using a suite of DMO  $N$ -body simulations, which cover a wide range of resolutions and box sizes, we found that, in nDGP, the concentration is typically suppressed relative



to GR, varying from a few percent in N5 to up to  $\sim 15\%$  in N0.5 (see Sec. 6.3.2). Using stacked density profiles at different mass bins, we have shown that this behaviour is caused by a reduced (increased) density at the inner (outer) halo regions. Including full baryonic physics significantly affects the concentration-mass relation; however, our results show that, for masses  $M_{200} \lesssim 10^{13} h^{-1} M_{\odot}$ , the model differences between nDGP and GR still have a similar magnitude compared to the DMO simulations.

By combining the data from our  $z \leq 1$  simulation snapshots, we calibrated a general model, given by Eq. (6.3), which is able to accurately predict the suppression of the halo concentration with respect to the GR results as a function of the halo mass and the  $H_0 r_c$  parameter of nDGP over ranges  $10^{12} h^{-1} M_{\odot} \lesssim M_{200} \lesssim 10^{15} h^{-1} M_{\odot}$  and 0.5-5, respectively. This model can be included in our MCMC pipeline for converting between mass definitions in case, for example, the theoretical predictions and observables are defined with respect to different spherical overdensities. Our model can also be used, along with the HMF, to predict the nonlinear matter power spectrum, which can also be used to constrain gravity.

We also used our DMO simulations to study the HMF over the mass range  $1.52 \times 10^{10} h^{-1} M_{\odot} \leq M_{500} \lesssim 10^{15} h^{-1} M_{\odot}$  at redshifts 0, 1 and 2 (see Sec. 6.3.3). Our results (Fig. 6.11), indicate that the nDGP HMF is enhanced at high masses (by up to  $\sim 60\%$  in N0.5) and suppressed at low masses (by  $\sim 10\%$  in N0.5) compared to GR. These results indicate the potential constraining power from using the observed mass function to probe the  $H_0 r_c$  parameter of nDGP. By combining the data from our  $z \leq 2$  snapshots, we have calibrated a general model, given by Eq. (6.4), which can accurately reproduce the HMF enhancement as a function of the halo mass, redshift and  $H_0 r_c$  parameter. This model can be used for theoretical predictions of the nDGP HMF (using a parameter-dependent GR calibration) in our MCMC pipeline.

In Chapter 4, we showed that a model for the  $f(R)$  dynamical mass enhancement can be used to predict observable-mass scaling relations in  $f(R)$  gravity using their GR counterparts. Such a model in nDGP could similarly be useful to help understand the enhancements of the temperature and SZ and X-ray  $Y$ -parameters observed in Chapter 6. For now, though, we note that the scaling relations in nDGP still appear to follow power-

law relations as a function of the mass: the  $\bar{T}_{\text{gas}}(M)$  relation in N1 can be related to the GR relation by a simple rescaling of the amplitude, whereas the  $Y_{\text{SZ}}(M)$  and  $Y_{\text{X}}(M)$  relations appear to have shallower slopes in N5 and N1 than in GR. Therefore, in our future MCMC pipeline for obtaining constraints of nDGP, we can still assume the GR power-law form of the scaling relations by allowing the parameters controlling the amplitude and slope to vary along with the cosmological and nDGP parameters (e.g., [de Haan et al., 2016](#); [Bocquet et al., 2019](#)).

Although our simulations have only been run for a single choice of cosmological parameters, we expect that our models for the enhancements of the halo concentration and HMF will have a reasonable accuracy for other (not too exotic) parameter values. The gravitational force enhancement in nDGP, given by  $[1 + 1/(3\beta)]$ , has only a weak dependence on  $\Omega_{\text{M}}$ : for the N1 model ( $\Omega_{\text{rc}} = 0.25$ ), the force enhancement varies within a very small range (roughly 12.1% – 12.6%) for  $\Omega_{\text{M}} \in [0.25, 0.35]$  at the present day, and the range of variation is even smaller at higher redshifts. Therefore, for now we assume that the effects of the cosmological parameters on the concentration and HMF are approximately cancelled out in the ratios  $\Delta c/c_{\text{GR}}$  and  $\Delta n/n_{\text{GR}}$ . However, we will revisit this in a future work, using a large number of nDGP simulations that are currently being run for different combinations of cosmological parameters, before these models are used in tests of gravity using observational data.

### 9.1.6 Sunyaev-Zel'dovich effect in $f(R)$ gravity and nDGP

Over the past couple of decades, great advances have been made in the measurement of the secondary anisotropies of the CMB caused by the SZ effect, including its thermal component and even its much smaller kinematic component. The angular power spectrum of the tSZ effect has been increasingly adopted as a probe of cosmological parameters that influence the growth of large-scale structures. Also, as observations of the kSZ power spectrum continue to improve, the latter has been identified as another potentially powerful probe of cosmology. The next generation of ground-based observatories ([Ade et al., 2019](#); [Abazajian et al., 2016](#)) look set to revolutionise the constraining power of these probes.

In Chapter 7, we used the SHYBONE simulations to look at the viability of using the angular power spectra of the tSZ and kSZ effects as large-scale probes of HS  $f(R)$  gravity and nDGP. We generated mock maps of the tSZ and kSZ signals (Sec. 7.2.2), and used these maps to measure the angular power spectra. Our results (Figs. 7.6 and 7.7) indicate that the fifth force, present in  $f(R)$  gravity and nDGP, and the subgrid baryonic physics have different effects on the tSZ and kSZ power spectra. The former enhances the power on all scales probed by our maps ( $500 \lesssim l \lesssim 8 \times 10^4$ ) by boosting the abundance and peculiar velocity of large-scale structures (e.g., dark matter haloes and free electrons inside them), while the latter brings about a suppression on scales  $l \gtrsim 3000$  for the tSZ effect and on all tested scales for the kSZ effect. Even with both of these effects present, we find that the power can be significantly enhanced in  $f(R)$  gravity and nDGP: by up to 60% for the tSZ effect and 35% for the kSZ effect for the F5 and N1 models; and by 5%-10% for F6 and N5, which correspond to relatively weak modifications of GR.

In addition, we computed the power spectrum of the transverse component of the electron momentum field (Sec. 7.3.3), which is closely related to the kSZ angular power spectrum. In particular, we showed (Fig. 7.9) that at angular sizes  $l \geq 600$  the kSZ signal is dominantly contributed by  $k$ -modes in the transverse-momentum power spectrum which are in the non-linear regime, and which are affected strongly by MG. The  $k$ -modes in the linear regime may contribute more to smaller  $l$ , but at least for  $f(R)$  gravity the impact of MG at those  $l$  values will be much less significant due to the finite range of the fifth force, as we can already see in Fig. 7.6.

We found that the relative difference between the MG models and GR is significantly affected by the additional baryonic processes that act in the full-physics simulations. Given that these processes are still relatively less well-constrained, this adds to the uncertainty in our theoretical predictions of the kSZ angular power spectra on small angular scales, e.g.,  $l > 600$ . Therefore, further work should be carried out using a range of full-physics parameters to precisely identify the scales on which constraints can be reliably made before the tSZ and kSZ power are used to probe  $f(R)$  gravity and nDGP.

### 9.1.7 Realistic simulations to study clusters in modified gravity

Running large-box cosmological simulations which simultaneously incorporate screened modified gravity and full baryonic physics can be computationally expensive, necessitating the use of lower mass resolutions so that the calculations can involve fewer particles. However, this means that the gas density field is smoothed, resulting in high-density peaks being lost and consequently an overall reduction in star formation. This can result in poor agreement with observations of the stellar and gaseous properties of galaxies.

In Chapter 8, we retuned the IllustrisTNG baryonic model so that it can be used to run full-physics simulations at a much lower resolution while still retaining a high level of agreement with galaxy observations. Calibrated using runs with a box size of  $68h^{-1}\text{Mpc}$  and, initially,  $256^3$  gas cells, our model uses updated values for the following parameters (Sec. 8.2.1): the threshold gas density for star formation,  $\rho_*$ , is reduced from  $\approx 0.1$  to 0.08; the parameter  $\bar{e}_w$  which controls the energy released by the stellar-driven wind feedback is reduced from 3.6 to 0.5; and the black hole radiative efficiency  $\epsilon_r$  is increased from 0.2 to 0.22. In addition to these changes, we have also increased the gravitational softening to a factor 1/20 of the mean interparticle separation. By reducing the heating and blowing out of gas by feedback and two-body interactions, and lowering the threshold density of star formation, these changes boost the amount of star formation at our lowered resolution, resulting in good agreement with observations of galaxy properties including the stellar mass fraction, the stellar mass function, the SFRD and the gas mass fraction (Fig. 8.1).

Using our retuned model, we have run GR and F5 simulations with a box size  $301.75h^{-1}\text{Mpc}$  (Sec. 8.2.2). The predictions of stellar and gaseous properties in both gravity models show a very good match with galaxy observations, particularly for group- and cluster-sized masses (Fig. 8.2), which shows that for F5 it is not necessary to further retune the baryonic parameters. Using these simulations, we have studied, for redshifts  $0 \leq z \leq 1$  and masses  $10^{13}M_\odot \leq M_{500} \lesssim 10^{15}M_\odot$ , the scaling relations between the cluster mass and four observable mass proxies (Sec. 8.3.1): the SZ Compton  $Y$ -parameter  $Y_{\text{SZ}}$  and its X-ray analogue  $Y_X$ , the mass-weighted gas temperature  $\bar{T}_{\text{gas}}$ , and the X-ray luminosity  $L_X$ .

For the  $Y_{\text{SZ}}(M)$  and  $Y_X(M)$  relations, our mapping between the F5 and GR relations,

which involves dividing the F5  $Y$ -parameter by the ratio of the dynamical mass to the true mass, is accurate to within  $\sim 12\%$  for galaxy groups and just a few percent for galaxy clusters. This validates our method for accounting for the effect of the fifth force on the  $Y_{\text{SZ}}(M)$  relation, which is currently used in our  $f(R)$  constraint pipeline (Chapter 4). For the  $\bar{T}_{\text{gas}}(M)$  relation, the same rescaling is again reasonable, with  $\lesssim 7\%$  accuracy for the full range of masses. Our rescaling does not work as well for the  $L_X(M)$  relation, which is likely due to the greater susceptibility of the X-ray luminosity to feedback processes.

We have also shown (Sec. 8.3.2) that the  $Y_X$ -temperature and  $L_X$ -temperature scaling relations can differ in F5 and GR by up to 45%. These relations could potentially be used for large-scale tests of gravity that do not involve measuring the cluster mass, and hence not only eliminating one potential source of uncertainty but also including additional information in the model constraints.

## 9.2 Outlook and future work

The results in this thesis show that the  $p_2$  parameter, defined in Chapter 2, can be very useful in the description and modelling of complicated effects in  $f(R)$  gravity. In addition to its relatively simple one-parameter definition, it can allow the combining of data generated by simulations run for different cosmological parameters, as  $p_2$  encapsulates the values of  $\Omega_M$  and  $\Omega_\Lambda$ . Indeed, the data for the concentration enhancement from AREPO and Diamond F6 shows excellent agreement (see Fig. 3.6), even though these two simulations were run for different cosmological parameters and using very different codes. It will be interesting to see where else  $p_2$  can be used in  $f(R)$  studies. Of particular interest would be to see how it can simplify the modelling of the HMF. The enhancement of the HMF in  $f(R)$  gravity peaks at a particular halo mass which depends on the strength of the scalar field. A stronger scalar field allows higher-mass haloes to be unscreened, and therefore results in an enhancement of the HMF at a higher mass. At the very least, the mass of the peak enhancement of the HMF can be expected to be strongly correlated to  $p_2$ . The enhancement of the matter power spectrum could also be investigated via a similar treatment.

Our  $f(R)$  constraint pipeline (Chapter 5) can be improved in a couple of ways. First,

while the HMF model of [Cataneo et al. \(2016\)](#) is accurate, it only covers the redshift range  $[0, 0.5]$ . An extended model that works for a larger redshift range, as well as for wider ranges of other cosmological parameters (not restricted to the  $\Omega_M$  and  $\sigma_8$  parameters as we have focused on here), would be very useful. Calibrating this model for spherical overdensity  $\Delta = 500$  would also mean that conversions between the halo mass definitions  $M_{500}$  and  $M_{300m}$  would no longer be required. Second, the MCMC pipeline should be extended so that independent cluster data, such as weak lensing, can be included in the model constraint. This can improve the precision of the mass calibration and potentially prevent degeneracies between the MG and scaling relation parameters. Once these tasks are completed, this pipeline can be used to constrain  $f(R)$  gravity using observations. It is also straightforward to extend our framework to other gravity models; as discussed, we have already started to do this for the nDGP model.

Applying the pipeline to real survey data will involve additional considerations. For example, our mock cluster samples were generated assuming a fixed cut in the observable and a constant fractional measurement uncertainty. Using real survey data will require more accurate modellings, including a selection function which describes the probability of a cluster with a given redshift and flux being selected. Therefore, tests should be carried out using mocks with more realistic selection criteria to confirm that the pipeline can infer stable and reliable constraints. It will be interesting to apply our pipeline to previous data (for example, the Planck SZ cluster sample), to check whether our detailed modellings of the dynamical mass, concentration and thermal gas properties significantly impact the final constraints. If so, this would indicate that previous constraints inferred using similar samples may have been subject to biases. Another interesting check will be to infer constraints using both X-ray and SZ selected samples, to check if one can give more powerful constraints than the other (for example, SZ samples are known to have nearly redshift-independent selection functions). Our pipeline can then be applied to much larger catalogues from ongoing surveys, such as the SZ AdvACT cluster survey and the X-ray eROSITA survey, and as mentioned above it will be helpful to make use of weak lensing data from additional surveys like DES and Euclid.

Our models for the enhancements of the dynamical mass and halo concentration in  $f(R)$  gravity are important components of the constraint pipeline. However, these have only been calibrated using DMO simulations. It would therefore be useful to carefully verify these results using cluster data from full-physics hydrodynamical simulations run for  $f(R)$  gravity. We have already carried out preliminary checks of the dynamical mass model for masses  $M_{500} \lesssim 10^{14.5} M_\odot$  using the full-physics SHYBONE simulations (Fig. 4.6). With the baryonic model presented in Chapter 8, which can be used to run much larger full-physics simulations of screened modified gravity, it will now be possible to carry out these checks with many more haloes spanning masses  $10^{13} M_\odot \lesssim M_{500} \lesssim 10^{15} M_\odot$ .

On a similar note, our baryonic model can also be used to run full-physics simulations of nDGP for boxes that are much larger than the current  $62h^{-1}\text{Mpc}$  runs. These could be used for a robust study of the observable-mass scaling relations in nDGP up to  $M_{500} \sim 10^{15} M_\odot$ , and they could also be used to check the accuracy of our general models for the enhancements of the concentration and the HMF (which were calibrated using DMO simulations) in the presence of full baryonic physics.

Finally, in our study of the tSZ and kSZ angular power spectra in  $f(R)$  gravity and nDGP, we were unable to study larger angular scales ( $l \lesssim 500$ ), again due to the relatively small box size of the SHYBONE simulations. This can potentially be rectified by using the larger simulations discussed above. Our results in Chapter 7 indicate that the gravity model differences are less sensitive to baryonic physics at larger scales. The precise scales at which the SZ power spectra can be used to reliably probe gravity should be confirmed by using a combination of full-physics and non-radiative simulations, as well as a range of baryonic physics parameters. This will also help to better-assess the potential constraining power from current and upcoming CMB experiments.

# Modelling the dynamical mass of haloes in $f(R)$ gravity

## A.1 Weighted fitting of tanh curve

In Sec. 2.4.2, we discussed and compared, for a few selected cases, two schemes to fit the  $M_{\text{dyn}}/M_{\text{true}}$  mass ratio data using a tanh curve. We found that, although the weighted and unweighted fitting schemes give broadly consistent results, the latter scheme, by assuming that all data points have the same error, leads to fitted tanh curves that have better visual agreement with the data points. This is because in some snapshots the data for the median  $M_{\text{dyn}}/M_{\text{true}}$  ratio has big disparities in the uncertainties because there are few high-mass haloes due to box size constraints, or because the ratio data in screened and unscreened regimes shows too little variation. The estimated median ratio values therein are not biased because of this, and so we presented our main results (see Sec. 2.5) using the unweighted scheme. This gives all bins equal weight regardless of the large disparities in the uncertainty, allowing the fitted curve to more easily go through the data points. However, one could still argue that the strong variation of median  $M_{\text{dyn}}/M_{\text{true}}$  ratio uncertainties in the different mass bins is at least partly physical (e.g., in the completely unscreened regime there is intrinsically little uncertainty in the ratio). Therefore here we present our results from using the weighted approach, which show that the choice of method does not have a



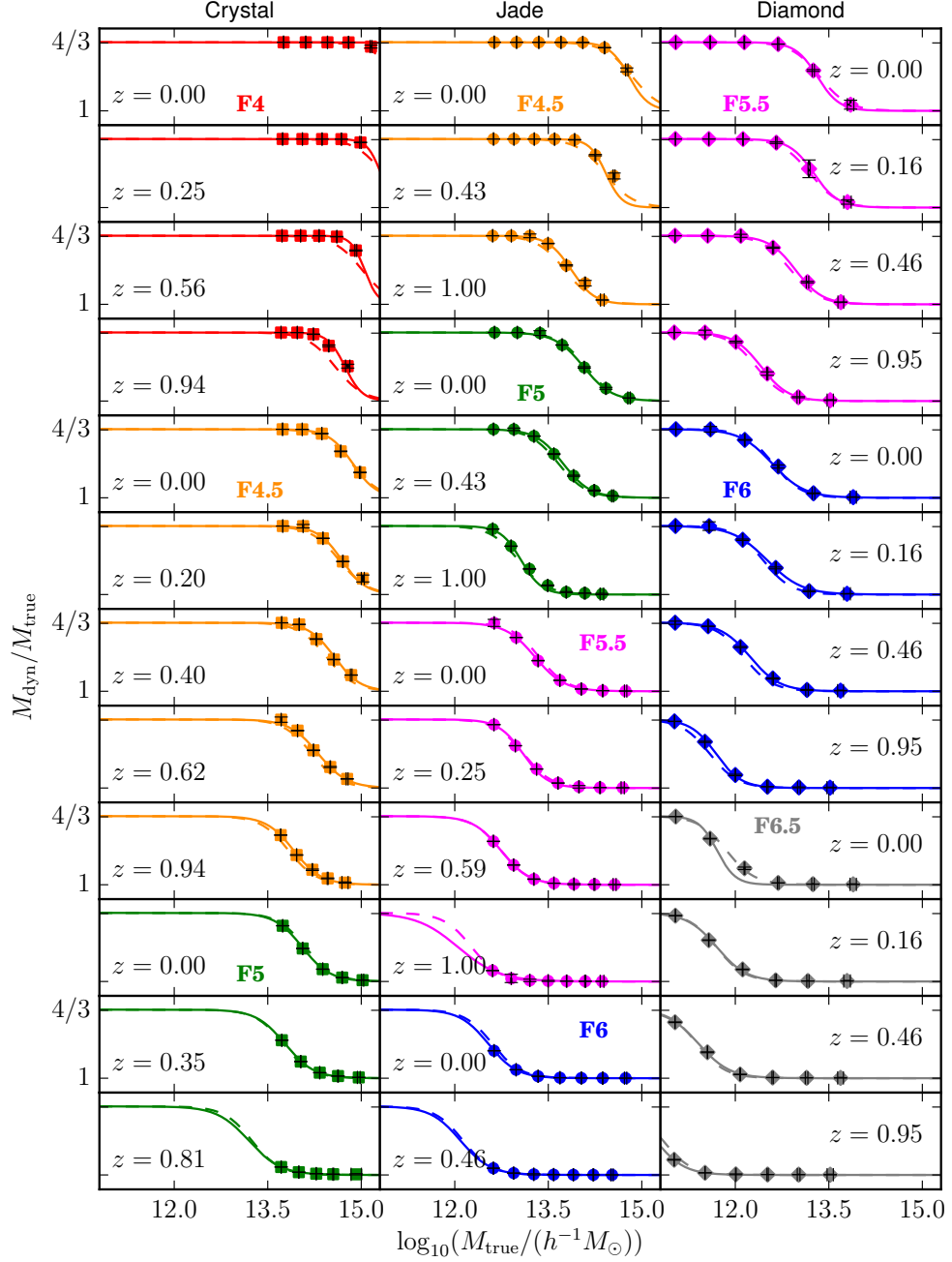


Figure A.1: Dynamical mass to lensing mass ratio as a function of the lensing mass for F4 (red), F4.5 (orange), F5 (green), F5.5 (magenta), F6 (blue) and F6.5 (grey) at various redshifts as annotated. The data has been generated using the Crystal (left column), Jade (middle column) and Diamond (right column) modified ECOSMOG simulations (see Table 2.1). The data points, corresponding to mass bins represented by their median ratio and mean mass, and their one standard deviation error bars are produced using jackknife resampling. Jackknife errors less than  $10^{-4}$  are replaced with half of the range between the 16th and 84th percentiles. *Solid line*: Eq. (2.17) with  $p_1$  and  $p_2$  determined by weighted least squares fitting for the given snapshot; *Dashed line*: Eq. (2.17) with best-fit constant  $p_1$  result ( $p_1 = 2.23$ ) and linear  $p_2$  result (Eq. (A.1)) from Figs. A.3 and A.2, respectively.

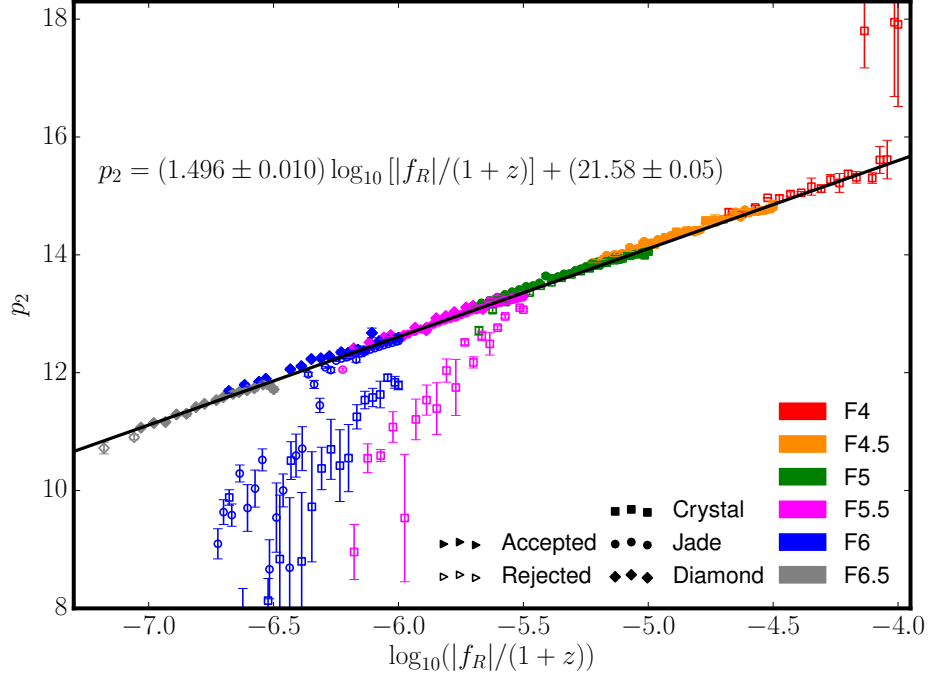


Figure A.2: Parameter  $p_2$  in Eq. (2.17) plotted as a function of the background scalar field at redshift  $z$ ,  $\bar{f}_R(z)$ , divided by  $(1+z)$ , for several present day field strengths  $f_{R0}$  (see legends) of Hu-Sawicki  $f(R)$  gravity with  $n = 1$ .  $p_2$  is measured via a weighted least squares optimization of Eq. (2.17) to data from modified ECOSMOG simulations, described by Table 2.1, at simulation snapshots with redshift  $z < 1$ .  $\bar{f}_R(z)$  is calculated for each snapshot using Eq. (1.15). The trend line has been produced via a weighted least squares linear fit, using the one standard deviation error bars, of the solid data points, which correspond to snapshots for which the mass bins contain at least half of the median mass ratio range 1 to 4/3. The hollow data does not meet this criterion, so is deemed unreliable and neglected from the fit, which is given by Eq. (A.1).

significant effect on the final results, namely on the constant and linear fits of  $p_1$  and  $p_2$  respectively.

To check the reliability of the weighted fit across all redshifts, field strengths and simulations, Fig. A.1 has been produced, which is analogous to Fig. 2.3 and covers the same snapshots. The solid line trends are the weighted fits of the simulation data at the given snapshots, and in general these show very good agreement with the simulation data. However the disparities in the sizes of the error bars now have a stronger impact on the fit and significant deviation from the simulation data is observed for several snapshots, including the Crystal F4  $z = 0.00$ , the Jade F4.5 snapshots and Diamond F6.5  $z = 0.00$ .

The results for  $p_2$ , produced through the weighted approach, are shown in Fig. A.2.

The lowest redshift snapshots of F4 are now observed to peel off from the linear trend due to the large disparities in the uncertainties of the mass bin data, as discussed in Sec. 2.4.2 (see Fig. 2.2). The disparity in uncertainty in part results from the limited number of high-mass haloes which could be screened in F4; such massive haloes are very rare and the only way to resolve this issue is to have a simulation with a much larger box size. However, as is shown in Fig. 2.4 in Sec. 2.5, using unweighted least squares to measure  $p_2$  has the effect of smoothing out the F4 data for  $p_2$ , although this does not reduce the general scatter in F4. In general the data is more scattered across all models in Fig. A.2 than in Fig. 2.4, although for F4 there is now a more even scatter, with the data showing better alignment with the trend line than for the unweighted case.

The criterion for the rejection of the measured  $p_2$  values is the same as for the unweighted approach, and so the outliers for low-redshift F4 in Fig. A.2 do not affect the linear fit of this data. As can be seen from Fig. A.2, all of the solid data points, which meet this criterion, lie along a clear linear trend, while the hollow data points of F5.5 and F6 are all observed to peel off from this trend in a similar manner to the data in Fig. 2.4. The result of the linear fit, using the one standard deviation error bars, is:

$$p_2 = (1.496 \pm 0.010) \log_{10} \left( \frac{|\bar{f}_R|}{1+z} \right) + (21.58 \pm 0.05). \quad (\text{A.1})$$

Agreement of the slope with the theoretical prediction of 1.5 from Eq. (2.19) is excellent. The best-fit linear parameters of  $1.496 \pm 0.010$  and  $21.58 \pm 0.05$  also show strong agreement with the linear fit of the unweighted results (see Fig. 2.4), implying that the choice of whether to use weighted or unweighted least squares fitting of Eq. (2.17) is not of particular importance as far as  $p_2$  is concerned.

The results for  $p_1$ , measured via weighted least squares, are given in Fig. A.3, which is plotted on the same axes range as Fig. 2.5. Once again, the same selection criteria is used as for the unweighted least squares approach, and the hollow data points are left out of any fitting. The points are now significantly more scattered, and all models now contain notable outliers which include several of the solid data points.

Taking F6.5  $z = 0.00$  as an example, we can clearly see from Fig. A.1 that the width of the mass transition has been under-estimated by the weighted least squares approach,

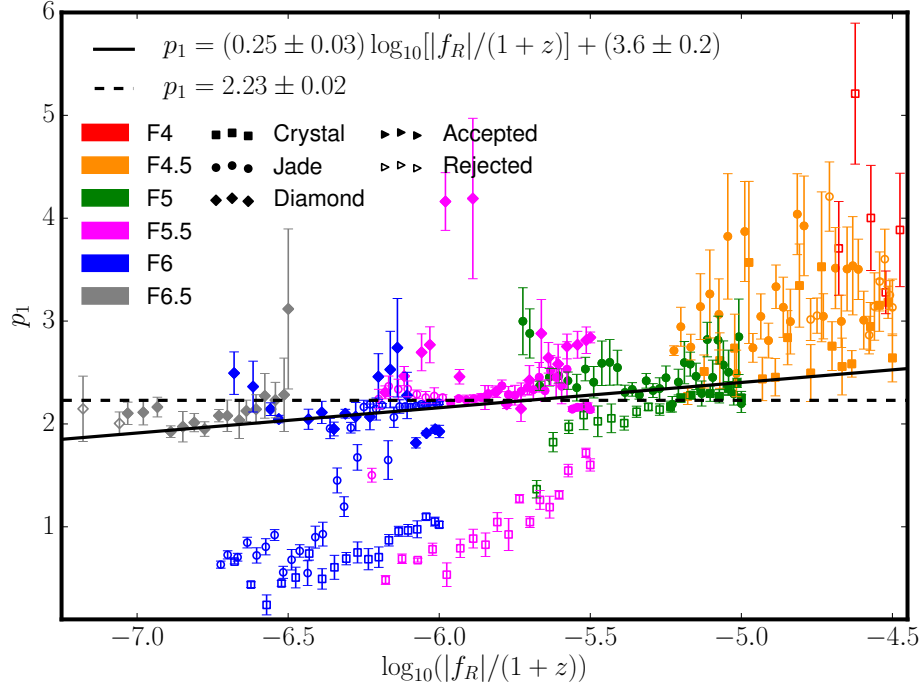


Figure A.3: Parameter  $p_1$  in Eq. (2.17) plotted as a function of the background scalar field at redshift  $z$ ,  $\bar{f}_R(z)$ , divided by  $(1+z)$ , for several present day field strengths  $f_{R0}$  (see legends) of Hu-Sawicki  $f(R)$  gravity with  $n = 1$ .  $p_1$  is measured via a weighted least squares optimization of Eq. (2.17) to data from modified ECOSMOG simulations, described by Table 2.1, at simulation snapshots with redshift  $z < 1$ .  $\bar{f}_R(z)$  is calculated for each snapshot using Eq. (1.15). Weighted least squares linear (*solid line*) and constant (*dashed line*) fits, using the one standard deviation error bars, of the solid data points, which correspond to snapshots for which the mass bins contain at least half of the median mass ratio range 1 to 4/3, are shown. The hollow data does not meet this criteria, so is deemed unreliable and neglected from the fits, which are given by Eq. (A.2) and  $p_1 = (2.23 \pm 0.02)$  respectively.

probably because of the large error bar on one of the data points lying within the transition. A similar effect applies to the other strongly over-estimated data points in Fig. A.3, and as discussed above this comes down to limitations in using a weighted least squares fit.

The result of the constant fit, which is motivated by the theoretical prediction of Eq. (2.18), using the solid data points only, is  $p_1 = (2.23 \pm 0.02)$ , which is shown by the dashed line. This shows excellent agreement with the constant fit to the unweighted data of Fig. 2.5. Again, a linear model was also fitted, shown by the solid line, and is given by:

$$p_1 = (0.25 \pm 0.03) \log_{10} \left( \frac{|\bar{f}_R|}{1+z} \right) + (3.6 \pm 0.2). \quad (\text{A.2})$$

The gradient is still not in agreement with the prediction of zero. Accounting for higher order effects, e.g., environmental screening and the non-sphericity of haloes, may bring

these results into better agreement with the theoretical predictions; however, since we are interested in an empirical fitting function that can be of practical use, we prefer a simple over a sophisticated theoretical model.

As with the unweighted least squares fitting, the validity of these fits of  $p_1$  and  $p_2$  can be checked through an examination of Fig. A.1. This time the dashed lines are produced using Eq. (2.17) along with the linear fit of  $p_2$  from Fig. A.2 (solid line) and the constant fit of  $p_1$  from Fig. A.3 (dashed line), which are motivated by theory. Agreement is now not quite as strong between the dashed and solid lines as in Fig. 2.3, though still very good for most snapshots shown. Disagreement with the simulation data still exists for F4, which partly results from the lack of high-mass haloes and the flatness of the data in the unscreened regime, as for the unweighted approach. However, in Fig. A.1, disparities in F4 also result from the limitations in the weighted least squares fit in finding agreement with mass bins of large error, and this affects other models as well. Examples include the Jade F4.5 snapshots, Jade F5.5  $z = 1.00$  and Diamond F6.5  $z = 0.00$ . In these panels the theoretical dashed line fits actually show better agreement with the simulation data than the solid lines, as they depend on fits from all snapshots and are therefore effectively not error bar dependent. On the whole, the dashed lines show excellent agreement with the simulation data, providing further validation of the analytical model given by Eqs. (2.17), (2.18) and (2.19), even if agreement is not quite as strong as for the unweighted approach.

## A.2 Consistency tests

As was explained in the main text, the issue of an insufficient mass range is resolved through the use of three simulations with varying resolutions, whereas the use of silver data ensures an extended set of present-day scalar field values from  $|f_{R0}| = 10^{-4}$  right down to  $|f_{R0}| = 10^{-6.5}$ . This allows the theoretical model to be rigorously tested for all present-day field strengths in this range, not just for F4, F5 and F6, for which full simulation data are available.

The validity of using silver data was tested by generating F5 silver data from the Crystal simulation  $\Lambda$ CDM data, to be directly compared with the F5 gold data from the

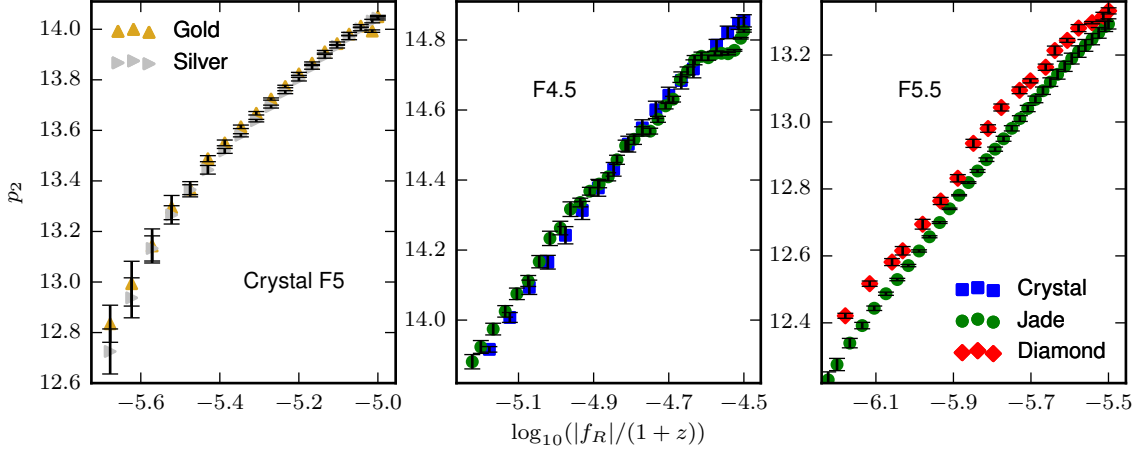


Figure A.4: Parameter  $p_2$  in Eq. (2.17) plotted as a function of the background scalar field at redshift  $z$ ,  $\bar{f}_R(z)$ , divided by  $(1+z)$ , for Hu-Sawicki  $f(R)$  gravity with  $n = 1$ .  $p_2$  is measured via an unweighted least squares optimization of Eq. (2.17) to data from modified ECOSMOG simulations, described by Table 2.1, at simulation snapshots with redshift  $z < 1$ . The one standard deviation error bars are included. *Left to right*: comparison of gold and silver data from the Crystal simulation with present day scalar field value  $|f_{R0}| = 10^{-5}$ ; comparison of the Crystal and Jade data with  $|f_{R0}| = 10^{-4.5}$ ; comparison of the Jade and Diamond data with  $10^{-5.5}$ . The legend in the right plot applies to both the middle and right plots.

same simulation. A comparison of the values of the Eq. (2.17) parameter  $p_2$  is shown in the left panel of Fig. A.4, where the percentage error is measured at around 0.1% for the unweighted approach. This is clearly low enough so that the use of silver data is justified. Physically, this makes sense, because major differences between a full  $f(R)$  simulation (used to generate gold data) and its  $\Lambda$ CDM counterpart (used to generate silver data) include the halo density profile and halo mass, but the difference is generally small enough to not have a strong impact on the scalar field profile. The averaging of the halo mass distribution in the top-hat approximation is shown to be a very good approximation, and further makes the differences in the halo density profiles irrelevant from the point of view of thin-shell modelling.

When combining simulations of different resolutions, the dispersion between these simulations can also lead to a significant systematic source of uncertainty. This can be tested by looking at a few model parameters  $f_{R0}$  for which the mass range necessary to fit  $p_1$  and  $p_2$  as a function of  $\bar{f}_R(z)/(1+z)$  for  $0 \leq z \leq 1$  is offered by simulations of different resolutions. In the middle panel of Fig. A.4 the Crystal and Jade simulations are compared

for F4.5, and found to agree to within an accuracy of 0.3%. A similar test on the Jade and Diamond simulations for F5.5 yielded an error of 0.4-0.8% (right panel of Fig. A.4). These agreements are good enough that the disparity between the results of the simulations is negligible and combination of different simulations is justified. Note that these two checks are also done using the unweighted least squares approach.

A limitation of the current study is that we do not have simulations that allow us to test the fitting functions of  $p_1$  and  $p_2$  for other cosmological parameters, such as  $\Omega_M$  and  $\sigma_8$ , as these are fixed in the original simulations and cannot be changed for producing the silver data. While this is something that would be good to explicitly check in future work, we believe that the excellent agreement between the physically motivated thin-shell modelling and the simulation data, in spite of the approximations employed, has indicated that the theoretical model has successfully captured the essential physics. Therefore we expect the fitting functions we found in this paper to apply to other values of  $\Omega_M$  and  $\sigma_8$  as well. For example, in the paragraph below Eqs. (2.18, 2.19) we have discussed that, according to the thin-shell model,  $p_1$  and  $p_2$  should depend only on Newton's constant  $G$  and the halo mass definition  $\Delta$  (with  $H_0 = 100h \text{ kms}^{-1}\text{Mpc}^{-1}$ ), and in particular they do not depend on cosmological parameters such as  $\Omega_m$  and  $\sigma_8$ . Note that varying  $\Omega_M$  and  $\sigma_8$  will modify the halo abundances and density profiles, and in the check of silver vs. gold data above we have already confirmed that slight changes to these quantities do not affect our fitting functions noticeably.

Another check that is not included in this study is whether the fitting functions work for forms of  $f(R)$  other than Hu-Sawicki as well. While a detailed investigation of this is of interest, we do not find a compelling justification to make substantial effort to include it here, for two reasons. First, as for the case of varying  $\Omega_M$  and  $\sigma_8$ , the effects on the modelling of  $M_{\text{dyn}}/M_{\text{true}}$  through a modified halo abundance and density profile are expected to be small/negligible. Second, the different  $f(R)$  models generally have a different transition from screened to unscreened regimes, though the details of this transition depends on the model itself and its parameters. This indicates that, even though the slope of  $p_2$ , which is 1.5, is expected to remain for general  $f(R)$  models, the intercept of  $p_2$  could be model

dependent. For  $p_1$ , which denotes how the transition from screened to unscreened regimes takes place, the discussion after Eqs. (2.18, 2.19) implies it does not depend on the details of  $f(R)$ , though more explicit checks using simulations are necessary to confirm this or to calibrate its (probably constant) value for general  $f(R)$  models. As mentioned above, it is not feasible to do simulations for all possible models. And nor is this necessary, given that any  $f(R)$  model studied in a cosmological context is phenomenological and not fundamental, and the focus should really be how to get precise stringent constraints on a representative example, which can then be interpreted in the context of general cases.



# A self-consistent pipeline for unbiased constraints of $f(R)$ gravity

## B.1 Modelling the dynamical mass scatter

In Fig. 5.1, the data points show the binned mass ratio scatter as a function of the rescaled logarithmic mass,  $\log_{10}(M_{500}M_{\odot}^{-1}h) - p_2 = \log_{10}(M_{500}/10^{p_2})$ . To generate this, we have evaluated the difference between the actual dynamical mass enhancement and the value predicted by Eq. (5.1) for each halo, and measured the root-mean-square difference within the same mass bins as used to fit Eq. (5.1) in Chapter 2. We have modelled this data using a 6-parameter fitting formula which is made up of two parts. A skewed normal distribution is used to capture the shape of the peak: this includes parameters for the normalisation  $\lambda_s$ , the position  $\mu_s$  and width  $\sigma_s$  with respect to the  $x = \log_{10}(M_{500}/10^{p_2})$  axis, and a parameter  $\alpha$  quantifying the skewness. On its own, this distribution would fall to zero at both low and high  $x$ ; however, we see from Fig. 5.1 that the scatter is slightly greater on average at high  $x$  than at low  $x$ . To account for this, we add on a tanh function with two parameters: an amplitude  $\lambda_t$  and a shift  $y_t$  along the vertical axis. Our full model is then given by:

$$\sigma_{\mathcal{R}} = \frac{\lambda_s}{\sigma_s} \phi(x') \left[ 1 + \operatorname{erf} \left( \frac{\alpha x'}{\sqrt{2}} \right) \right] + (\lambda_t \tanh(x) + y_t), \quad (\text{B.1})$$

where  $x' = (x - \mu_s)/\sigma_s$ .  $\phi(x')$  represents the normal distribution:

$$\phi(x') = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x'^2}{2}\right), \quad (\text{B.2})$$

and  $\text{erf}(x')$  is the error function:

$$\text{erf}(x') = \frac{2}{\sqrt{\pi}} \int_0^{x'} e^{-t^2} dt. \quad (\text{B.3})$$

Since we have many more data points at higher masses than at lower masses in Fig. 5.1, we have used a weighted least squares approach which ensures that different parts of the  $\log_{10}(M_{500}/10^{p_2})$  range have an equal contribution to the fitting of Eq. (B.1). To do this, we have split the rescaled mass range into 10 equal-width bins and counted the number,  $N_i$ , of data points within each bin  $i$ . In the least squares fitting, each data point is then weighted by  $1/N_i$ . This means that points found at lower masses, where there are fewer data points, are each given a greater weight than points found at higher masses. The resulting best-fit parameter values are:  $\lambda_s = 0.0532 \pm 0.0008$ ,  $\sigma_s = 0.58 \pm 0.03$ ,  $\mu_s = -0.35 \pm 0.03$ ,  $\alpha = 1.09 \pm 0.18$ ,  $\lambda_t = 0.0012 \pm 0.0003$  and  $y_t = 0.0019 \pm 0.0002$ .

## B.2 Mass conversions

The following formula can be used to convert the HMF from mass definition  $M_\Delta$  to a new definition  $M_{\Delta'}$ :

$$n'(M_{\Delta'}) = n(M_\Delta(M_{\Delta'})) \left( \frac{d \ln M_{\Delta'}}{d \ln M_\Delta} \right)^{-1}, \quad (\text{B.4})$$

where  $n'$  is the HMF in the new mass definition and  $n$  is the HMF in the old definition. This requires a relation between the mass definitions. For this, we use the following (Hu and Kravtsov, 2003):

$$\frac{M_\Delta}{M_{200}} = \frac{\Delta}{200} \left( \frac{c_\Delta}{c_{200}} \right)^3, \quad (\text{B.5})$$

where  $c_\Delta$  is the concentration with respect to generic overdensity  $\Delta$ . The latter can be computed from  $c_{200}$  using:

$$\frac{1}{c_\Delta} = x \left[ f_\Delta = \frac{\Delta}{200} f \left( \frac{1}{c_{200}} \right) \right], \quad (\text{B.6})$$

where the function  $f(x)$  is given by:

$$f(x) = x^3 \left[ \ln(1 + x^{-1}) - (1 + x)^{-1} \right]. \quad (\text{B.7})$$

Eq. (B.6) is computed using the inverse of this function. [Hu and Kravtsov \(2003\)](#) provide an analytical formula which can accurately solve this:

$$x(f) = \left[ a_1 f^{2p} + \left( \frac{3}{4} \right)^2 \right]^{-\frac{1}{2}} + 2f, \quad (\text{B.8})$$

where  $p = a_2 + a_3 \ln f + a_4 (\ln f)^2$  and the parameters have values  $a_1 = 0.5116$ ,  $a_2 = -0.4283$ ,  $a_3 = -3.13 \times 10^{-3}$  and  $a_4 = -3.52 \times 10^{-5}$ . The authors state that this formula has  $\sim 0.3\%$  accuracy for galaxy and cluster scales.

### B.3 Test of the constraint pipeline on a stronger $f(R)$ model

For the main results of Chapter 5, we have tested our constraint pipeline using GR and F5 mocks. For the F5 mock (cf. Fig. 5.6), our pipeline produces a marginalised distribution of  $\log_{10} |f_{R0}|$  which peaks close to  $-5$ , but features a long tail extending to  $-7$ , which is the lowest value of  $\log_{10} |f_{R0}|$  considered in this work. As discussed in Sec. 5.4.2, this can be explained by parameter degeneracies, which can make it more difficult to fully distinguish this model from GR.

To check whether our pipeline can successfully distinguish stronger  $f(R)$  models than F5, and whether such models suffer from the same degeneracies, we show, in Fig. B.1, constraints obtained using an F4.5 ( $\log_{10} |f_{R0}| = -4.5$ ) mock along with the F5 results from Fig. 5.6. The F4.5 constraint features smaller contours and a tight peak at  $\log_{10} |f_{R0}| \approx -4.5$  which does not feature long tails towards lower or higher values of  $\log_{10} |f_{R0}|$ . The median and 68% range is given by  $-4.47^{+0.06}_{-0.07}$ , which is in excellent agreement with the fiducial value of  $-4.50$ . This indicates that our pipeline can clearly distinguish different values of  $|f_{R0}|$  and it provides further evidence that it can distinguish  $f(R)$  models from GR in an unbiased manner.

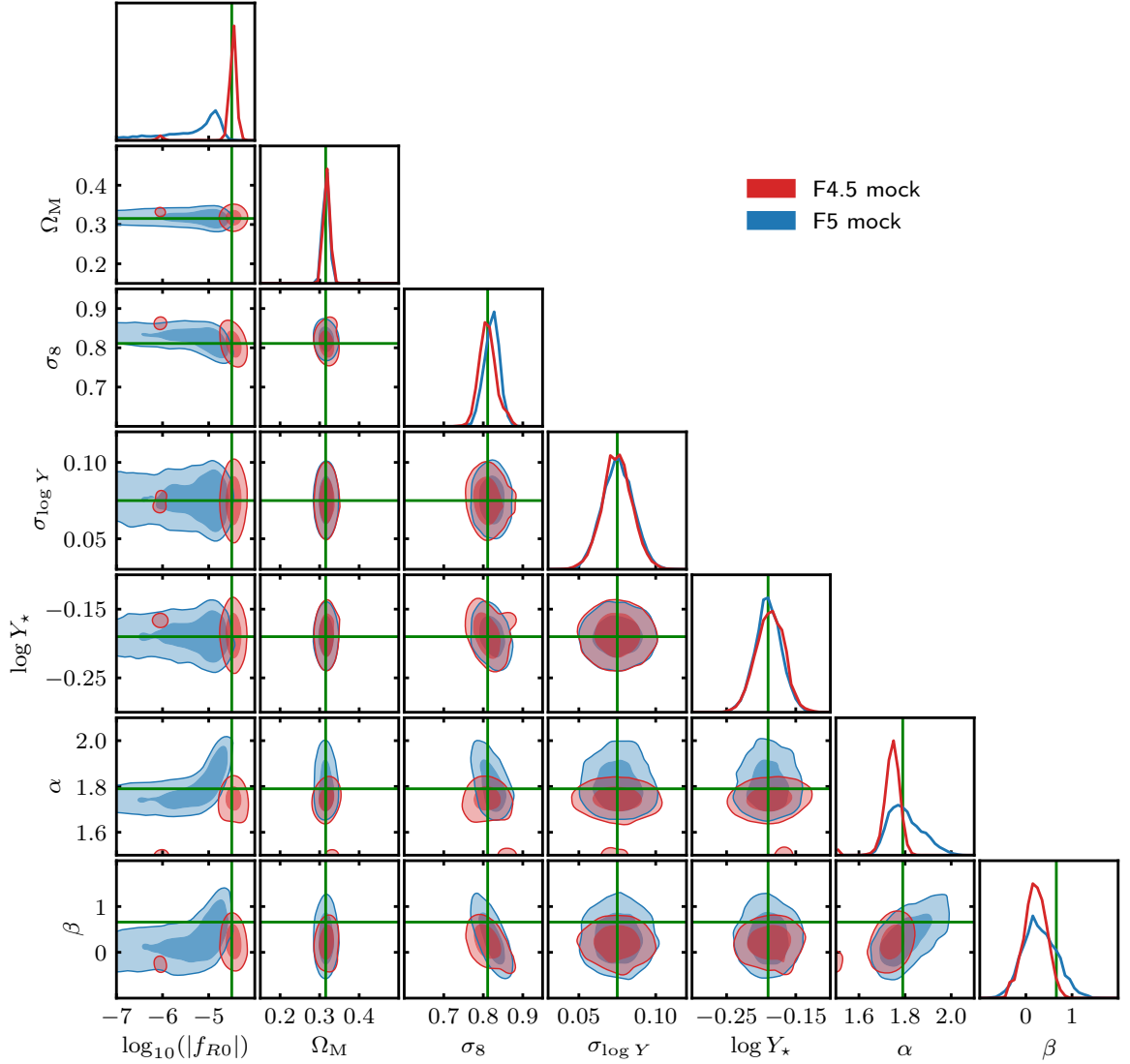


Figure B.1: Parameter constraints obtained by applying our full  $f(R)$  pipeline, as detailed in Sec. 5.3, to F4.5 (red) and F5 (blue) mocks, with observational flux threshold  $Y_{\text{cut}} = 1.5 \times 10^{-5} \text{Mpc}^2$ . The dark and light regions of the contours represent 68% and 95% confidences, respectively. The marginalised distributions of the sampled parameter values are shown in the top panels of each column. The fiducial cosmological parameter values of the F4.5 mock are indicated by the green lines.

---

# Cluster and halo properties in DGP gravity

## C.1 Simulation consistency

In Secs. 6.3.2 and 6.3.3, we combined the halo data from our DMO simulations in order to study the effects of nDGP on the halo concentration and the HMF over a wide mass range. In doing this, it is important to verify that the data from the simulations, which have different resolutions, are consistent. We therefore show, in Figs. C.1 and C.2, the concentration and HMF data, respectively, from each of our DMO simulations for GR and N1.

In Fig. C.1, we show the binned concentration from all four of our DMO simulations, including L1000 which was excluded from our results in Sec. 6.3.2. At redshift 0, where the simulations all have sufficient resolution, the concentration follows a smooth power-law relation as a function of the mass, with the simulations showing excellent agreement at overlapping masses for both GR and N1. The agreement is not as strong at redshifts 1 and 2, where we see, for example, gaps between the L200 (blue) and L500 (green) concentrations. The concentration is slightly underestimated for haloes that are not well-resolved, affecting the data at the low-mass end (close to the lower mass cut of 2000 particles) of the L500 and L1000 data at  $z = 1$  and the L200 and L500 data at  $z = 2$ .

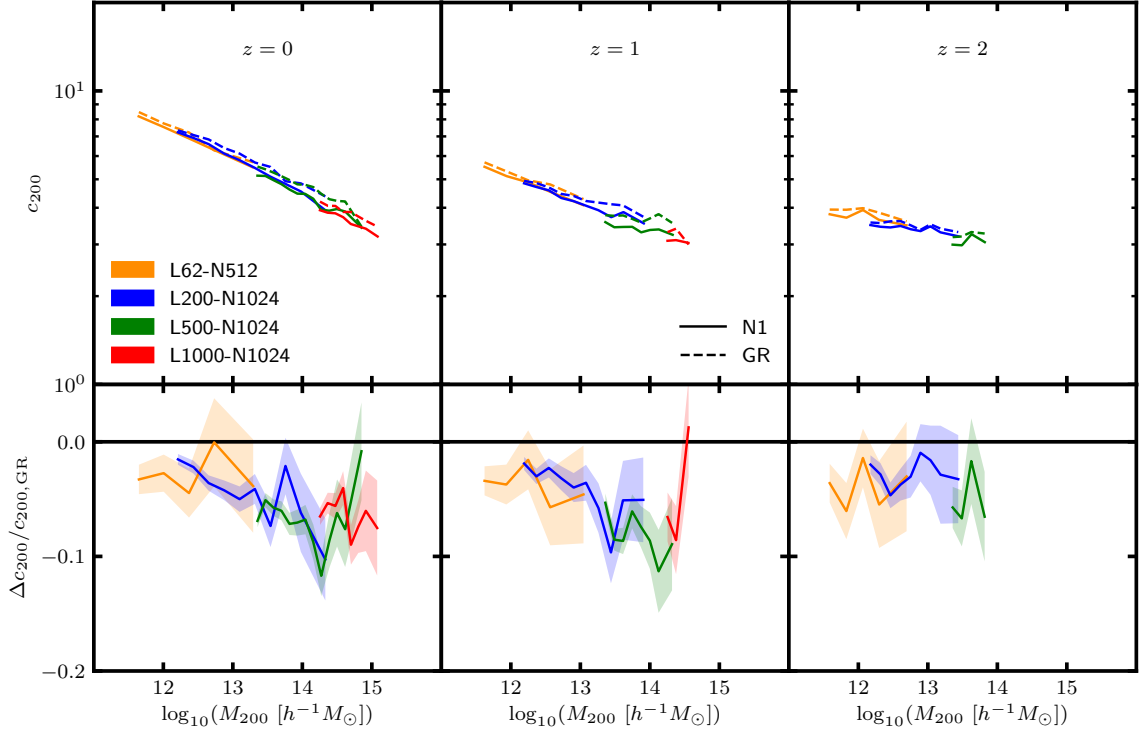


Figure C.1: Median halo concentration (*top row*) and relative difference with respect to GR (*bottom row*) as a function of the mean logarithm of the halo mass at redshifts 0, 1 and 2. The data is generated using the dark-matter-only simulations L62 (*orange*), L200 (*blue*), L500 (*green*) and L1000 (*red*), the specifications of which are given in Table 2.1. Data is shown for GR (*dashed lines*) and the nDGP model N1 (*solid lines*). The shaded regions in the lower panels show the  $1\sigma$  uncertainty in the relative difference.

These resolution issues are potentially problematic for studies of the absolute concentration; however, in this work, we are more interested in the relative difference between the nDGP and GR concentration. From the lower panels of Fig. C.1, it appears that the L62, L200 and L500 simulations give consistent predictions of the relative difference at overlapping masses for each redshift shown. This justifies using a halo mass cut of 2000 particles to study and model the relative difference in Sec. 6.3.2. This cut ensures that there are plenty of haloes at overlapping masses, which is important for the combined binning of the halo data, while it does not give rise to inconsistencies in the relative difference for these three simulations. We decided to exclude the L1000 simulation for a couple of reasons: the concentration suppression does not appear to be fully consistent with the data from the higher-resolution simulations – for example, at  $z = 0$ , the suppression in L1000 appears to be lower than the predictions from L500 at low masses and greater at high masses – and at

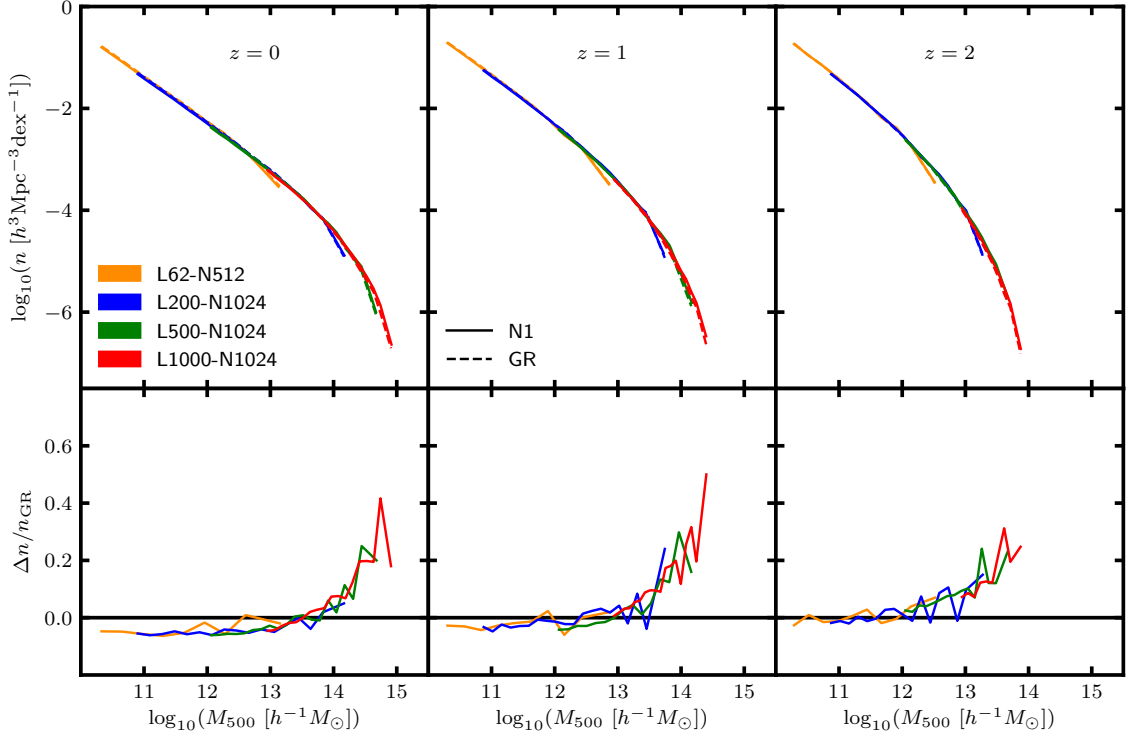


Figure C.2: Halo mass function (*top row*) and its relative difference in nDGP with respect to GR (*bottom row*), as a function of the mean logarithm of the halo mass at redshifts 0, 1 and 2. The data is generated using the dark-matter-only simulations L62 (*orange*), L200 (*blue*), L500 (*green*) and L1000 (*red*), the specifications of which are given in Table 2.1. Data is shown for GR (*dashed lines*) and the nDGP model N1 (*solid lines*).

higher redshifts it does not have many resolved haloes.

In Fig. C.2, we show the binned HMF from DMO simulations. The predictions of the absolute HMF, shown in the top row, agree very well. The HMF is slightly underestimated at the high-mass end of each simulation: this is a natural consequence of the limited box sizes, which causes the high-mass HMF to be incomplete. We note that combining the halo data of the four simulations and summing the total volume in the way that we have described in Sec. 6.3.3 means that incompleteness is only really present for the highest-mass bins shown in Fig. 6.11. The lower panels of Fig. C.2 show the relative differences between GR and N1. The predictions from the four simulations show excellent agreement, again indicating that these simulations can be safely combined.

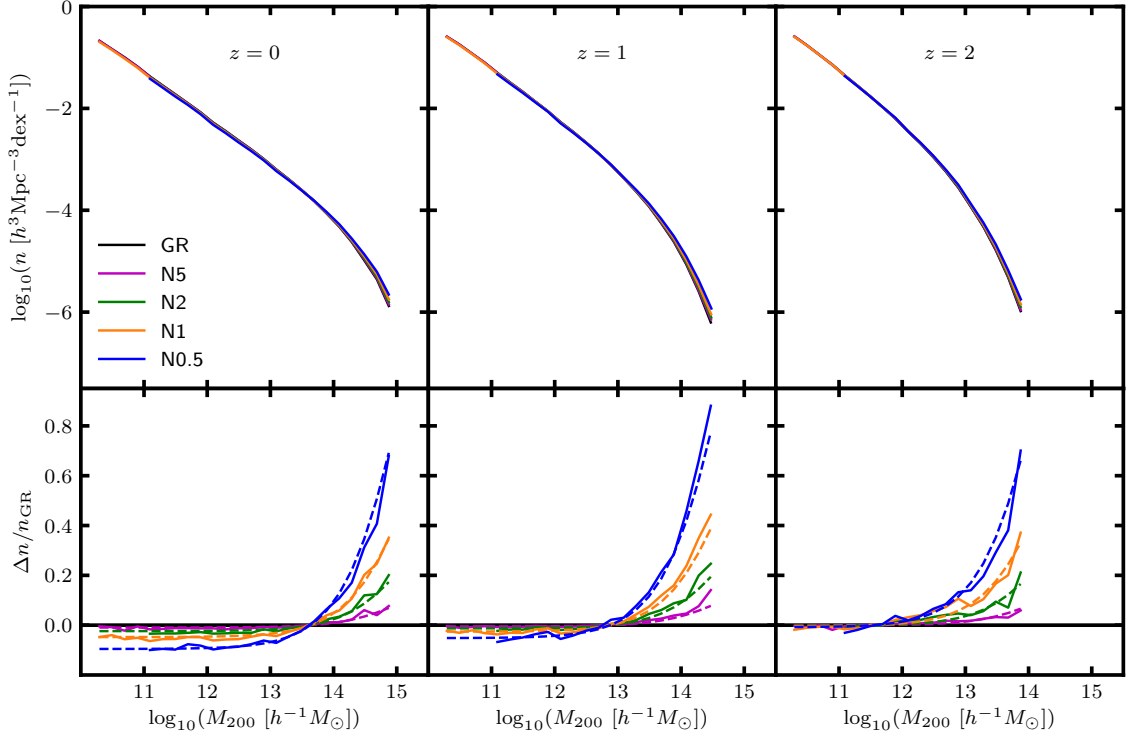


Figure C.3: Halo mass function (*top row*) and its relative difference in nDGP with respect to GR (*bottom row*), as a function of the mean logarithm of the halo mass at redshifts 0, 1 and 2. The results shown are similar to Fig. 6.11; however, here we use mass definition  $M_{200}$  instead of  $M_{500}$ , and the dashed lines show the predictions from the model given by Eqs. (6.4, C.1).

## C.2 $M_{200}$ mass function

In Sec. 6.3.3, we presented our results and model for the nDGP HMF in terms of the  $M_{500}$  mass definition. For completeness, we also show, in Fig. C.3, the HMF in terms of the  $M_{200}$  mass definition. This has again been calculated by combining the haloes from all four DMO simulations, although here we impose a lower mass threshold of 100 particles within the radius  $R_{200}$  rather than  $R_{500}$ . We use the same set of logarithmic mass bins (with fixed width 0.2) and again show all bins that contain at least 100 haloes.

The results in Fig. C.3 are very similar to Fig. 6.11, with the nDGP fifth force suppressing the HMF at lower masses and enhancing the HMF at higher masses. Therefore, we are able to use the same fitting formula to model the relative difference. Replacing  $M_{500}$



with  $M_{200}$  in Eq. (6.4), the best-fit parameter are now:

$$\begin{aligned} A(H_0 r_c) &= (0.59 \pm 0.03)(H_0 r_c)^{-1}, \\ B(z) &= (15.22 \pm 0.03) - (0.441 \pm 0.006)z, \\ C(z) &= (0.919 \pm 0.005) + (0.037 \pm 0.003)z. \end{aligned} \tag{C.1}$$

The predictions of this model are also in very good agreement with the simulation measurement. As in the case of  $M_{500}$ , we note that this model should only be used to predict the HMF within the mass range  $10^{11} h^{-1} M_\odot \lesssim M_{200} \lesssim M_{\max}(z)$ , where  $M_{\max}(z)$  is the maximum mass used for the calibration. For definition  $M_{200}$ , this can be estimated using:

$$\log_{10} \left( M_{\max} M_\odot^{-1} h \right) = 14.93 - 0.52z. \tag{C.2}$$

---

# Realistic galaxy formation simulations to study clusters in modified gravity

In Chapter 8, we presented our new baryonic model for low-resolution, full-physics cosmological simulations. In particular, we focused on the changes that we made to the IllustrisTNG model and described only a small subset of the  $\sim 200$  calibration runs. In this appendix, we will provide a more detailed description of the calibration procedure, including an outline of our simulations and details of the parameter search.

Table D.1 shows the specifications of the simulations used to tune the baryonic model. The primary goal of the tuning was to find a model that can produce sufficient star formation in low-resolution simulations to match galaxy observations. We studied in detail simulations with three different mass resolutions before settling on the resolution of the L68-N256 simulations, which have already been mentioned in Chapter 8.

## D.1 L100-N256 simulations

To start with, we used simulations with a box size of  $100h^{-1}\text{Mpc}$ , containing  $256^3$  dark matter particles and (initially) the same number of gas cells (L100-N256). With an average gas cell mass of  $\sim 8.3 \times 10^8 h^{-1} M_\odot$ , these have 512 times lower mass resolution than the simulations used to calibrate the fiducial TNG model and the same resolution as the BAHAMAS simulations (McCarthy et al., 2017), which were run using GADGET-3 (Springel,

Specifications	Simulations			
	L100-N256	L86-N256	L68-N256	L136-N512
box size / $h^{-1}\text{Mpc}$	100	86	68	136
DM particle number	$256^3$	$256^3$	$256^3$	$512^3$
$m_{\text{DM}} / h^{-1}M_{\odot}$	$4.29 \times 10^9$	$2.73 \times 10^9$	$1.35 \times 10^9$	$1.35 \times 10^9$
$m_{\text{gas}} / h^{-1}M_{\odot}$	$8.3 \times 10^8$	$5.3 \times 10^8$	$2.6 \times 10^8$	$2.6 \times 10^8$
number of runs	$\sim 100$	$\sim 60$	$\sim 50$	3

Table D.1: Specifications of the AREPO simulations that have been used to tune our baryonic model. These are labelled L100-N256, L86-N256, L68-N256 and L136-N512, according to their box size and dark matter particle number (we note that there are initially the same number of gas cells as dark matter particles). The simulations have all been run with standard gravity (GR).

2005) rather than AREPO. We ran  $\sim 100$  simulations at this resolution, varying the following baryonic parameters: the threshold gas density for star formation  $\rho_{\star}$  (see Sec. 8.2.1.3) was varied in the range  $[0.00, 0.13] \text{ cm}^{-3}$ ; the parameter  $\bar{e}_{\text{w}}$  controlling the stellar wind energy (see Eq. (8.1)) was varied in the range  $[0.0, 3.6]$ ; the parameters  $\kappa_{\text{w}}$  and  $v_{\text{w},\text{min}}$  controlling the stellar wind speed (see Eq. (8.2)) were varied over ranges  $[0.0, 29.6]$  and  $[0, 500] \text{ kms}^{-1}$ , respectively; and the black hole radiative efficiency  $\epsilon_{\text{r}}$  (see Sec. 8.2.1.4) was varied in the range  $[0.02, 0.20]$ . We also tested gravitational softening lengths in the range  $1/40$  to  $1/10$  times the mean inter-particle separation.

These runs provided a very useful insight into the effects of changing each parameter, however all of the tested parameter combinations at the L100-N256 resolution resulted in insufficient star formation within haloes of mass  $M_{200} \lesssim 10^{13} M_{\odot}$ , and at higher halo masses it was difficult to simultaneously match observations for different galaxy properties. For example, parameter combinations which yielded a sufficiently high stellar mass function typically resulted in the stellar mass fraction being overestimated, and in order to match the SFRD observations it was necessary to set either  $\rho_{\star}$  or the stellar wind energy close to zero. We therefore decided to look at higher resolutions.

## D.2 L86-N256 simulations

Keeping the dark matter particle number (and initial gas cell number) unchanged, we initially reduced the box size to  $86h^{-1}\text{Mpc}$  (L86-N256), and executed  $\sim 60$  runs at this

higher resolution. Our best model used a gravitational softening length of  $1/20$  times the mean inter-particle separation and the following parameter combination:  $\rho_\star = 0.05\text{cm}^{-3}$ ,  $\bar{e}_w = 0.5$ ,  $\kappa_w = 2$ ,  $v_{w,\text{min}} = 200\text{ km s}^{-1}$  and  $\epsilon_r = 0.15$ . This gave a reasonable match with high-mass observations of the stellar mass fraction ( $M_{200} \gtrsim 2 \times 10^{12} h^{-1} M_\odot$ ) and stellar mass function ( $M_\star \gtrsim 10^{11} h^{-1} M_\odot$ ), however the agreement was still poor at lower masses and the SFRD was significantly underestimated for redshifts  $z \gtrsim 2$ . We made some efforts to rectify this. For example, we switched off feedback entirely by setting the wind energy to zero and preventing the formation of black hole particles, and we tried using much lower values of  $\rho_\star$ . While these efforts resulted in more star formation at lower masses, it was still not enough to match observational data, while at higher masses and lower redshifts there was far too much star formation.

### D.3 L68-N256 and L136-N512 simulations

We finally settled on the  $68h^{-1}\text{Mpc}$  (L68-N256) box size, where we ran a further  $\sim 50$  runs to calibrate the final model presented in Chapter 8. Our final model, with  $\rho_\star = 0.08\text{cm}^{-3}$ ,  $\bar{e}_w = 0.5$ ,  $\epsilon_r = 0.22$  and a softening length of  $1/20$  times the mean inter-particle separation, is able to produce sufficient star formation at lower masses and higher redshifts than the above L86-N256 model, and only requires changes to three of the TNG model parameters (we take the TNG values for  $\kappa_w$  and  $v_{w,\text{min}}$ ). The predictions from five of the L68-N256 runs are shown in Fig. 8.1 to illustrate the effects of each of the changes to the fiducial TNG parameter values.

In order to assess the effects of sample variance, we also ran three simulations with an increased box size of  $136h^{-1}\text{Mpc}$  (L136-N512) and the same mass resolution as the L68-N256 runs. The results from these simulations, which were run using our three most promising baryonic models, indicated that the stellar mass fraction and stellar mass function are slightly reduced in the larger box. This is why we selected the above model, even though it slightly overestimates the stellar mass function in Fig. 8.1.

---

# Bibliography

- K. N. Abazajian et al. CMB-S4 Science Book, First Edition. *ArXiv e-prints*, 10 2016.
- T. Abbott et al. The Dark Energy Survey: more than dark energy – an overview. *Mon. Not. Roy. Astron. Soc.*, 460(2):1270–1299, 2016. doi: 10.1093/mnras/stw641.
- I. Achitouv, M. Baldi, E. Puchwein, and J. Weller. Imprint of  $f(R)$  gravity on nonlinear structure formation. *Phys. Rev. D*, 93(10):103522, 2016. doi: 10.1103/PhysRevD.93.103522.
- R. Adam et al. Euclid preparation III. Galaxy cluster detection in the wide photometric survey, performance and algorithm selection. *Astron. Astrophys.*, 627(627):A23, 2019. doi: 10.1051/0004-6361/201935088.
- P. Ade et al. Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts. *Astron. Astrophys.*, 594:A24, 2016. doi: 10.1051/0004-6361/201525833.
- P. Ade et al. The Simons Observatory: Science goals and forecasts. *J. Cosmo. Astropart. Phys.*, 02:056, 2019. doi: 10.1088/1475-7516/2019/02/056.
- P. A. R. Ade et al. Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.*, 571:A16, 2014. doi: 10.1051/0004-6361/201321591.
- N. Aghanim et al. Planck 2015 results. XXII. A map of the thermal Sunyaev-Zeldovich effect. *Astron. Astrophys.*, 594:A22, 2016. doi: 10.1051/0004-6361/201525826.

- N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. doi: 10.1051/0004-6361/201833910.
- L. Anderson et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample. *Mon. Not. Roy. Astron. Soc.*, 427(4):3435–3467, Dec. 2012. doi: 10.1111/j.1365-2966.2012.22066.x.
- P. Arnalte-Mur, W. A. Hellwing, and P. Norberg. Real- and redshift-space halo clustering in  $f(R)$  cosmologies. *Mon. Not. Roy. Astron. Soc.*, 467(2):1569–1585, 2017. doi: 10.1093/mnras/stx196.
- C. Arnold, E. Puchwein, and V. Springel. Scaling relations and mass bias in hydrodynamical  $f(R)$  gravity simulations of galaxy clusters. *Mon. Not. Roy. Astron. Soc.*, 440(1):833–842, 2014. doi: 10.1093/mnras/stu332.
- C. Arnold, V. Springel, and E. Puchwein. Zoomed cosmological simulations of Milky Way-sized haloes in  $f(R)$  gravity. *Mon. Not. Roy. Astron. Soc.*, 462:1530–1541, Oct. 2016. doi: 10.1093/mnras/stw1708.
- C. Arnold, P. Fosalba, V. Springel, E. Puchwein, and L. Blot. The modified gravity light-cone simulation project – I. Statistics of matter and halo distributions. *Mon. Not. Roy. Astron. Soc.*, 483(1):790–805, 2019a. doi: 10.1093/mnras/sty3044.
- C. Arnold, M. Leo, and B. Li. Realistic simulations of galaxy formation in  $f(R)$  modified gravity. *Nature Astron.*, 2019b. doi: 10.1038/s41550-019-0823-y.
- E. Artis, J.-B. Melin, J. G. Bartlett, and C. Murray. Impact of the calibration of the Halo Mass Function on galaxy cluster number count cosmology. *Astron. Astrophys.*, 649:A47, 2021. doi: 10.1051/0004-6361/202140293.
- I. K. Baldry et al. Galaxy And Mass Assembly (GAMA): the galaxy stellar mass function at  $z < 0.06$ . *Mon. Not. Roy. Astron. Soc.*, 421(1):621–634, Mar. 2012. doi: 10.1111/j.1365-2966.2012.20340.x.

- X. Barcons, D. Barret, A. Decourchelle, J. W. den Herder, A. C. Fabian, H. Matsumoto, D. Lumb, K. Nandra, L. Piro, R. K. Smith, and R. Willingale. Athena: ESA's X-ray observatory for the late 2020s. *Astronomische Nachrichten*, 338(153):153–158, Mar. 2017. doi: 10.1002/asna.201713323.
- A. Barreira, B. Li, C. M. Baugh, and S. Pascoli. Spherical collapse in Galileon gravity: fifth force solutions, halo mass function and halo bias. *JCAP*, 1311:056, 2013. doi: 10.1088/1475-7516/2013/11/056.
- A. Barreira, B. Li, W. A. Hellwing, L. Lombriser, C. M. Baugh, and S. Pascoli. Halo model and halo properties in Galileon gravity cosmologies. *JCAP*, 1404:029, 2014. doi: 10.1088/1475-7516/2014/04/029.
- A. Barreira, B. Li, E. Jennings, J. Merten, L. King, C. Baugh, and S. Pascoli. Galaxy cluster lensing masses in modified lensing potentials. *Mon. Not. Roy. Astron. Soc.*, 454(4):4085–4102, 2015. doi: 10.1093/mnras/stv2211.
- A. Barreira, A. G. Sánchez, and F. Schmidt. Validating estimates of the growth rate of structure with modified gravity simulations. *Phys. Rev. D*, 94(8):084022, 2016. doi: 10.1103/PhysRevD.94.084022.
- P. S. Behroozi, R. H. Wechsler, and C. Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from  $z = 0$ -8. *Astrophys. J.*, 770(1):57, June 2013. doi: 10.1088/0004-637X/770/1/57.
- B. A. Benson et al. SPT-3G: A Next-Generation Cosmic Microwave Background Polarization Experiment on the South Pole Telescope. *Proc. SPIE Int. Soc. Opt. Eng.*, 9153: 91531P, 2014. doi: 10.1117/12.2057305.
- M. Bernardi, A. Meert, R. K. Sheth, V. Vikram, M. Huertas-Company, S. Mei, and F. Shankar. The massive end of the luminosity and stellar mass functions: Dependence on the fit to the light profile. *Mon. Not. Roy. Astron. Soc.*, 436:697, 2013. doi: 10.1093/mnras/stt1607.

- F. Bianchini and A. Silvestri. Kinetic Sunyaev-Zel'dovich effect in modified gravity. *Phys. Rev.*, D93(6):064026, 2016. doi: 10.1103/PhysRevD.93.064026.
- S. Bocquet et al. Cluster Cosmology Constraints from the 2500 deg<sup>2</sup> SPT-SZ Survey: Inclusion of Weak Gravitational Lensing Data from Magellan and the Hubble Space Telescope. *Astrophys. J.*, 878(1):55, 2019. doi: 10.3847/1538-4357/ab1f10.
- B. Bolliet, B. Comis, E. Komatsu, and J. F. Macías-Pérez. Dark energy constraints from the thermal Sunyaev-Zeldovich power spectrum. *Mon. Not. Roy. Astron. Soc.*, 477(4):4957–4967, 2018. doi: 10.1093/mnras/sty823.
- J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. *Astrophys. J.*, 379:440, 1991. doi: 10.1086/170520.
- B. Bose and K. Koyama. A Perturbative Approach to the Redshift Space Correlation Function: Beyond the Standard Model. *JCAP*, 08:029, 2017. doi: 10.1088/1475-7516/2017/08/029.
- M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson. Resolving cosmic structure formation with the Millennium-II Simulation. *Mon. Not. Roy. Astron. Soc.*, 398(3):1150–1164, Sept. 2009. doi: 10.1111/j.1365-2966.2009.15191.x.
- P. Brax and P. Valageas. Impact on the power spectrum of Screening in Modified Gravity Scenarios. *Phys. Rev.*, D88(2):023527, 2013. doi: 10.1103/PhysRevD.88.023527.
- P. Brax, A.-C. Davis, and B. Li. Modified Gravity Tomography. *Phys. Lett.*, B715:38–43, 2012a. doi: 10.1016/j.physletb.2012.08.002.
- P. Brax, A.-C. Davis, B. Li, and H. A. Winther. A Unified Description of Screened Modified Gravity. *Phys. Rev.*, D86:044015, 2012b. doi: 10.1103/PhysRevD.86.044015.
- A. M. C. L. Brun, I. G. McCarthy, J. Schaye, and T. J. Ponman. Towards a realistic population of simulated galaxy groups and clusters. *Mon. Not. Roy. Astron. Soc.*, 441(2):1270–1290, 2014. doi: 10.1093/mnras/stu608.



- J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes. Evolution, scatter, and environment. *Mon. Not. Roy. Astron. Soc.*, 321:559–575, 2001. doi: 10.1046/j.1365-8711.2001.04068.x.
- M. Cataneo, D. Rapetti, L. Lombriser, and B. Li. Cluster abundance in chameleon  $f(R)$  gravity I: toward an accurate halo mass function prediction. *JCAP*, 1612(12):024, 2016. doi: 10.1088/1475-7516/2016/12/024.
- M. Cataneo, L. Lombriser, C. Heymans, A. Mead, A. Barreira, S. Bose, and B. Li. On the road to per-cent accuracy: nonlinear reaction of the matter power spectrum to dark energy and modified gravity. *arXiv e-prints*, art. arXiv:1812.05594, Dec. 2018.
- M. Cataneo et al. New constraints on  $f(r)$  gravity from clusters of galaxies. *Phys. Rev. D*, 92:044009, Aug 2015. doi: 10.1103/PhysRevD.92.044009. URL <https://link.aps.org/doi/10.1103/PhysRevD.92.044009>.
- M. Cautun, E. Paillas, Y.-C. Cai, S. Bose, J. Armijo, B. Li, and N. Padilla. The Santiago–Harvard–Edinburgh–Durham void comparison – I. SHEDding light on chameleon gravity tests. *Mon. Not. Roy. Astron. Soc.*, 476(3):3195–3217, 2018. doi: 10.1093/mnras/sty463.
- K. C. Chan and R. Scoccimarro. Large-Scale Structure in Brane-Induced Gravity II. Numerical Simulations. *Phys. Rev. D*, 80:104005, 2009. doi: 10.1103/PhysRevD.80.104005.
- W. Cui et al. The Three Hundred project: a large catalogue of theoretically modelled galaxy clusters for cosmological and astrophysical applications. *Mon. Not. Roy. Astron. Soc.*, 480(3):2898–2915, Nov. 2018. doi: 10.1093/mnras/sty2111.
- A. De Felice and S. Tsujikawa.  $f(R)$  theories. *Living Rev. Rel.*, 13:3, 2010. doi: 10.12942/lrr-2010-3.
- T. de Haan et al. Cosmological Constraints from Galaxy Clusters in the 2500 square-degree SPT-SZ Survey. *Astrophys. J.*, 832(1):95, 2016. doi: 10.3847/0004-637X/832/1/95.

- I. De Martino.  $f(R)$  -gravity model of the Sunyaev-Zeldovich profile of the Coma cluster compatible with Planck data. *Phys. Rev. D*, 93(12):124043, 2016. doi: 10.1103/PhysRevD.93.124043.
- C. Deffayet, G. Esposito-Farese, and A. Vikman. Covariant Galileon. *Phys. Rev.*, D79:084003, 2009. doi: 10.1103/PhysRevD.79.084003.
- A. Del Popolo, F. Pace, and D. F. Mota. Mass-temperature relation in  $\Lambda$ CDM and modified gravity. *Phys. Rev. D*, 100(2):024013, 2019. doi: 10.1103/PhysRevD.100.024013.
- G. Despali, C. Giocoli, R. E. Angulo, G. Tormen, R. K. Sheth, G. Baso, and L. Moscardini. The universality of the virial halo mass function and models for non-universality of other halo definitions. *Mon. Not. Roy. Astron. Soc.*, 456(3):2486–2504, 2016. doi: 10.1093/mnras/stv2842.
- R. D’Souza, S. Vegetti, and G. Kauffmann. The massive end of the stellar mass function. *Mon. Not. Roy. Astron. Soc.*, 454(4):4027–4036, Dec. 2015. doi: 10.1093/mnras/stv2234.
- A. R. Duffy, J. Schaye, S. T. Kay, and C. Dalla Vecchia. Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology. *Mon. Not. Roy. Astron. Soc.*, 390:L64, 2008. doi: 10.1111/j.1745-3933.2008.00537.x. [Erratum: *Mon. Not. Roy. Astron. Soc.* 415, L85(2011)].
- A. A. Dutton and A. V. Macciò. Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles. *Mon. Not. Roy. Astron. Soc.*, 441(4):3359–3374, 2014. doi: 10.1093/mnras/stu742.
- G. Dvali, G. Gabadadze, and M. Porrati. 4d gravity on a brane in 5d minkowski space. *Physics Letters B*, 485(1):208 – 214, 2000. ISSN 0370-2693. doi: [https://doi.org/10.1016/S0370-2693\(00\)00669-9](https://doi.org/10.1016/S0370-2693(00)00669-9). URL <http://www.sciencedirect.com/science/article/pii/S0370269300006699>.
- D. J. Eisenstein et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophys. J.*, 633:560–574, 2005. doi: 10.1086/466512.

- J. Erler, K. Basu, J. Chluba, and F. Bertoldi. Planck’s view on the spectrum of the Sunyaev–Zeldovich effect. *Mon. Not. Roy. Astron. Soc.*, 476(3):3360–3381, 2018. doi: 10.1093/mnras/sty327.
- D. Fabjan, S. Borgani, E. Rasia, A. Bonafede, K. Dolag, and L. Murante, G. and Tornatore. X-ray mass proxies from hydrodynamic simulations of galaxy clusters - I. *Mon. Not. Roy. Astron. Soc.*, 416:801, 2011. doi: 10.1111/j.1365-2966.2011.18497.x.
- B. Falck, K. Koyama, G.-b. Zhao, and B. Li. The Vainshtein Mechanism in the Cosmic Web. *JCAP*, 07:058, 2014. doi: 10.1088/1475-7516/2014/07/058.
- B. Falck, K. Koyama, and G.-B. Zhao. Cosmic Web and Environmental Dependence of Screening: Vainshtein vs. Chameleon. *JCAP*, 07:049, 2015. doi: 10.1088/1475-7516/2015/07/049.
- B. Falck, K. Koyama, G.-B. Zhao, and M. Cautun. Using Voids to Unscreen Modified Gravity. *Mon. Not. Roy. Astron. Soc.*, 475(3):3262–3272, 2018. doi: 10.1093/mnras/stx3288.
- A. Finoguenov et al. 4MOST Consortium Survey 5: eROSITA Galaxy Cluster Redshift Survey. *The Messenger*, 175:39–41, 2019. doi: 10.18727/0722-6691/5124.
- D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC Hammer. *PASP*, 125(925):306, Mar. 2013. doi: 10.1086/670067.
- E. George et al. A measurement of secondary cosmic microwave background anisotropies from the 2500-square-degree SPT-SZ survey. *Astrophys. J.*, 799(2):177, 2015. doi: 10.1088/0004-637X/799/2/177.
- S. P. D. Gill, A. Knebe, and B. K. Gibson. The Evolution substructure 1: A New identification method. *Mon. Not. Roy. Astron. Soc.*, 351:399, 2004. doi: 10.1111/j.1365-2966.2004.07786.x.
- A. H. Gonzalez, S. Sivanandam, A. I. Zabludoff, and D. Zaritsky. Galaxy Cluster Baryon Fractions Revisited. *Astrophys. J.*, 778:14, 2013. doi: 10.1088/0004-637X/778/1/14.

- M. Gronke, D. F. Mota, and H. A. Winther. Universal predictions of screened modified gravity on cluster scales. *Astron. Astrophys.*, 583:A123, 2015. doi: 10.1051/0004-6361/201526611.
- M. Gronke, A. Hammami, D. F. Mota, and H. A. Winther. Estimates of cluster masses in screened modified gravity. *Astron. Astrophys.*, 595:A78, 2016. doi: 10.1051/0004-6361/201628644.
- A. Hammami and D. F. Mota. Probing modified gravity via the mass-temperature relation of galaxy clusters. *Astron. Astrophys.*, 598:A132, 2017. doi: 10.1051/0004-6361/201629003.
- M. Hasselfield et al. The Atacama Cosmology Telescope: Sunyaev-Zel’dovich selected galaxy clusters at 148 GHz from three seasons of data. *JCAP*, 1307:008, 2013. doi: 10.1088/1475-7516/2013/07/008.
- J.-h. He and B. Li. Accurate method of modeling cluster scaling relations in modified gravity. *Phys. Rev.*, D93(12):123512, 2016. doi: 10.1103/PhysRevD.93.123512.
- J.-h. He, B. Li, A. J. Hawken, and B. R. Granett. Revisiting the screening mechanism in  $f(R)$  gravity. *Phys. Rev.*, D90(10):103505, 2014. doi: 10.1103/PhysRevD.90.103505.
- J.-h. He, A. J. Hawken, B. Li, and L. Guzzo. Effective Dark Matter Halo catalog in  $f(R)$  gravity. *Phys. Rev. Lett.*, 115(7):071306, 2015. doi: 10.1103/PhysRevLett.115.071306.
- J.-h. He, L. Guzzo, B. Li, and C. M. Baugh. No evidence for modifications of gravity from galaxy motions on cosmological scales. *Nature Astronomy*, 2:967–972, Sept. 2018. doi: 10.1038/s41550-018-0573-2.
- C. Hernández-Aguayo, J. Hou, B. Li, C. M. Baugh, and A. G. Sánchez. Large-scale redshift space distortions in modified gravity theories. *Mon. Not. Roy. Astron. Soc.*, 485(2): 2194–2213, 2019. doi: 10.1093/mnras/stz516.
- C. Hernández-Aguayo, C. Arnold, B. Li, and C. M. Baugh. Galaxy formation in the brane world I: overview and first results. *ArXiv e-prints*, 6 2020.

- M. Hilton et al. The Atacama Cosmology Telescope: A Catalog of  $>4000$  Sunyaev–Zel’dovich Galaxy Clusters. *Astrophys. J. Suppl.*, 253(1):3, 2021. doi: 10.3847/1538-4365/abd023.
- G. Hinshaw et al. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *Astrophys. J. Suppl.*, 208(2):19, Oct. 2013. doi: 10.1088/0067-0049/208/2/19.
- H. Hoekstra, R. Herbonnet, A. Muzzin, A. Babul, A. Mahdavi, M. Viola, and M. Cacciato. The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses. *Mon. Not. Roy. Astron. Soc.*, 449(1):685–714, 2015. doi: 10.1093/mnras/stv275.
- B. Horowitz and U. Seljak. Cosmological constraints from thermal Sunyaev–Zeldovich power spectrum revisited. *Mon. Not. Roy. Astron. Soc.*, 469(1):394–400, 2017. doi: 10.1093/mnras/stx766.
- B. Hu, X.-W. Liu, and R.-G. Cai. CHAM: a fast algorithm of modelling non-linear matter power spectrum in the sScreened HALO Model. *Mon. Not. Roy. Astron. Soc.*, 476(1): L65–L68, 2018. doi: 10.1093/mnrasl/sly032.
- W. Hu and A. V. Kravtsov. Sample variance considerations for cluster surveys. *Astrophys. J.*, 584:702–715, 2003. doi: 10.1086/345846.
- W. Hu and I. Sawicki. Models of  $f(R)$  Cosmic Acceleration that Evade Solar-System Tests. *Phys. Rev.*, D76:064004, 2007. doi: 10.1103/PhysRevD.76.064004.
- G. Hurier and F. Lacasa. Combined analysis of galaxy cluster number count, thermal Sunyaev–Zel’dovich power spectrum, and bispectrum. *Astron. Astrophys.*, 604(604):A71, 2017. doi: 10.1051/0004-6361/201630041.
- Z. Ivezić, J. A. Tyson, R. Allsman, J. Andrew, and R. Angel. LSST: from Science Drivers to Reference Design and Anticipated Data Products. *ArXiv e-prints*, 2008.

- B. Jain, V. Vikram, and J. Sakstein. Astrophysical Tests of Modified Gravity: Constraints from Distance Indicators in the Nearby Universe. *Astrophys. J.*, 779:39, 2013. doi: 10.1088/0004-637X/779/1/39.
- F. Jansen et al. XMM-Newton observatory. I. The spacecraft and operations. *Astron. Astrophys.*, 365:L1–L6, Jan. 2001. doi: 10.1051/0004-6361:20000036.
- A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida. The Mass function of dark matter halos. *Mon. Not. Roy. Astron. Soc.*, 321:372, 2001. doi: 10.1046/j.1365-8711.2001.04029.x.
- N. Kaiser. Evolution and clustering of rich clusters. *Mon. Not. Roy. Astron. Soc.*, 222: 323–345, Sept. 1986. doi: 10.1093/mnras/222.2.323.
- J. Khoury and A. Weltman. Chameleon fields: Awaiting surprises for tests of gravity in space. *Phys. Rev. Lett.*, 93:171104, 2004a. doi: 10.1103/PhysRevLett.93.171104.
- J. Khoury and A. Weltman. Chameleon cosmology. *Phys. Rev.*, D69:044026, 2004b. doi: 10.1103/PhysRevD.69.044026.
- J. Khoury and M. Wyman. N-Body Simulations of DGP and Degravitation Theories. *Phys. Rev. D*, 80:064023, 2009. doi: 10.1103/PhysRevD.80.064023.
- S. R. Knollmann and A. Knebe. AHF: Amiga’s Halo Finder. *Astrophys. J. Suppl.*, 182(2): 608–624, June 2009. doi: 10.1088/0067-0049/182/2/608.
- M. Kopp, S. A. Appleby, I. Achitouv, and J. Weller. Spherical collapse and halo mass function in  $f(R)$  theories. *Phys. Rev.*, D88(8):084015, 2013. doi: 10.1103/PhysRevD.88.084015.
- K. Koyama. Ghosts in the self-accelerating universe. *Class. Quant. Grav.*, 24(24):R231–R253, 2007. doi: 10.1088/0264-9381/24/24/R01.
- K. Koyama. Cosmological Tests of Modified Gravity. *Rept. Prog. Phys.*, 79(4):046902, 2016. doi: 10.1088/0034-4885/79/4/046902.

- K. Koyama and F. P. Silva. Nonlinear interactions in a cosmological background in the dvali-gabadadze-porrati braneworld. *Phys. Rev.*, D75:084040, Apr 2007. doi: 10.1103/PhysRevD.75.084040. URL <https://link.aps.org/doi/10.1103/PhysRevD.75.084040>.
- A. Kravtsov, A. Vikhlinin, and A. Meshcheryakov. Stellar mass – halo mass relation and star formation efficiency in high-mass halos. *Astron. Lett.*, 44(1):8–34, 2018. doi: 10.1134/S1063773717120015.
- A. V. Kravtsov, A. Vikhlinin, and D. Nagai. A new robust low-scatter x-ray mass indicator for clusters of galaxies. *Astrophys. J.*, 650:128–136, 2006. doi: 10.1086/506319.
- J. Kwan, S. Bhattacharya, K. Heitmann, and S. Habib. Cosmic Emulation: The Concentration-Mass Relation for  $\Lambda$ CDM Universes. *Astrophys. J.*, 768:123, May 2013. doi: 10.1088/0004-637X/768/2/123.
- T. Y. Lam and B. Li. Excursion set theory for modified gravity: correlated steps, mass functions and halo bias. *Mon. Not. Roy. Astron. Soc.*, 426:3260, 2012. doi: 10.1111/j.1365-2966.2012.21746.x.
- R. Laureijs et al. Euclid Definition Study Report. *ArXiv e-prints*, Oct. 2011.
- A. Lawrence et al. The UKIRT Infrared Deep Sky Survey (UKIDSS). *Mon. Not. Roy. Astron. Soc.*, 379:1599–1617, 2007. doi: 10.1111/j.1365-2966.2007.12040.x.
- A. M. Le Brun, I. G. McCarthy, J. Schaye, and T. J. Ponman. The scatter and evolution of the global hot gas properties of simulated galaxy cluster populations. *Mon. Not. Roy. Astron. Soc.*, 466(4):4442–4469, 2017. doi: 10.1093/mnras/stw3361.
- M. Levi et al. The DESI Experiment, a whitepaper for Snowmass 2013. *ArXiv e-prints*, 2013.
- A. Lewis, A. Challinor, and A. Lasenby. Efficient computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, 538:473–476, 2000. doi: 10.1086/309179.

- B. Li and G. Efstathiou. An extended excursion set approach to structure formation in chameleon models. *Mon. Not. Roy. Astron. Soc.*, 421:1431, 2012. doi: 10.1111/j.1365-2966.2011.20404.x.
- B. Li and T. Y. Lam. Excursion set theory for modified gravity: Eulerian versus Lagrangian environments. *Mon. Not. Roy. Astron. Soc.*, 425:730, 2012. doi: 10.1111/j.1365-2966.2012.21592.x.
- B. Li, G.-B. Zhao, R. Teyssier, and K. Koyama. ECOSMOG: An Efficient Code for Simulating Modified Gravity. *JCAP*, 1201:051, 2012. doi: 10.1088/1475-7516/2012/01/051.
- B. Li, G.-B. Zhao, and K. Koyama. Exploring Vainshtein mechanism on adaptively refined meshes. *JCAP*, 05:023, 2013. doi: 10.1088/1475-7516/2013/05/023.
- B. Li, J.-h. He, and L. Gao. Cluster gas fraction as a test of gravity. *Mon. Not. Roy. Astron. Soc.*, 456(1):146–155, 2016. doi: 10.1093/mnras/stv2650.
- C. Li and S. D. M. White. The distribution of stellar mass in the low-redshift Universe. *Mon. Not. Roy. Astron. Soc.*, 398(4):2177–2187, Oct. 2009. doi: 10.1111/j.1365-2966.2009.15268.x.
- X. Liu et al. Constraining  $f(R)$  Gravity Theory Using Weak Lensing Peak Statistics from the Canada-France-Hawaii-Telescope Lensing Survey. *Phys. Rev. Lett.*, 117(5):051101, 2016. doi: 10.1103/PhysRevLett.117.051101.
- L. Lombriser. A parametrisation of modified gravity on nonlinear cosmological scales. *JCAP*, 1611(11):039, 2016. doi: 10.1088/1475-7516/2016/11/039.
- L. Lombriser, K. Koyama, G.-B. Zhao, and B. Li. Chameleon  $f(R)$  gravity in the virialized cluster. *Phys. Rev. D*, 85:124054, 2012. doi: 10.1103/PhysRevD.85.124054.
- L. Lombriser, B. Li, K. Koyama, and G.-B. Zhao. Modeling halo mass functions in chameleon  $f(R)$  gravity. *Phys. Rev.*, D87(12):123511, 2013. doi: 10.1103/PhysRevD.87.123511.



- L. Lombriser, K. Koyama, and B. Li. Halo modelling in chameleon theories. *JCAP*, 03:021, 2014. doi: 10.1088/1475-7516/2014/03/021.
- L. Lovisari, T. Reiprich, and G. Schellenberger. Scaling Properties of a Complete X-ray Selected Galaxy Group Sample. *Astron. Astrophys.*, 573:A118, 2015. doi: 10.1051/0004-6361/201423954.
- L. Lovisari, S. Ettori, M. Gaspari, and P. A. Giles. Scaling Properties of Galaxy Groups. *Universe*, 7(5):139, May 2021. doi: 10.3390/universe7050139.
- A. D. Ludlow, J. F. Navarro, R. E. Angulo, M. Boylan-Kolchin, V. Springel, C. Frenk, and S. D. M. White. The mass–concentration–redshift relation of cold dark matter haloes. *Mon. Not. Roy. Astron. Soc.*, 441(1):378–388, 2014. doi: 10.1093/mnras/stu483.
- Y.-Z. Ma and G.-B. Zhao. Dark energy imprints on the kinematic Sunyaev–Zel’dovich signal. *Phys. Lett. B*, 735:402–411, 2014. doi: 10.1016/j.physletb.2014.06.066.
- S. Majumdar and J. J. Mohr. Self-calibration in cluster studies of dark energy: Combining the cluster redshift distribution, the power spectrum, and mass measurements. *Astrophys. J.*, 613(1):41–50, Sep 2004. ISSN 1538-4357. doi: 10.1086/422829. URL <http://dx.doi.org/10.1086/422829>.
- D. S. Y. Mak, E. Pierpaoli, F. Schmidt, and N. Macellari. Constraints on modified gravity from Sunyaev-Zeldovich cluster surveys. *Phys. Rev. D.*, 85(12):123513, June 2012. doi: 10.1103/PhysRevD.85.123513.
- A. Mantz, S. W. Allen, D. Rapetti, and H. Ebeling. The observed growth of massive galaxy clusters – I. Statistical methods and cosmological constraints. *Mon. Not. Roy. Astron. Soc.*, 406(3):1759–1772, 08 2010. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2010.16992.x. URL <https://doi.org/10.1111/j.1365-2966.2010.16992.x>.
- A. B. Mantz, S. W. Allen, R. G. Morris, D. A. Rapetti, D. E. Applegate, P. L. Kelly, A. von der Linden, and R. W. Schmidt. Cosmology and astrophysics from relaxed galaxy clusters – II. Cosmological constraints. *Mon. Not. Roy. Astron. Soc.*, 440(3):2077–2098, 2014a. doi: 10.1093/mnras/stu368.

- A. B. Mantz et al. Weighing the giants – IV. Cosmology and neutrino mass. *Mon. Not. Roy. Astron. Soc.*, 446(3):2205–2225, 11 2014b. ISSN 0035-8711. doi: 10.1093/mnras/stu2096. URL <https://doi.org/10.1093/mnras/stu2096>.
- F. Marinacci et al. First results from the IllustrisTNG simulations: radio haloes and magnetic fields. *Mon. Not. Roy. Astron. Soc.*, 480(4):5113–5139, 2018. doi: 10.1093/mnras/sty2206.
- J. Martin. Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask). *Comptes Rendus Physique*, 13:566–665, 2012. doi: 10.1016/j.crhy.2012.04.008.
- I. G. McCarthy, A. M. C. L. Brun, J. Schaye, and G. P. Holder. The thermal Sunyaev-Zel’dovich effect power spectrum in light of Planck. *Mon. Not. Roy. Astron. Soc.*, 440(4):3645–3657, 2014. doi: 10.1093/mnras/stu543.
- I. G. McCarthy, J. Schaye, S. Bird, and A. M. C. Le Brun. The BAHAMAS project: Calibrated hydrodynamical simulations for large-scale structure cosmology. *Mon. Not. Roy. Astron. Soc.*, 465(3):2936–2965, 2017. doi: 10.1093/mnras/stw2792.
- N. J. McConnell and C.-P. Ma. Revisiting the Scaling Relations of Black Hole Masses and Host Galaxy Properties. *Astrophys. J.*, 764(2):184, Feb. 2013. doi: 10.1088/0004-637X/764/2/184.
- A. Merloni et al. eROSITA Science Book: Mapping the Structure of the Energetic Universe. *ArXiv e-prints*, Sept. 2012.
- D. F. Mota and D. J. Shaw. Evading Equivalence Principle Violations, Cosmological and other Experimental Constraints in Scalar Field Theories with a Strong Coupling to Matter. *Phys. Rev.*, D75:063501, 2007. doi: 10.1103/PhysRevD.75.063501.
- S. Murray, C. Power, and A. Robotham. HMFcalc: An Online Tool for Calculating Dark Matter Halo Mass Functions. *ArXiv e-prints*, 6 2013.

- D. Nagai, A. Vikhlinin, and A. V. Kravtsov. Testing x-ray measurements of galaxy clusters with cosmological simulations. *The Astrophysical Journal*, 655(1):98–108, jan 2007. doi: 10.1086/509868. URL <https://doi.org/10.1086/509868>.
- J. P. Naiman et al. First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium. *Mon. Not. Roy. Astron. Soc.*, 477(1): 1206–1224, June 2018. doi: 10.1093/mnras/sty618.
- J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *Astrophys. J.*, 490:493–508, Dec. 1997. doi: 10.1086/304888.
- D. Nelson et al. First results from the IllustrisTNG simulations: the galaxy colour bimodality. *Mon. Not. Roy. Astron. Soc.*, 475(1):624–647, 2018. doi: 10.1093/mnras/stx3040.
- A. F. Neto et al. The statistics of lambda CDM Halo Concentrations. *Mon. Not. Roy. Astron. Soc.*, 381:1450–1462, 2007. doi: 10.1111/j.1365-2966.2007.12381.x.
- A. Nicolis, R. Rattazzi, and E. Trincherini. The Galileon as a local modification of gravity. *Phys. Rev.*, D79:064036, 2009. doi: 10.1103/PhysRevD.79.064036.
- P. Norberg, C. M. Baugh, E. Gaztanaga, and D. J. Croton. Statistical Analysis of Galaxy Surveys - I. Robust error estimation for 2-point clustering statistics. *Mon. Not. Roy. Astron. Soc.*, 396:19, 2009. doi: 10.1111/j.1365-2966.2009.14389.x.
- E. Paillas, M. Cautun, B. Li, Y.-C. Cai, N. Padilla, J. Armijo, and S. Bose. The Santiago–Harvard–Edinburgh–Durham void comparison II: unveiling the Vainshtein screening using weak lensing. *Mon. Not. Roy. Astron. Soc.*, 484(1):1149–1165, 01 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz022. URL <https://doi.org/10.1093/mnras/stz022>.
- R. Pakmor and V. Springel. Simulations of magnetic fields in isolated disc galaxies. *Mon. Not. Roy. Astron. Soc.*, 432(1):176–193, June 2013. doi: 10.1093/mnras/stt428.
- R. Pakmor, A. Bauer, and V. Springel. Magnetohydrodynamics on an unstructured moving grid. *Mon. Not. Roy. Astron. Soc.*, 418(2):1392–1401, Dec. 2011. doi: 10.1111/j.1365-2966.2011.19591.x.

- H. Park, M. A. Alvarez, and J. R. Bond. The Impact of Baryonic Physics on the Kinetic Sunyaev-Zel'dovich Effect. *Astrophys. J.*, 853(2):121, 2018. doi: 10.3847/1538-4357/aaa0da.
- S. Peirone, M. Raveri, M. Viel, S. Borgani, and S. Ansoldi. Constraining  $f(R)$  Gravity with Planck Sunyaev-Zel'dovich Clusters. *Phys. Rev.*, D95(2):023521, 2017. doi: 10.1103/PhysRevD.95.023521.
- A. Pillepich, T. H. Reiprich, C. Porciani, K. Borm, and A. Merloni. Forecasts on dark energy from the X-ray cluster survey with eROSITA: constraints from counts and clustering. *Mon. Not. Roy. Astron. Soc.*, 481(1):613–626, 2018a. doi: 10.1093/mnras/sty2240.
- A. Pillepich et al. First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *Mon. Not. Roy. Astron. Soc.*, 475(1):648–675, 2018b. doi: 10.1093/mnras/stx3112.
- A. Pillepich et al. Simulating Galaxy Formation with the IllustrisTNG Model. *Mon. Not. Roy. Astron. Soc.*, 473(3):4077–4106, 2018c. doi: 10.1093/mnras/stx2656.
- L. Pizzuti et al. CLASH-VLT: constraints on  $f(R)$  gravity models with galaxy clusters using lensing and kinematic analyses. *JCAP*, 1707(07):023, 2017. doi: 10.1088/1475-7516/2017/07/023.
- F. Prada, A. A. Klypin, A. J. Cuesta, J. E. Betancort-Rijo, and J. Primack. Halo concentrations in the standard LCDM cosmology. *Mon. Not. Roy. Astron. Soc.*, 423:3018–3030, 2012. doi: 10.1111/j.1365-2966.2012.21007.x.
- G. W. Pratt et al. Gas entropy in a representative sample of nearby X-ray galaxy clusters (REXCESS): relationship to gas mass fraction. *Astron. Astrophys.*, 511:A85, Feb. 2010. doi: 10.1051/0004-6361/200913309.
- D. Reed, R. Bower, C. Frenk, A. Jenkins, and T. Theuns. The halo mass function from the dark ages through the present day. *Mon. Not. Roy. Astron. Soc.*, 374:2–15, 2007. doi: 10.1111/j.1365-2966.2006.11204.x.

- C. Reichardt et al. An Improved Measurement of the Secondary Cosmic Microwave Background Anisotropies from the SPT-SZ + SPTpol Surveys. *ArXiv e-prints*, 2 2020.
- A. G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998. doi: 10.1086/300499.
- M. Roncarelli, M. Baldi, and F. Villaescusa-Navarro. The kinematic Sunyaev–Zel’dovich effect of the large-scale structure (II): the effect of modified gravity. *Mon. Not. Roy. Astron. Soc.*, 481(2):2497–2506, 2018. doi: 10.1093/mnras/sty2225.
- J. Sakstein, B. Jain, and V. Vikram. Detecting modified gravity in the stars. *Int. J. Mod. Phys.*, D23(12):1442002, 2014. doi: 10.1142/S0218271814420024.
- L. Salvati, M. Douspis, and N. Aghanim. Constraints from thermal Sunyaev-Zel’dovich cluster counts and power spectrum combined with CMB. *Astron. Astrophys.*, 614:A13, 2018. doi: 10.1051/0004-6361/201731990.
- J. Schaye et al. The EAGLE project: Simulating the evolution and assembly of galaxies and their environments. *Mon. Not. Roy. Astron. Soc.*, 446:521–554, 2015. doi: 10.1093/mnras/stu2058.
- F. Schmidt. Cosmological simulations of normal-branch braneworld gravity. *Phys. Rev. D.*, 80(12):123003, Dec. 2009. doi: 10.1103/PhysRevD.80.123003.
- F. Schmidt. Self-Consistent Cosmological Simulations of DGP Braneworld Gravity. *Phys. Rev. D*, 80:043001, 2009. doi: 10.1103/PhysRevD.80.043001.
- F. Schmidt. Dynamical Masses in Modified Gravity. *Phys. Rev.*, D81:103002, 2010. doi: 10.1103/PhysRevD.81.103002.
- F. Schmidt, A. Vikhlinin, and W. Hu. Cluster constraints on  $f(R)$  gravity. *PRD*, 80:083505, 2009. doi: 10.1103/PhysRevD.80.083505.
- F. Schmidt, W. Hu, and M. Lima. Spherical collapse and the halo model in braneworld gravity. *Phys. Rev. D*, 81(6):063005, Mar. 2010. doi: 10.1103/PhysRevD.81.063005.

- P. Schuecker, H. Böhringer, C. A. Collins, and L. Guzzo. The REFLEX galaxy cluster survey. VII.  $\Omega_m$  and  $\sigma_8$  from cluster abundance and large-scale clustering. *Astron. Astrophys.*, 398:867–877, Feb. 2003. doi: 10.1051/0004-6361:20021715.
- L. D. Shaw, D. H. Rudd, and D. Nagai. Deconstructing the Kinetic SZ Power Spectrum. *Astrophys. J.*, 756(1):15, Sept. 2012. doi: 10.1088/0004-637X/756/1/15.
- R. K. Sheth and G. Tormen. Large scale bias and the peak background split. *Mon. Not. Roy. Astron. Soc.*, 308:119, 1999. doi: 10.1046/j.1365-8711.1999.02692.x.
- R. K. Sheth and G. Tormen. An Excursion set model of hierarchical clustering : Ellipsoidal collapse and the moving barrier. *Mon. Not. Roy. Astron. Soc.*, 329:61, 2002. doi: 10.1046/j.1365-8711.2002.04950.x.
- R. K. Sheth, H. J. Mo, and G. Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *Mon. Not. Roy. Astron. Soc.*, 323:1, 2001. doi: 10.1046/j.1365-8711.2001.04006.x.
- D. Shi, B. Li, J. Han, L. Gao, and W. A. Hellwing. Exploring the liminality: properties of haloes and subhaloes in borderline  $f(R)$  gravity. *Mon. Not. Roy. Astron. Soc.*, 452(3): 3179–3191, 2015. doi: 10.1093/mnras/stv1549.
- D. Shi, B. Li, and J. Han. Environmental screening of dark matter haloes in  $f(R)$  gravity. *Mon. Not. Roy. Astron. Soc.*, 469(1):705–715, 2017. doi: 10.1093/mnras/stx865.
- J. L. Sievers et al. The Atacama Cosmology Telescope: Cosmological parameters from three seasons of data. *J. Cosmo. Astropart. Phys.*, 10:060, 2013. doi: 10.1088/1475-7516/2013/10/060.
- T. P. Sotiriou and V. Faraoni.  $f(R)$  Theories Of Gravity. *Rev. Mod. Phys.*, 82:451–497, 2010. doi: 10.1103/RevModPhys.82.451.
- D. N. Spergel et al. First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters. *Astrophys. J. Suppl.*, 148:175–194, 2003. doi: 10.1086/377226.

- V. Springel. The Cosmological simulation code GADGET-2. *Mon. Not. Roy. Astron. Soc.*, 364:1105–1134, 2005. doi: 10.1111/j.1365-2966.2005.09655.x.
- V. Springel. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *Mon. Not. Roy. Astron. Soc.*, 401:791–851, Jan. 2010. doi: 10.1111/j.1365-2966.2009.15715.x.
- V. Springel and L. Hernquist. Cosmological SPH simulations: A Hybrid multi-phase model for star formation. *Mon. Not. Roy. Astron. Soc.*, 339:289, 2003. doi: 10.1046/j.1365-8711.2003.06206.x.
- V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann. Populating a cluster of galaxies - I. Results at  $z = 0$ . *Mon. Not. Roy. Astron. Soc.*, 328(3):726–750, Dec. 2001. doi: 10.1046/j.1365-8711.2001.04912.x.
- V. Springel et al. Simulating the joint evolution of quasars, galaxies and their large-scale distribution. *Nature*, 435:629–636, 2005. doi: 10.1038/nature03597.
- V. Springel et al. The Aquarius Project: the subhalos of galactic halos. *Mon. Not. Roy. Astron. Soc.*, 391:1685–1711, 2008. doi: 10.1111/j.1365-2966.2008.14066.x.
- V. Springel et al. First results from the IllustrisTNG simulations: matter and galaxy clustering. *Mon. Not. Roy. Astron. Soc.*, 475(1):676–698, 2018. doi: 10.1093/mnras/stx3304.
- A. Stark, C. J. Miller, N. Kern, D. Gifford, G.-B. Zhao, B. Li, K. Koyama, and R. C. Nichol. Probing Theories of Gravity with Phase Space-Inferred Potentials of Galaxy Clusters. *Phys. Rev.*, D93(8):084036, 2016. doi: 10.1103/PhysRevD.93.084036.
- M. Sun, G. M. Voit, M. Donahue, C. Jones, and W. Forman. Chandra studies of the X-ray gas properties of galaxy groups. *Astrophys. J.*, 693:1142–1172, 2009. doi: 10.1088/0004-637X/693/2/1142.

- R. A. Sunyaev and I. B. Zeldovich. Microwave background radiation as a probe of the contemporary structure and history of the universe. *Annual review of astronomy and astrophysics*, 18:537–560, Jan. 1980. doi: 10.1146/annurev.aa.18.090180.002541.
- R. A. Sunyaev and Y. B. Zeldovich. The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies. *Comments on Astrophysics and Space Physics*, 4:173, Nov. 1972.
- A. Terukina, L. Lombriser, K. Yamamoto, D. Bacon, K. Koyama, and R. C. Nichol. Testing chameleon gravity with the Coma cluster. *JCAP*, 1404:013, 2014. doi: 10.1088/1475-7516/2014/04/013.
- A. Terukina, K. Yamamoto, N. Okabe, K. Matsushita, and T. Sasaki. Testing a generalized cubic Galileon gravity model with the Coma Cluster. *JCAP*, 10:064, 2015. doi: 10.1088/1475-7516/2015/10/064.
- R. Teyssier. Cosmological hydrodynamics with adaptive mesh refinement: a new high resolution code called ramses. *Astron. Astrophys.*, 385:337–364, 2002. doi: 10.1051/0004-6361:20011817.
- J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, G. Yepes, S. Gottlober, and D. E. Holz. Toward a halo mass function for precision cosmology: The Limits of universality. *Astrophys. J.*, 688:709–728, 2008. doi: 10.1086/591439.
- J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, and S. Gottlöber. The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. *Astrophys. J.*, 724:878, 2010. doi: 10.1088/0004-637X/724/2/878.
- N. Truong et al. Cosmological hydrodynamical simulations of galaxy clusters: X-ray scaling relations and their evolution. *Mon. Not. Roy. Astron. Soc.*, 474(3):4089–4111, 2018. doi: 10.1093/mnras/stx2927.
- A. Vainshtein. To the problem of nonvanishing gravitation mass. *Physics Letters B*, 39(3): 393 – 394, 1972. ISSN 0370-2693. doi: [https://doi.org/10.1016/0370-2693\(72\)90147-5](https://doi.org/10.1016/0370-2693(72)90147-5). URL <http://www.sciencedirect.com/science/article/pii/0370269372901475>.



- A. Vikhlinin et al. Chandra Cluster Cosmology Project. II. Samples and X-Ray Data Reduction. *Astrophys. J.*, 692(2):1033–1059, Feb. 2009. doi: 10.1088/0004-637X/692/2/1033.
- V. Vikram, A. Cabre, B. Jain, and J. T. VanderPlas. Astrophysical Tests of Modified Gravity: the Morphology and Kinematics of Dwarf Galaxies. *JCAP*, 1308:020, 2013. doi: 10.1088/1475-7516/2013/08/020.
- M. Vogelsberger, S. Genel, D. Sijacki, P. Torrey, V. Springel, and L. Hernquist. A model for cosmological simulations of galaxy formation physics. *Mon. Not. Roy. Astron. Soc.*, 436(4):3031–3067, Dec. 2013. doi: 10.1093/mnras/stt1789.
- F. von Braun-Bates and J. Devriendt. Universality of the halo mass function in screened gravity theories. *JCAP*, 12:028, 2018. doi: 10.1088/1475-7516/2018/12/028.
- M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro. Precision determination of the mass function of dark matter halos. *Astrophys. J.*, 646:881–885, 2006. doi: 10.1086/504962.
- D. H. Weinberg, J. S. Bullock, F. Governato, R. Kuzio de Naray, and A. H. G. Peter. Cold dark matter: controversies on small scales. *Proc. Nat. Acad. Sci.*, 112:12249–12255, 2015. doi: 10.1073/pnas.1308716112.
- R. Weinberger et al. Simulating galaxy formation with black hole driven thermal and kinetic feedback. *Mon. Not. Roy. Astron. Soc.*, 465(3):3291–3308, Mar 2017. doi: 10.1093/mnras/stw2944.
- M. C. Weisskopf, H. D. Tananbaum, L. P. van Speybroeck, and S. L. O’Dell. Chandra x-ray observatory (cxo):overview. *Proc. SPIE Int. Soc. Opt. Eng.*, 4012:2, 2000. doi: 10.1117/12.391545.
- S. D. M. White, J. F. Navarro, A. E. Evrard, and C. S. Frenk. The baryon content of galaxy clusters: a challenge to cosmological orthodoxy. *Nature*, 366(6454):429–433, Dec. 1993. doi: 10.1038/366429a0.

- H. Wilcox, R. C. Nichol, G.-B. Zhao, D. Bacon, K. Koyama, and A. K. Romer. Simulation tests of galaxy cluster constraints on chameleon gravity. *Mon. Not. Roy. Astron. Soc.*, 462(1):715–725, 2016. doi: 10.1093/mnras/stw1617.
- H. Wilcox et al. The XMM Cluster Survey: Testing chameleon gravity using the profiles of clusters. *Mon. Not. Roy. Astron. Soc.*, 452(2):1171–1183, 2015. doi: 10.1093/mnras/stv1366.
- C. M. Will. The Confrontation between General Relativity and Experiment. *Living Rev. Rel.*, 17:4, 2014. doi: 10.12942/lrr-2014-4.
- H. A. Winther et al. Modified Gravity N-body Code Comparison Project. *Mon. Not. Roy. Astron. Soc.*, 454(4):4208–4234, 2015. doi: 10.1093/mnras/stv2253.
- P.-J. Zhang, U.-L. Pen, and H. Trac. Precision era of the kinetic Sunyaev-Zeldovich effect: Simulations, analytical models and observations and the power to constrain reionization. *Mon. Not. Roy. Astron. Soc.*, 347:1224, 2004. doi: 10.1111/j.1365-2966.2004.07298.x.
- G.-B. Zhao, B. Li, and K. Koyama. N-body Simulations for  $f(R)$  Gravity using a Self-adaptive Particle-Mesh Code. *Phys. Rev.*, D83:044007, 2011a. doi: 10.1103/PhysRevD.83.044007.
- G.-B. Zhao, B. Li, and K. Koyama. Testing General Relativity using the Environmental Dependence of Dark Matter Halos. *Phys. Rev. Lett.*, 107:071303, 2011b. doi: 10.1103/PhysRevLett.107.071303.